



(86) Date de dépôt PCT/PCT Filing Date: 2004/03/29
(87) Date publication PCT/PCT Publication Date: 2005/02/10
(85) Entrée phase nationale/National Entry: 2005/09/28
(86) N° demande PCT/PCT Application No.: US 2004/009510
(87) N° publication PCT/PCT Publication No.: 2005/012502
(30) Priorité/Priority: 2003/03/28 (60/458,026) US

(51) Cl.Int.⁷/Int.Cl.⁷ G01N 33/68
(71) Demandeur/Applicant:
IDM PHARMA, INC., US
(72) Inventeurs/Inventors:
BAKER, DENISE M., US;
LIVINGSTON, BRIAN D., US;
CHESNUT, ROBERT W., US;
SETTE, ALESSANDRO, US;
NEWMAN, MARK J., US
(74) Agent: MBM & CO.

(54) Titre : PROCEDES D'IDENTIFICATION DE VARIANTS OPTIMAUX D'EPITOPES PEPTIDIQUES
(54) Title: METHODS OF IDENTIFYING OPTIMAL VARIANTS OF PEPTIDE EPITOPES

(57) **Abrégé/Abstract:**

The present invention is directed to methods for selecting a variant of a peptide epitope which induces a CTL response against another variant(s) of the peptide epitope, by determining whether the variant comprises only conserved residues, as defined herein, at non-anchor positions in comparison to the other variant(s). The present invention is also directed to variants identified by the methods above; peptides comprising such variants; nucleic acids encoding such variants and peptides; cells comprising such variants, and/or peptides, and/or nucleic acids; compositions comprising such variants, and/or peptides, and/or nucleic acids, and/or cells; as well as therapeutic and diagnostic methods for using such variants, peptides, nucleic acids, cells, and compositions.



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

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
10 February 2005 (10.02.2005)

PCT

(10) International Publication Number
WO 2005/012502 A3

(51) International Patent Classification⁷: **C06F 19/00**

(21) International Application Number:
PCT/US2004/009510

(22) International Filing Date: 29 March 2004 (29.03.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/458,026 28 March 2003 (28.03.2003) US

(71) Applicant (for all designated States except US): **EPIM-
MUNE INC.** [US/US]; 5820 Nancy Ridge Drive, San
Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BAKER, Denise,
M.** [US/US]; 13635 Janeen Place, Poway, CA 92064 (US).
LIVINGSTON, Brian, D. [US/US]; 13555 Chaco Court,
San Diego, CA 92129 (US). **CHESNUT, Robert, W.**
[US/US]; 1473 Kings Cross Drive, Cardiff-by-the-Sea, CA
92007 (US). **SETTE, Alessandro** [IT/US]; 5551 Linda
Rosa Avenue, La Jolla, CA 92037 (US). **NEWMAN,
Mark, J.** [US/US]; 6603A Santa Isabel Street, Carlsbad,
CA 92009 (US).

(74) Agents: **STEFFE, Eric, K.** et al.; Sterne, Kessler, Gold-
stein & Fox P.L.L.C., 1100 New York Avenue, N.W., Wash-
ington, DC 20005 (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, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,
ZW.

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

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted
a patent (Rule 4.17(ii)) for the following designations AE,
AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ,
CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE,
EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS,
JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA,
MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM,
PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ,
TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM,
ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD,
SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY,
KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT,
LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF,
BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN,
TD, TG)
- as to the applicant's entitlement to claim the priority of the
earlier application (Rule 4.17(iii)) for all designations
- of inventorship (Rule 4.17(iv)) for US only

Published:

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

(88) Date of publication of the international search report:

14 July 2005

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

(54) Title: METHODS OF IDENTIFYING OPTIMAL VARIANTS OF PEPTIDE EPITOPES

(57) Abstract: The present invention is directed to methods for selecting a variant of a peptide epitope which induces a CTL response against another variant(s) of the peptide epitope, by determining whether the variant comprises only conserved residues, as defined herein, at non-anchor positions in comparison to the other variant(s). The present invention is also directed to variants identified by the methods above; peptides comprising such variants; nucleic acids encoding such variants and peptides; cells comprising such variants, and/or peptides, and/or nucleic acids; compositions comprising such variants, and/or peptides, and/or nucleic acids, and/or cells; as well as therapeutic and diagnostic methods for using such variants, peptides, nucleic acids, cells, and compositions.

WO 2005/012502 A3

DEMANDE OU BREVET VOLUMINEUX

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

CECI EST LE TOME 1 DE 3
CONTENANT LES PAGES 1 À 208

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

JUMBO APPLICATIONS/PATENTS

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

THIS IS VOLUME 1 OF 3
CONTAINING PAGES 1 TO 208

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

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

METHODS OF IDENTIFYING OPTIMAL VARIANTS OF PEPTIDE EPITOPES

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] This invention relates to the field of biology. In a particular embodiment, it relates to peptides, polynucleotides, and compositions useful to monitor or elicit an immune response to selected antigens, and methods of identifying such peptides and polynucleotides.

Related Art

[0002] HLA class I molecules are expressed on the surface of almost all nucleated cells. Following intracellular processing of antigens, epitopes from the antigens are presented as a complex with the HLA class I molecules on the surface of such cells. CTL recognize the peptide-HLA class I complex, which then results in the destruction of the cell bearing the HLA-peptide complex directly by the CTL and/or via the activation of non-destructive mechanisms *e.g.*, the production of interferon, that inhibit viral replication.

[0003] **Human Immunodeficiency Virus.** Acquired immunodeficiency syndrome (AIDS) caused by infection with human immunodeficiency virus-1 (HIV-1) represents a major world health problem. Estimates indicate that about 16,000 people worldwide are infected with HIV each day.

[0004] The development of anti-viral drugs has been a major advancement in reducing viral loads in HIV infected patients. Highly active retroviral therapy (HAART) has been shown to reduce viremia to nearly undetectable levels. However, current drug therapies are not practicable as a long term solution to the HIV epidemic. HAART therapy is severely limited due to poor tolerance for the drugs and the emergence of drug-resistant virus. Moreover, replication competent HIV persists in the lymphoid tissue of patients who have responded to HAART, thus serving as a reservoir of virus. Lastly, current anti-retroviral drug therapies have little impact upon the global epidemic: almost 90% of the world's HIV infected population resides within countries lacking financial resources for these drugs. Thus, a need exists for an efficacious vaccine to both prevent and treat HIV infection.

[0005] Virus-specific, human leukocyte antigen (HLA) class I-restricted cytotoxic T lymphocytes (CTL) are known to play a major role in the prevention and clearance of virus infections in vivo (Oldstone *et al.*, *Nature* 321:239, 1989; Jamieson *et al.*, *J. Virol.* 61:3930, 1987; Yap *et al.*, *Nature* 273:238, 1978; Lukacher *et al.*, *J. Exp. Med.* 160:814, 1994; McMichael *et al.*, *N. Engl. J. Med.* 309:13, 1983; Sethi *et al.*, *J. Gen. Virol.* 64:443, 1983; Watari *et al.*, *J. Exp. Med.* 165:459, 1987; Yasukawa *et al.*, *J. Immunol.* 143:2051, 1989; Tigges *et al.*, *J. Virol.* 66:1622, 1993; Reddenhase *et al.*, *J. Virol.* 55:263, 1985; Quinnan *et al.*, *N. Engl. J. Med.* 307:6, 1982).

[0006] While immune correlates of protective immunity against HIV infection are not well defined, there is a growing body of evidence that suggests CTL are important in controlling HIV infection. HIV-specific CTL responses can be detected early in infection and the appearance of the responses corresponds to the time in infection at which initial viremia is reduced (Pantaleo *et al.*, *Nature* 370:463, 1994; Walker *et al.*, *Proc. Natl. Acad. Sci.* 86:9514, 1989). In addition, HIV replication in infected lymphocytes can be inhibited by incubation with autologous CTL (*see, e.g.*, Tsubota *et al.*, *J. Exp. Med.* 169:1421, 1989). These data are supported by recent studies that indicate CTL are required for controlling viral replication in a SIV/rhesus animal model (Schmitz *et al.*, *Science* 283:857, 1999), and additionally supported by studies that demonstrate that CTL exert selective pressure on HIV populations as evidenced by the eventual predominance of viruses with amino acid replacements in those regions of the virus to which CTL responses are directed (*see, e.g.*, Borrow *et al.*, *Nature Med.* 3:205-211, 1997; Price *et al.*, *Proc. Nat. Acad. Sci.* 94:12890-1895, 1997; Koenig *et al.*, *Nature Med.* 1:330-336, 1995; and Haas *et al.*, *J. Immunol.* 157:4212-4221, 1996).

[0007] Virus-specific T helper lymphocytes are also known to be critical for maintaining effective immunity in chronic viral infections. Historically, HTL responses were viewed as primarily supporting the expansion of specific CTL and B cell populations; however, more recent data indicate that HTL may directly contribute to the control of virus replication. For example, a decline in CD4⁺ T cells and a corresponding loss in HTL function characterize infection with HIV (Lane *et al.*, *New Engl. J. Med.* 313:79, 1985). Furthermore, studies in HIV infected patients have also shown that there is an inverse relationship between virus-specific HTL responses and viral load, suggesting that HTL play a role in viremia (*see, e.g.*, Rosenberg *et al.*, *Science* 278:1447, 1997).

[0008] A fundamental challenge in the development of an efficacious HIV vaccine is the heterogeneity observed in HIV. The virus, like some other infectious agents including retroviruses, rapidly mutates during replication resulting in the generation of virus that can escape anti-viral therapy and immune recognition (Borrow *et al.*, *Nature Med.* 3:205, 1997). In addition, HIV can be classified into a variety of subtypes that exhibit significant sequence divergence (*see, e.g.*, Lukashov *et al.*, *AIDS* 12:S43, 1998). In view of the heterogeneous nature of HIV, and the

heterogeneous immune response observed with HIV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple HIV epitopes appears to be important for the development of an efficacious vaccine against HIV. There is a need to establish such vaccine embodiments which elicit immune responses of sufficient breadth and vigor to prevent and/or clear HIV infection.

[0009] Hepatitis B Virus. Chronic infection by hepatitis B virus (HBV) affects at least 5% of the world's population and is a major cause of cirrhosis and hepatocellular carcinoma (Hoofnagle, J., *N. Engl. J. Med.* 323:337, 1990; Fields, B. and Knipe, D., In: *Fields Virology* 2:2137, 1990). The World Health Organization lists hepatitis B as a leading cause of death worldwide, close behind chronic pulmonary disease, and more prevalent than AIDS. Chronic HBV infection can range from an asymptomatic carrier state to continuous hepatocellular necrosis and inflammation, and can lead to hepatocellular carcinoma.

[0010] The immune response to HBV is believed to play an important role in controlling hepatitis B infection. A variety of humoral and cellular responses to different regions of the HBV nucleocapsid core and surface antigens have been identified. T cell mediated immunity, particularly involving class I human leukocyte antigen-restricted cytotoxic T lymphocytes (CTL), is believed to be crucial in combatting established HBV infection.

[0011] Several studies have emphasized the association between self-limiting acute hepatitis and multispecific CTL responses (Penna, A. et al., *J. Exp. Med.* 174:1565, 1991; Nayarsina, R. et al., *J. Immunol.* 150:4659, 1993). Spontaneous and interferon-related clearance of chronic HBV infection is also associated with the resurgence of a vigorous CTL response (Guidotti, L. G. et al., *Proc. Natl. Acad. Sci. USA* 91:3764, 1994). In all such cases the CTL responses are polyclonal, and specific for multiple viral proteins including the HBV envelope, core and polymerase antigens. By contrast, in patients with chronic hepatitis, the CTL activity is usually absent or weak, and antigenically restricted.

[0012] The crucial role of CTL in resolution of HBV infection has been further underscored by studies using HBV transgenic mice. Adoptive transfer of HBV-specific CTL into mice transgenic for the HBV genome resulted in suppression of virus replication. This effect was primarily mediated by a non-lytic, lymphokine-based mechanism (Guidotti, L. G. et al., *Proc. Natl. Acad. Sci. USA* 91:3764, 1994; Guidotti, L.

G., Guilhot, S., and Chisari, F. V. *J. Virol.* 68:1265, 1994; Guidotti, L. G. et al., *J. Virol.* 69:6158, 1995; Gilles, P. N., Fey, G., and Chisari, F. V., *J. Virol.* 66:3955, 1992).

[0013] As is the case for HLA class I restricted responses, HLA class II restricted T cell responses are usually detected in patients with acute hepatitis, and are absent or weak in patients with chronic infection (Chisari, F. V. and Ferrari, C., *Annu. Rev. Immunol.* 13:29, 1995). HLA Class II responses are tied to activation of helper T cells (HTLs) Helper T lymphocytes, which recognize Class II HLA molecules, may directly contribute to the clearance of HBV infection through the secretion of cytokines which suppress viral replication (Franco, A. et al., *J. Immunol.* 159:2001, 1997). However, their primary role in disease resolution is believed to be mediated by inducing activation and expansion of virus-specific CTL and B cells.

[0014] In view of the heterogeneous immune response observed with HBV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple epitopes appears to be important for the development of an efficacious vaccine against HBV. There is a need to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear HBV infection. Epitope-based vaccines appear useful.

[0015] **Hepatitis C Virus.** Hepatitis C virus (HCV) infection is a global human health problem with approximately 150,000 new reported cases each year in the U.S. alone. HCV is a single stranded RNA virus, and is the etiological agent identified in most cases of non-A, non-B post-transfusion and post-transplant hepatitis, and is a common cause of acute sporadic hepatitis (Choo *et al.*, *Science* 244:359, 1989; Kuo *et al.*, *Science* 244:362, 1989; and Alter *et al.*, in: *Current Perspective in Hepatology*, p. 83, 1989). It is estimated that more than 50% of patients infected with HCV become chronically infected and, of those, 20% develop cirrhosis of the liver within 20 years (Davis *et al.*, *New Engl. J. Med.* 321:1501, 1989; Alter *et al.*, in: *Current Perspective in Hepatology*, p. 83, 1989; Alter *et al.*, *New Engl. J. Med.* 327:1899, 1992; and Dienstag, J. L. *Gastroenterology* 85:430, 1983). Moreover, the only therapy available for treatment of HCV infection is interferon- α . Most patients are unresponsive, however, and among the responders, there is a high recurrence rate within 6-12 months of cessation of treatment (Liang *et al.*, *J. Med. Virol.* 40:69, 1993). Ribavirin, a guanosine analog with a broad spectrum activity against many RNA and DNA viruses, has been shown in clinical trials to be effective against chronic

HCV infection when used in combination with interferon- α (*see, e.g.,* Poynard *et al., Lancet* 352:1426-1432, 1998; Reichard *et al., Lancet* 351:83-87, 1998) However, the response rate is still well below 50%.

[0016] Virus-specific, human leukocyte antigen (HLA) class I-restricted cytotoxic T lymphocytes (CTL) are known to play a major role in the prevention and clearance of virus infections *in vivo* (Oldstone *et al., Nature* 321:239, 1989; Jamieson *et al., J. Virol.* 61:3930, 1987; Yap *et al., Nature* 273:238, 1978; Lukacher *et al., J. Exp. Med.* 160:814, 1994; McMichael *et al., N. Engl. J. Med.* 309:13, 1983; Sethi *et al., J. Gen. Virol.* 64:443, 1983; Watari *et al., J. Exp. Med.* 165:459, 1987; Yasukawa *et al., J. Immunol.* 143:2051, 1989; Tigges *et al., J. Virol.* 66:1622, 1993; Reddenhase *et al., J. Virol.* 55:263, 1985; Quinnan *et al., N. Engl. J. Med.* 307:6, 1982).

[0017] In view of the heterogeneous immune response observed with HCV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple HCV epitopes appears to be important for the development of an efficacious vaccine against HCV. There is a need, however, to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear HCV infection.

[0018] **Human Papillomavirus.** Human papillomavirus (HPV) is a member of the papillomaviridae, a group of small DNA viruses that infect a variety of higher vertebrates. More than 80 types of HPVs have been identified. Of these, more than 30 can infect the genital tract. Some types, generally types 6 and 11, may cause genital warts, which are typically benign and rarely develop into cancer. Other strains of HPV, "cancer-associated", or "high-risk" types, can more frequently lead to the development of cancer. The primary mode of transmission of these strains of HPV is through sexual contact.

[0019] The main manifestations of the genital warts are cauliflower-like condylomata acuminata that usually involve moist surfaces; keratotic and smooth papular warts, usually on dry surfaces; and subclinical "flat" warts, which are found on any mucosal or cutaneous surface (Handsfield, H., *Am. J. Med.* 102(5A):16-20, 1997). These warts are typically benign but are a source of inter-individual spread of the virus (Ponten, J. & Guo, Z., *Cancer Surv.* 32:201-29, 1998). At least three HPV strains associated with genital warts have been identified: type 6a (*see, e.g.,* Hofmann, K.J., *et al., Virology* 209(2):506-518,

1995), type 6b (see, *e.g.*, Hofmann *et al.*, *supra*) and type 11 (see, *e.g.*, Dartmann, K. *et al.*, *Virology* 151(1):124-130, 1986).

[0020] Cancer-associated HPVs have been linked with cancer in both men and women; they include, but are not limited to, HPV-16, HPV-18, HPV-31, HPV-45, HPV-33 and HPV-56. Other HPV strains, including types 6 and 11 as well as others, *e.g.*, HPV-5 and HPV-8, are less frequently associated with cancer. The high risk types are typically associated with the development of cervical carcinoma and premalignant lesions of the cervix in women, but are also associated with similar malignant and premalignant lesions at other anatomic sites within the lower genital or anogenital tract. These lesions include neoplasia of the vagina, vulva, perineum, the penis, and the anus. HPV infection has also been associated with respiratory tract papillomas, and rarely, cancer, as well as abnormal growth or neoplasia in other epithelial tissues. See, *e.g.* VIROLOGY, 2ND ED, Fields *et al.*, Eds. Raven Press, New York, 1990, Chapters 58 and 59, for a review of HPV association with cancer.

[0021] The HPV genome consists of three functional regions, the early region, the late region, and the "long control region". The early region gene products control viral replication, transcription and cellular transformation. They include the HPV E1 and E2 proteins, which play a role in HPV DNA replication, and the E6 and E7 oncoproteins, which are involved in the control of cellular proliferation. The late region include the genes that encode the structural proteins L1 and L2, which are the major and minor capsid proteins, respectively. The "long control region" contains such sequences as enhancer and promoter regulatory regions.

[0022] HPV expresses different proteins at different stages of the infection, for example early, as well as late, proteins. Even in latent infections, however, early proteins are often expressed and are therefore useful targets for vaccine-based therapies. For example, high-grade dysplasia and cervical squamous cell carcinoma continue to express E6 and E7, which therefore can be targeted to treat disease at both early and late stages of infection.

[0023] Treatment for HPV infection is often unsatisfactory because of persistence of virus after treatment and recurrence of clinically apparent disease is common. The treatment may require frequent visits to clinics and is not directed at elimination of the virus but at clearing warts. Because of persistence of virus after treatment, recurrence of clinically apparent disease is common.

- [0024] Thus, a need exists for an efficacious vaccine to both prevent and treat HPV infection and to treat cancer that is associated with HPV infection. Effective HPV vaccines would be a significant advance in the control of sexually transmissible infections and could also protect against clinical disease, particularly cancers such as cervical cancer. (see, e.g., Rowen, P. & Lacey, C., *Dermatologic Clinics* 16(4):835-838, 1998).
- [0025] Virus-specific, human leukocyte antigen (HLA) class I-restricted cytotoxic T lymphocytes (CTL) are known to play a major role in the prevention and clearance of virus infections *in vivo* (Oldstone *et al.*, *Nature* 321:239, 1989; Jamieson *et al.*, *J. Virol.* 61:3930, 1987; Yap *et al.*, *Nature* 273:238, 1978; Lukacher *et al.*, *J. Exp. Med.* 160:814, 1994; McMichael *et al.*, *N. Engl. J. Med.* 309:13, 1983; Sethi *et al.*, *J. Gen. Virol.* 64:443, 1983; Watari *et al.*, *J. Exp. Med.* 165:459, 1987; Yasukawa *et al.*, *J. Immunol.* 143:2051, 1989; Tigges *et al.*, *J. Virol.* 66:1622, 1993; Reddenhase *et al.*, *J. Virol.* 55:263, 1985; Quinnan *et al.*, *N. Engl. J. Med.* 307:6, 1982).
- [0026] Virus-specific T helper lymphocytes are also known to be critical for maintaining effective immunity in chronic viral infections. Historically, HTL responses were viewed as primarily supporting the expansion of specific CTL and B cell populations; however, more recent data indicate that HTL may directly contribute to the control of virus replication. For example, a decline in CD4⁺ T cells and a corresponding loss in HTL function characterize infection with HIV (Lane *et al.*, *New Engl. J. Med.* 313:79, 1985). Furthermore, studies in HIV infected patients have also shown that there is an inverse relationship between virus-specific HTL responses and viral load, suggesting that HTL plays a role in viremia (see, e.g., Rosenberg *et al.*, *Science* 278:1447, 1997).
- [0027] The development of vaccines with prophylactic and therapeutic efficacy against HPV is ongoing. Early vaccine development was hampered by the inability to culture HPV. With the introduction of cloning techniques and protein expression, however, some attempts have been made to stimulate humoral and CTL response to HPV (See, e.g., Rowen, P. & Lacey, C., *Dermatologic Clinics* 16(4):835-838 (1998)). Studies to date, however, have been inconclusive.
- [0028] Activation of T helper cells and cytotoxic lymphocytes (CTLs) in the development of vaccines has also been analyzed. Lehtinen, M., *et al.* for instance, has shown that some peptides from the E2 protein of HPV type 16 activate T helper cells and CTLs (*Biochem. Biophys. Res. Commun.* 209(2):541-6 (1995)). Similarly, Tarpey *et al.*, has shown that some peptides from HPV type 11 E7 protein can stimulate human HPV-specific CTLs *in*

vitro (*Immunology* 81:222-227 (1994)) and Borysiewicz *et al.* have reported a recombinant vaccinia virus expressing HPV 16 and HPV 17 E6 and E7 that stimulated CTL responses in at least one patient (*Lancet* 347:1347-1357, 1996).

[0029] ***Plasmodium falciparum* and Malaria.** Malaria, which is caused by infection with the parasite *Plasmodium falciparum* (PF), represents a major world health problem. Approximately 500 million people in the world are at risk from the disease, with approximately 200 million people actually harboring the parasites. An estimated 1 to 2 million deaths occur each year due to malaria. (Miller *et al.*, *Science* 234:1349, 1986).

[0030] Fatal outcomes are not confined to first infections, and constant exposure is apparently a prerequisite for maintaining immunity. Naturally acquired sterile immunity is rare, if it exists at all. Accordingly, major efforts to develop an efficacious malaria vaccine have been undertaken.

[0031] Human volunteers injected with irradiated PF sporozoites are resistant to subsequent sporozoite challenges, which demonstrates that development of a malaria vaccine is indeed immunologically feasible. Furthermore, these immune individuals developed a vigorous response, including antibodies, and cytotoxic T lymphocyte (CTL) and helper T lymphocyte (HTL) components, directed against multiple antigens. Reproducing the breadth and multiplicity of this response in a vaccine, however, is a task of large proportions. The epitope approach, as described herein, may represent a solution to this challenge, in that it allows the incorporation of various antibody, CTL and HTL epitopes, from various proteins, in a single vaccine composition.

[0032] Anti-sporozoite antibodies are by themselves, in general, not completely efficacious in clearing the infection (Egan *et al.*, *Science* 236:453, 1987). However, high concentrations of antibodies directed against the repeated region of the major B cell antigen of the sporozoite/circumsporozoite protein (CSP) have been shown to prevent liver cell infection in certain experimental models (Egan *et al.*, *Science* 236:453, 1987; Potocnjak, P. *et al.*, *Science* 207:71, 1980). The present inventors have shown that constructs encompassing CSP-repeat B cell epitopes and the optimized helper epitope PADRE™ (San Diego, CA) are highly immunogenic, and can protect *in vitro* against sporozoite invasion in both mouse and human liver cells, and protect mice *in vivo* against live sporozoite challenge (Franke *et al.*, *Vaccine* 17:1201-1205, 1999)

- [0033] PF-specific CD4⁺ T cells also have a role in malarial immunity beyond providing help for B cell and CTL responses. Experiments by Renia *et al.* (Renia, *et al.*, *Proc. Natl. Acad. Sci. USA* 88:7963, 1991) demonstrated that HTLs directed against the *Plasmodium yoelli* CS protein could in fact adoptively transfer protection against malaria.
- [0034] Considerable data implicate CTLs in protection against pre-erythrocytic-stage malaria. CD8⁺ CTLs can eliminate *Plasmodium berghei*- or *Plasmodium yoelii*-infected mouse hepatocytes from in vitro culture in a major histocompatibility complex (MHC)-restricted and antigen-restricted manner (Hoffman *et al.*, *Science* 244:1078-1081, 1989; Weiss *et al.*, *J. Exp. Med.* 171:763-773, 1990). Further, it has also been shown that the immunity that developed in mice vaccinated with irradiated sporozoites is also dependent upon the presence of CD8⁺ T cells. These T cells accumulate in inflammatory liver infiltrates subsequent to challenge. Passive transfer of circumsporozoite (CSP)-specific CTL clones as long as three hours after inoculation of sporozoites (*i.e.*, after the parasites have left the bloodstream and infected liver cells) were capable of protecting animals against infection (Romero *et al.*, *Nature* 341:323, 1989).
- [0035] It is notable that CTL-restricted responses directed against a single antigen are insufficient to protect mice with different MHC alleles, and a combination of multiple antigens was required even to protect mice from the most common laboratory strains of *Plasmodium*. These data indicate that a combination of epitopes from several antigens is necessary to elicit a protective CTL response.
- [0036] Indirect evidence that CTLs are important in protective immunity against Pf in humans has also accumulated. It has been reported that cytotoxic CD8⁺ T cells can be identified in humans immunized with PF sporozoites (Moreno, *et al.*, *Int. Immunol.* 3:997, 1991). Further, humans immunized with irradiated sporozoites or naturally exposed to malaria can generate a CTL response to the pre-erythrocytic-stage antigens, CSP, sporozoite surface protein 2 (SSP2), liver-stage antigen-1 (LSA-1), and exported protein-1 (Exp-1) (*see, e.g.* Malik *et al.*, *Proc. Natl. Acad. Sci. USA* 88, 3300-3304, 1991; Doolan *et al.*, *Int. Immunol.* 3:511-516, 1991; Hill *et al.*, *Nature* 360:434-439, 1992). Additionally, there is evidence that the polymorphism within the CSP may be the result of selection by CTLs of parasites that express variant forms (McCutchan and Water, *Immunol. Lett.* 25:23-26, 1990). This is based on the observation that the variation is nonsynonymous at the nucleotide level, thereby indicating selective pressure at the protein level. The polymorphism primarily maps to identified CTL and T helper epitopes (Doolan *et al.*, *Int.*

Immunol. 5:27-46, 1993); and CTL responses to some of the parasite variants do not cross-react (Hill *et al.*, *supra*). Finally, the MHC class I human leukocyte antigen (HLA)-Bw53 has been associated with resistance to severe malaria in The Gambia, and CTLs to a conserved epitope restricted by the HLA-Bw53 allele have been identified on *P. falciparum* LSA-1 (Hill *et al.*, *Nature* 352:595-600, 1991; Hill *et al.*, *Nature* 340:434-439, 1992). Since HLA-Bw53 is found in 15%-40% of the population of sub-Saharan Africa but in less than 1% of Caucasians and Asians, these data suggest evolutionary selection on the basis of protection against severe malaria.

[0037] Thus, antibody, and both HLA class I and class II restricted responses directed against multiple sporozoite antigens appear to be involved in generating protective immunity to malaria. Furthermore, several important antigenic epitopes against which humoral and cellular immunity is focused have already been exactly delineated.

[0038] In view of the heterogeneous immune response observed with PF infection, induction of a multi-specific cellular immune response directed simultaneously against multiple PF epitopes appears to be important for the development of an efficacious vaccine against PF. There is a need, however, to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear PF infection.

[0039] **Epitope-Based Vaccines.** The use of epitope-based vaccines has several advantages over current vaccines. The epitopes for inclusion in such a vaccine are to be selected from conserved regions of viral or tumor-associated antigens, in order to reduce the likelihood of escape mutants. The advantage of an epitope-based approach over the use of whole antigens is that there is evidence that the immune response to whole antigens is directed largely toward variable regions of the antigen, allowing for immune escape due to mutations. Furthermore, immunosuppressive epitopes that may be present in whole antigens can be avoided with the use of epitope-based vaccines.

[0040] Additionally, with an epitope-based vaccine approach, there is an ability to combine selected epitopes (CTL and HTL) and additionally to modify the composition of the epitopes, achieving, for example, enhanced immunogenicity. Accordingly, the immune response can be modulated, as appropriate, for the target disease. Similar engineering of the response is not possible with traditional approaches.

[0041] Another major benefit of epitope-based immune-stimulating vaccines is their safety. The possible pathological side effects caused by infectious agents or whole protein antigens, which might have their own intrinsic biological activity, is eliminated.

[0042] An epitope-based vaccine also provides the ability to direct and focus an immune response to multiple selected antigens from the same pathogen. Thus, patient-by-patient variability in the immune response to a particular pathogen may be alleviated by inclusion of epitopes from multiple antigens from that pathogen in a vaccine composition. A "pathogen" may be an infectious agent or a tumor associated molecule.

[0043] One of the most formidable obstacles to the development of broadly efficacious epitope-based immunotherapeutics has been the extreme polymorphism of HLA molecules. In the past, effective non-genetically biased coverage of a population has been a task of considerable complexity; such coverage has required that epitopes be used specific for HLA molecules corresponding to each individual HLA allele. Therefore, impractically large numbers of epitopes would be required in order to cover ethnically diverse populations. Recently, methods have been developed that allow the identification of epitopes that bind multiple HLA molecules. Therefore, epitope-based vaccines can be designed that contain epitopes which, either individually or in combination, bind a greater number of HLA molecules. The resulting epitope-based vaccines have a greater breadth of population coverage across one or more continents and even worldwide.

[0044] **Variation in Epitopes of Infectious Agents.** A challenge in the development of effective vaccines against infectious agents such as hepatitis B virus (HBV) (47, 60) hepatitis C virus (HCV) (61-63), human papilloma virus (HPV) (64, 65) *Plasmodium falciparum* (66), and human immunodeficiency virus (HIV-1) is the protein sequence variation associated with different isolates. This variation is the result of gene sequence mutations. When such mutations occur in regions encoding epitopes recognized by cytotoxic T-lymphocytes (CTL), they provide a mechanism for escape of the agent from immune system control.

[0045] HIV-1 represents an infectious agent with an especially high frequency of sequence variation. The sequence variation associated with HIV-1 proteins from related isolates, members of the same clades or types, as well as unrelated isolates, is well documented (1). Viral escape from CTL induced as the result of natural infection or vaccines was documented in nonhuman primate models where the mechanism behind this

escape was mutation of the primary anchor residues in dominant CTL epitopes (5-9). Viral escape from HIV-specific CTL has also been strongly implied by data obtained from HIV-1 infected individuals whose disease status change, including the transition from acute to chronic infection (10, 11), loss of stable control of viral replication and subsequent progression to AIDS (4, 12) or mother-to-child transmission (13). Thus, HIV-1 genetic and protein sequence variation represent a significant challenge to immune system-based control of viral replication, both within infected individuals and within populations.

[0046] While the public health need for a vaccine against HIV-1 is well recognized and accepted, the genetic variation of HIV-1 isolates represents a highly significant obstacle (1, 14-16). Several strategies have been proposed, some of which include:

- (1) Designing vaccines on HIV-1 types prevalent within small, well defined populations or geographical regions, such as individual countries or regions, and producing multiple different vaccines for exclusive use within these countries or regions (16).
- (2) Use of HIV-1 ancestral or consensus sequences based on HIV types present in larger target populations, such as groups of neighboring countries or continents (15, 17-19).
- (3) Incorporation of viral gene products obtained from multiple different virus isolates, representing diversely different types or clades, into a single 'multi-valent' vaccine.

[0047] Related vaccine design concepts that incorporate many of the advantages associated with the approaches described above are the use of highly conserved regions or epitopes derived from these regions as the basis of the vaccine. The logic behind this approach is that conserved regions of the viral genome are those that have been maintained through the evolution of HIV-1 because changes impact gene product function and general viral fitness. This theory is consistent with analyses of HIV-1 protein sequence data which demonstrated that CTL epitopes are concentrated in conserved regions and that regions devoid of CTL epitopes are the most variable (20). Additional support comes from published reports describing CTL responses, induced as the result of

natural infection or vaccination, that recognize viral proteins or epitopes common to viral isolates from diverse types or clades (21-26). Broad function CTL responses are also known to be correlated with slower progression to AIDS, at least for certain carefully studied populations (27, 28). Despite these reports and the clustering of CTL epitopes in conserved regions of HIV-1 gene products, amino acid sequence variation of analogous regions and epitopes from different viral isolates, both within the same type or clade and from different types, remains significant. There are currently no rules guiding the selection of conserved regions of CTL epitopes for use in vaccines other than the use of amino acid sequence identity (29).

[0048] A clear understanding of how CTL recognize pathogen infected cells has emerged over the past decade. It is now well established that small fragments of pathogen-derived proteins are generated, defined as peptide epitopes generally 8-11 amino acids in length, which bind to HLA-A, -B, or -C (human Class I Major Histocompatibility Molecules) molecules expressed on the cell surface. Sequencing of naturally processed peptides bound to HLA molecules provided a means to identify the amino acid residues required for allele-specific epitope-peptide binding (30-32). Data obtained from X-ray crystallographic analysis of HLA-epitope peptide complexes, allowed for the identification and structural characterization of 'binding pockets' within the peptide binding cleft of HLA molecules. More refined epitope anchor motif definitions were then developed using data obtained from *in vitro* peptide-MHC binding assays. It is now well known that the main anchor residues typically occur at position 2 and the carboxyl terminus of peptides 8-11 amino acids in length, thus positions 8, 9, 10 or 11 (33-40). The definition of epitope peptide binding anchor motifs is the key to most, if not all, epitope prediction methods.

[0049] Initial CTL epitope identification methods were developed using common HLA alleles, such as HLA-A2.1. Motifs defined using different HLA molecules were found to be similar and this lead to the definition of HLA supertype families (41). The biological effect of this supertype relationship was first demonstrated for HIV-1 epitopes in a study where the HLA-A3 and -A11 epitope peptide binding patterns repertoires were demonstrated to be overlapping, not only with each other but also with HLA-A31, -A33 and -A*6801 (42). This binding specificity was defined as the HLA-A3 supertype. A significant overlap in peptide binding patterns was also demonstrated amongst several serologically distant HLA-B alleles (43, 44), and multiple HLA-A2 alleles (45, 46),

resulting in the definition of the HLA-B7 and HLA-A2 supertype families. Recognition of epitopes by CTL in a supertype manner has since been demonstrated to occur naturally in infectious diseases and cancer (47-53).

[0050] While only two positions within CTL epitopes are typically characterized as the primary binding anchor positions, the amino acids that can serve as the anchor residues are more variable. The preferred and tolerated amino acids that can serve as anchor residues for the HLA-A2, -A3 and -B7 supertype families of epitopes are listed in Table 1. It is possible for analogous HIV-1 epitope peptides derived from different isolates, which differ with respect to the amino acids used as anchor residues, to bind to HLA molecules similarly. This type of variation can be as conserved since it is likely that CTL produced against one epitope would recognize the related epitope. Thus, variation limited to changes in anchor residues that result in sufficient epitope peptide binding to HLA molecules does not result in immune escape from CTL. Epitopes that contain this type of variation can be identified using the appropriately designed motif search algorithms.

[0051] The TCR of CTL has been reported to be somewhat flexible or promiscuous with respect to recognition of epitope peptides bound to HLA molecules. For HIV-1, this flexibility was demonstrated as CTL recognition of related, but slightly variable, epitopes by single clones of CTL produced following natural infection (54, 55). Similar flexibility of CTL epitope recognition was demonstrated using rhesus macaques and natural infection with SIV or immunization (56, 57). This observation is not unique to HIV-1 and SIV but rather the TCR appears to have evolved to allow promiscuous recognition of peptide epitope bound to MHC molecules (58).

[0052] Selective replacement of certain amino acids in CTL epitope peptides, amino acids thought to represent TCR contact points, is not only tolerated but can increase the recognition of the epitopes by CTL clones (59). The types of amino acid substitutions that can be incorporated, typically amino acids that are similar in chemical properties are best tolerated, and their positions, independent of primary anchor positions, within a selected number of CTL epitopes from tumor associated antigens were also defined.

[0053] For HIV-1 and other infectious agents, reproducible methods for predicting the CTL recognition of related variant epitopes that occur amongst isolates have not been developed. Nor have methods for identifying CTL epitopes that are most likely to induce broadly functional responses when used in vaccine. Thus, there exists a need to develop

such methods to overcome the challenge associated with protein sequence variation in HIV and other infectious agents.

1. Åsjö, B., et al. *AIDS* 11:A17-A36. 1997.
2. Phillips, R. E., et al. *Nature* 354:453-459. 1991.
3. Phillips, R. E., et al. *Chem. Immunol.* 56:150-164. 1993.
4. McMichael, A. J., et al. *Annu. Rev. Immunol.* 15:271-296. 1997.
5. Allen, T. M., et al. *Nature* 407:386-390. 2000.
6. Barouch, D. H., et al. *Nature* 415:335-339. 2002.
7. Evans, D. T., et al. *Nat. Med.* 5:1270-1276. 1999.
8. Vogel, T. U., et al. *J. Virol.* 76:11623-11636. 2002.
9. Nacsa, J., et al. *Virology* 305:210-218. 2003.
10. Borrow, P., et al. *Nat. Med.* 3:205-211. 1997.
11. Price, D. A., et al. *Proc. Natl. Acad. Sci. U.S.A* 94:1890-1895. 1997.
12. Goulder, P. J., et al. *Nat. Med.* 3:212-217. 1997.
13. Goulder, P. J., et al. *Immunol. Lett.* 79:109-116. 2001.
14. Walker, B. D. et al. *Nat. Immunol.* 2:473-475. 2001.
15. Korber, B., et al. *Br. Med. Bull.* 58:19-42. 2001.
16. Gaschen, B., et al. *Science* 296:2354-2360. 2002.
17. Novitsky, V., et al. *J. Virol.* 76:5435-5451. 2002.
18. Davis, N. L., et al. *IUBMB. Life* 53:209-211. 2002.
19. Ellenberger, D. L., et al. *Virology* 302:155-163. 2002.
20. Yusim, K., et al. *J. Virol.* 76:8757-8768. 2002.

21. Ferrari, G., et al. *Proc. Natl. Acad. Sci. U.S.A.* 94:1396-1401. 1997.
22. Fukada, K., et al. *AIDS* 16:701-711. 2002.
23. Ferrari, G., et al. *AIDS Res. Hum. Retroviruses* 16:1433-1443. 2000.
24. Buseyne, F., et al. *Virology* 250:316-324. 1998.
25. Lynch, J. A., et al. *J. Infect. Dis.* 178:1040-1046. 1998.
26. Wilson, S. E., et al. *AIDS Res. Hum. Retroviruses* 14:925-937. 1998.
27. Rowland-Jones, S. L., et al. *Immunol. Lett.* 66:9-14. 1999.
28. Gillespie, G. M., et al. *AIDS* 16:961-972. 2002.
29. Altfeld, M. A., et al. *J. Virol.* 75:1301-1311. 2001.
30. Sette, A. et al. *Curr. Opin. Immunol.* 4:79-86. 1992.
31. Sinigaglia, F. et al. *Curr. Opin. Immunol.* 6:52-56. 1994.
32. Engelhard, V. H. *Curr. Opin. Immunol.* 6:13-23. 1994.
33. Brown, K., et al. *Nature* 364:33-39. 1994.
34. Guo, H. C., et al. *Proc. Natl. Acad. Sci. U.S.A.* 90:8053-8057. 1993.
35. Guo, H. C., et al. *Nature* 360:364-366. 1992.
36. Silver, M. L., et al. *Nature* 360:367-369. 1992.
37. Matsumura, M., et al. *Science* 257:927-934. 1992.
38. Madden, D. R., et al. *Cell* 70:1035-1048. 1995.
39. Fremont, D. H., et al. *Science* 257:919-927. 1992.
40. Sapp, M., et al. *J. Mol. Biol.* 219:277-319. 1991.
41. Kubo, R. T., et al. *J. Immunol.* 152:3913-24. 1994.

42. Sidney, J., et al. *Hum. Immunol.* 45:79-93. 1996.
43. Sidney, J., et al. *J. Immunol.* 154:247-259. 1995.
44. Sidney, J., et al. *J. Immunol.* 157:3480-3490. 1996.
45. Del Guercio, M.-F., et al. *J. Immunol.* 154:685-693. 1995.
46. Fruci, D., et al. *Hum. Immunol.* 38:187-192. 1993.
47. Bertoni, R., et al. *J. Clin. Invest.* 100:503-513. 1997.
48. Threlkeld, S. C., et al. *J. Immunol.* 159:1648-1657.
49. Khanna, R., et al. *J. Virol.* 71:7429-7435. 1997.
50. Bertoletti, A., et al. *Hepatology* 26:1027-1034. 1997.
51. Fleischhauer, K., et al. *J. Immunol.* 157:787-297. 1996.
52. Kawashima, I., et al. *Hum. Immunol.* 59:1-14. 1998.
53. Wang, R. F., et al. *J. Immunol.* 160:890-897. 1998.
54. Tomiyama, H., et al. *J. Immunol.* 30:2521-2530. 2000.
55. Buseyne, F., et al. *Int. Immunol.* 13:941-950. 2001.
56. Shen, L., et al. *J. Immunol.* 153:12-15-1994. 2000.
57. Charini, W. A., et al. *J. Immunol.* 167:4996-5003. 2001.
58. Mason, D. *Immunol. Today.* 19:395-404. 1998.
59. Tangri, S., et al. *J. Exp. Med.* 194:833-846. 2001.
60. Rehmann, B., et al. *J. Exp. Med.* 181:1047-1058. 1995.
61. Chang, K. M., et al. *J. Immunol.* 162:1156-1164. 1999.
62. Chang, K. M., et al. *J. Clin. Invest.* 100:2376-2385. 1997.

63. Propato, A., et al. *Hum. Immunol.* 62:561-576. 2001.
64. Kast, W. M., et al. *J. Immunol.* 152:3904-3912. 1994.
65. Rensing, M. E., et al. *J. Immunol.* 154:5934-5943. 1995.
66. Doolan, D. L., et al. *Immunity.* 7:97-112. 1997.

SUMMARY OF THE INVENTION

[0054] The present invention is directed to methods for selecting a variant of a peptide epitope which induces a CTL response against another variant(s) of the peptide epitope, by determining whether the variant comprises only conserved residues, as defined herein, at non-anchor positions in comparison to the other variant(s).

[0055] In some embodiments, antigen sequences from a population of an infectious agent, said antigens comprising variants of a peptide epitope, are optionally aligned (manually or by computer) along their length, preferably their full length. Variant(s) of a peptide epitope (preferably naturally occurring variants), each 8-11 amino acids in length and comprising the same MHC class I supermotif or motif, are identified manually or with the aid of a computer. In some embodiments, a variant is optionally chosen which comprises preferred anchor residues of said motif and/or which occurs with high frequency within the population of variants. In other embodiments, a variant is randomly chosen. The randomly or otherwise chosen variant is compared to from one to all the remaining variant(s) to determine whether it comprises only conserved residues in the non-anchor positions relative to from one to all the remaining variant(s).

[0056] The present invention is also directed to variants identified by the methods above; peptides comprising such variants; nucleic acids encoding such variants and peptides; cells comprising such variants, and/or peptides, and/or nucleic acids; compositions comprising such variants, and/or peptides, and/or nucleic acids, and/or cells; as well as therapeutic and diagnostic methods for using such variants, peptides, nucleic acids, cells, and compositions.

BRIEF DESCRIPTION OF THE DRAWINGS/FIGURES

[0057] FIGS. 1A-1E. Recognition of variant peptides by CTL generated against a single epitope. Variant peptides were identified from 167 HIV strains for 5 HIV epitopes, 3 HLA-A2 restricted (Env 134, A, Gag 386, B, and Vpr 62, C) and 2 HLA-A11 restricted (Pol 98, D, and Env 47, E). These are listed according to their relationship to a previously determined parent (P) into single anchor substitutions (A), single non-anchor substitutions (NA) or multiple substitutions (M). Binding of each variant peptide is also shown. The number of viral sequences containing each variant peptide is shown in the column labeled # Isolates, and is reported for the total sequences, Clade B sequences (B), and Clade C sequences (C). Finally, the ability of CTL primed against the parent peptide to recognize the variant peptides is shown in the bar graphs.

[0058] FIGS. 2A-2C. Characterization of the peptide-specific T cell lines. A. FACS analysis of the TCRs expressed by peptide -stimulated cells after 0, 1, and 5 peptide stimulations, using a panel of commercially available mAb for mouse TCR 2-14. B-C. Peptide affinity. Parent and variant peptides were titrated against CTL that had been stimulated 5 times with the parent peptide.

[0059] FIGS. 2A-2B. Recognition of a panel of variant peptides by PBL from an HIV-infected individual.

[0060] FIG 4. Prediction of immunological conservation. Gag 271 variants and their binding are shown, along with the number of isolates that express each variant. Immunological recognition was predicted for each variant based on two different choices in the immunizing peptide. On the right, the immunogenicity for each variant is shown.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

[0061] The invention can be better understood with reference to the following definitions:

[0062] An "antigen" refers to a polypeptide encoded by the genome of an infectious agent, or other another source, but preferably an infectious agent in the present invention.

Examples of HIV antigens include Env, Gag, Nef, Pol, Tat, Rev, Vif, Vpr, Vpu, p17, p24, p2, p7, p1, p6, Protease, RT, Integrase, and gp160 (preferably Env, Gag, Nef, Pol, Tat, Rev, Vif, Vpr, Vpu). Examples of HBV antigens include Core, Env, and Pol. Examples of HCV antigens include Core, E1, E2, Ns1, Ns2, Ns3, Ns4, and Ns5. Examples of HPV antigens include E1, E2, E3, E4, E5, E6, E7, L1, and L2. Examples of *Plasmodium falciparum* antigens include CSP, SSP2, Exp1, and LSA1.

[0063] Throughout this disclosure, "binding data" results are often expressed in terms of "IC₅₀'s." IC₅₀ is the concentration of peptide in a binding assay at which 50% inhibition of binding of a reference peptide is observed. Given the conditions in which the assays are run (*i.e.*, limiting HLA proteins and labeled peptide concentrations), these values approximate K_D values. Assays for determining binding are described in detail, *e.g.*, in PCT publications WO 94/20127 and WO 94/03205, and other publications such as Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998); Sidney, *et al.*, *J. Immunol.* 154:247 (1995); and Sette, *et al.*, *Mol. Immunol.* 31:813 (1994). It should be noted that IC₅₀ values can change, often dramatically, if the assay conditions are varied, and depending on the particular reagents used (*e.g.*, HLA preparation, *etc.*). For example, excessive concentrations of HLA molecules will increase the apparent measured IC₅₀ of a given ligand.

[0064] Alternatively, binding is expressed relative to a reference peptide. Although as a particular assay becomes more, or less, sensitive, the IC₅₀'s of the peptides tested may change somewhat, the binding relative to the reference peptide will not significantly change. For example, in an assay run under conditions such that the IC₅₀ of the reference peptide increases 10-fold, the IC₅₀ values of the test peptides will also shift approximately 10-fold. Therefore, to avoid ambiguities, the assessment of whether a peptide is a good (*i.e.* high), intermediate, weak, or negative binder is generally based on its IC₅₀, relative to the IC₅₀ of a standard peptide. The Tables included in this application present binding data in a preferred biologically relevant form of IC₅₀ nM.

[0065] Binding may also be determined using other assay systems including those using: live cells (*e.g.*, Ceppellini *et al.*, *Nature* 339:392 (1989); Christnick *et al.*, *Nature* 352:67 (1991); Busch *et al.*, *Int. Immunol.* 2:443 (1990); Hill *et al.*, *J. Immunol.* 147:189 (1991); del Guercio *et al.*, *J. Immunol.* 154:685 (1995)), cell free systems using detergent lysates (*e.g.*, Cerundolo *et al.*, *J. Immunol.* 21:2069 (1991)), immobilized purified MHC (*e.g.*, Hill *et al.*, *J. Immunol.* 152, 2890 (1994); Marshall *et al.*, *J. Immunol.* 152:4946 (1994)), ELISA systems (*e.g.*, Reay *et al.*, *EMBO J.* 11:2829 (1992)), surface plasmon resonance (*e.g.*, Khilko *et al.*, *J. Biol. Chem.* 268:15425 (1993)); high flux soluble phase assays (Hammer *et al.*, *J. Exp. Med.* 180:2353 (1994)), and measurement of class I MHC stabilization or assembly (*e.g.*, Ljunggren *et al.*, *Nature* 346:476 (1990));

Schumacher *et al.*, *Cell* 62:563 (1990); Townsend *et al.*, *Cell* 62:285 (1990); Parker *et al.*, *J. Immunol.* 149:1896 (1992)).

[0066] As used herein, “high affinity” with respect to HLA class I molecules is defined as binding with an IC_{50} or K_D value, of 50 nM or less, “intermediate affinity” is binding with an IC_{50} or K_D value of between 50 and about 500 nM, weak affinity is binding with an IC_{50} or K_D value of between about 500 and about 5000 nM. “High affinity” with respect to binding to HLA class II molecules is defined as binding with an IC_{50} or K_D value of 100 nM or less; “intermediate affinity” is binding with an IC_{50} or K_D value of between about 100 and about 1000 nM.

[0067] A “computer” or “computer system” generally includes: a processor and related computer programs; at least one information storage/retrieval apparatus such as a hard drive, a disk drive or a tape drive; at least one input apparatus such as a keyboard, a mouse, a touch screen, or a microphone; and display structure, such as a screen or a printer. Additionally, the computer may include a communication channel in communication with a network. Such a computer may include more or less than what is listed above.

[0068] “Cross-reactive binding” indicates that a peptide is bound by more than one HLA molecule; a synonym is degenerate binding.

[0069] A “cryptic epitope” elicits a response by immunization with an isolated peptide, but the response is not cross-reactive *in vitro* when intact whole protein, which comprises the epitope, is used as an antigen.

[0070] The term “derived” when used to discuss an epitope is a synonym for “prepared.” A derived epitope can be isolated from a natural source, or it can be synthesized in accordance with standard protocols in the art. Synthetic epitopes can comprise artificial amino acids “amino acid mimetics,” such as D isomers of natural occurring L amino acids or non-natural amino acids such as cyclohexylalanine. A derived/prepared epitope can be an analog of a native epitope.

[0071] A “diluent” includes sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred diluent for pharmaceutical compositions. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as diluents, particularly for injectable solutions.

[0072] A “dominant epitope” is an epitope that induces an immune response upon immunization with a whole native antigen (see, *e.g.*, Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993). Such a response is cross-reactive *in vitro* with an isolated peptide epitope.

[0073] An “epitope” is the collective features of a molecule, such as primary, secondary and tertiary peptide structure, and charge, that together form a site recognized by an immunoglobulin, T cell receptor or HLA molecule. Alternatively, an epitope can be defined as a set of amino acid residues which is involved in recognition by a particular immunoglobulin, or in the context of T

cells, those residues necessary for recognition by T cell receptor proteins and/or Major Histocompatibility Complex (MHC) receptors. Epitopes are present in nature, and can be isolated, purified or otherwise prepared/derived by humans. For example, epitopes can be prepared by isolation from a natural source, or they can be synthesized in accordance with standard protocols in the art. Synthetic epitopes can comprise artificial amino acids, "amino acid mimetics," such as D isomers of naturally-occurring L amino acids or non-naturally-occurring amino acids such as cyclohexylalanine. Throughout this disclosure, epitopes may be referred to in some cases as peptides. The variants of the invention are set forth in Tables 6-9 and Figures 1A-4.

[0074] It is to be appreciated that proteins or peptides that comprise a variant of the invention as well as additional amino acid(s) are still within the bounds of the invention. In certain embodiments, the peptide comprises a fragment of an antigen. A "fragment of an antigen" or "antigenic fragment" or simply "fragment" is a portion of an antigen which has 100% identity with a wild type antigen or naturally-occurring variant thereof. The fragment may or may not comprise an epitope of the invention. The fragment may be less than or equal to 600 amino acids, less than or equal to 500 amino acids, less than or equal to 400 amino acids, less than or equal to 250 amino acids, less than or equal to 100 amino acids, less than or equal to 85 amino acids, less than or equal to 75 amino acids, less than or equal to 65 amino acids, or less than or equal to 50 amino acids in length. In certain embodiments, a fragment is e.g., less than 101 or less than 51 amino acids in length, in any increment down to 5 amino acids in length. For example, the fragment may be 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 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, or 100 amino acids in length.

[0075] In certain embodiments, there is a limitation on the length of a peptide of the invention. The embodiment that is length-limited occurs when the protein/peptide comprising an epitope of the invention comprises a region (i.e., a contiguous series of amino acids) having 100% identity with a native sequence. In order to avoid the definition of epitope from reading, e.g., on whole natural molecules, there is a limitation on the length of any region that has 100% identity with a native peptide sequence. Thus, for a peptide comprising an epitope of the invention and a region with 100% identity with a native peptide sequence, the region with 100% identity to a native sequence generally has a length of: less than or equal to 600 amino acids, often less than or equal to 500 amino acids, often less than or equal to 400 amino acids, often less than or equal to 250 amino acids, often less than or equal to 100 amino acids, often less than or equal to 85 amino acids, often less than or equal to 75 amino acids, often less than or equal to 65 amino acids, and often less than or equal to 50 amino acids. In certain embodiments, an "epitope" of the invention

is comprised by a peptide having a region with less than 51 amino acids that has 100% identity to a native peptide sequence, in any increment down to 5 amino acids.

[0076] Accordingly, peptide or protein sequences longer than 600 amino acids are within the scope of the invention, so long as they do not comprise any contiguous sequence of more than 600 amino acids that have 100% identity with a native peptide sequence. For any peptide that has five contiguous residues or less that correspond to a native sequence, there is no limitation on the maximal length of that peptide in order to fall within the scope of the invention. It is presently preferred that a peptide of the invention (*e.g.*, a peptide comprising an epitope of the invention) be less than 600 residues long in any increment down to eight amino acid residues.

[0077] A peptide epitope occurring with "high frequency" is one that occurs in at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the infectious agents in a population. A "high frequency" peptide epitope is one of the more common in a population, preferably the first most common, second most common, third most common, or fourth most common in a population of variant peptide epitopes.

[0078] "Human Leukocyte Antigen" or "HLA" is a human class I or class II Major Histocompatibility Complex (MHC) protein (*see, e.g.*, Stites, *et al.*, IMMUNOLOGY, 8TH ED., Lange Publishing, Los Altos, CA (1994)).

[0079] An "HLA supertype or HLA family", as used herein, describes sets of HLA molecules grouped on the basis of shared peptide-binding specificities. HLA class I molecules that share somewhat similar binding affinity for peptides bearing certain amino acid motifs are grouped into such HLA superotypes. The terms HLA superfamily, HLA supertype family, HLA family, and HLA xx-like molecules (where "xx" denotes a particular HLA type), are synonyms. See Tables 1-4.

[0080] As used herein, "high affinity" with respect to HLA class I molecules is defined as binding with an IC₅₀, or K_D value, of 50 nM or less; "intermediate affinity" is binding with an IC₅₀ or K_D value of between about 50 and about 500 nM; "weak affinity" is binding with an IC₅₀ or K_D value between about 500 and about 5000 nM. "High affinity" with respect to binding to HLA class II molecules is defined as binding with an IC₅₀ or K_D value of 100 nM or less; "intermediate affinity" is binding with an IC₅₀ or K_D value of between about 100 and about 1000 nM. See "binding data."

[0081] An "IC₅₀" is the concentration of peptide in a binding assay at which 50% inhibition of binding of a reference peptide is observed. Given the conditions in which the assays are run (*i.e.*, limiting HLA proteins and labeled peptide concentrations), these values approximate K_D values. See "binding data."

[0082] The terms "identical" or percent "identity," in the context of two or more peptide sequences or antigen fragments, refer to two or more sequences or subsequences that are the same

or have a specified percentage of amino acid residues that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using a sequence comparison algorithm or by manual alignment and visual inspection.

[0083] An "immunogenic" peptide or an "immunogenic" epitope or "peptide epitope" is a peptide that comprises an allele-specific motif or supermotif such that the peptide will bind an HLA molecule and induce a CTL and/or HTL response. Thus, immunogenic peptides of the invention are capable of binding to an appropriate HLA molecule and thereafter inducing a cytotoxic T lymphocyte (CTL) response, or a helper T lymphocyte (HTL) response, to the peptide.

[0084] An "infectious agent" refers to a disease-causing microorganism, including viruses, bacteria, fungi, and protozoa against which a cellular immune response, preferably a CTL response, plays a role in acquired immunity. Examples of infectious agents include viruses such as human immunodeficiency virus (HIV), hepatitis B virus (HBV), hepatitis C virus (HCV), human papillomavirus (HPV), Influenza virus, Dengue virus, Epstein-Barr virus, bacteria such as *Mycobacterium tuberculosis* and *Chlamydia*, fungi such as *Candida albicans*, *Cryptococcus neoformans*, *Coccidioides spp.*, *Histoplasma spp.*, and *Aspergillus fumigatis*, protozoa such as *Plasmodium spp.*, including *P. falciparum*, *Trypanosoma spp.*, *Schistosoma spp.*, *Leishmania spp.* and the like. Preferred infectious agents include HIV, HBV, HCV, HPV, Epstein-Barr virus, *Plasmodium falciparum*, Influenza virus and Dengue virus.

[0085] The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany the material as it is found in its native state. Thus, isolated peptides in accordance with the invention preferably do not contain materials normally associated with the peptides in their *in situ* environment. An "isolated" epitope refers to an epitope that does not include the whole sequence of the antigen or polypeptide from which the epitope was derived. Typically the "isolated" epitope does not have attached thereto additional amino acids that result in a sequence that has 100% identity with a native sequence. The native sequence can be a sequence such as a tumor-associated antigen from which the epitope is derived. Thus, the term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or peptide present in a living animal is not isolated, but the same polynucleotide or peptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such a polynucleotide could be part of a vector, and/or such a polynucleotide or peptide could be part of a composition, and still be "isolated" in that such vector or composition is not part of its natural environment. Isolated RNA molecules include *in vivo* or *in vitro* RNA

transcripts of the DNA molecules of the present invention, and further include such molecules produced synthetically.

[0086] "Major Histocompatibility Complex" or "MHC" is a cluster of genes that plays a role in control of the cellular interactions responsible for physiologic immune responses. In humans, the MHC complex is also known as the human leukocyte antigen (HLA) complex. For a detailed description of the MHC and HLA complexes, see, Paul, FUNDAMENTAL IMMUNOLOGY, 3RD ED., Raven Press, New York (1993).

[0087] The term "motif" refers to a pattern of residues in an amino acid sequence of defined length, preferably a peptide of less than about 15 amino acids in length, or less than about 13 amino acids in length, usually from about 8 to about 13 amino acids (*e.g.*, 8, 9, 10, 11, 12, or 13) for a class I HLA motif and from about 6 to about 25 amino acids (*e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25) for a class II HLA motif, which is recognized by a particular HLA molecule. Motifs are typically different for each HLA protein encoded by a given human HLA allele. These motifs often differ in their pattern of the primary and secondary anchor residues. See Tables 1-3.

[0088] A "native" or a "wild type" sequence refers to a sequence found in nature.

[0089] A "negative binding residue" or "deleterious residue" is an amino acid which, if present at certain positions (typically not primary anchor positions) in a peptide epitope, results in decreased binding affinity of the peptide for the peptide's corresponding HLA molecule.

[0090] The term "peptide" is used interchangeably with "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to the other, typically by peptide bonds between the α -amino and carboxyl groups of adjacent amino acids.

[0091] A "PanDR binding" peptide or "PADRE[®]" peptide (Epimmune, San Diego, CA) is a member of a family of molecules that binds more than one HLA class II DR molecule. The pattern that defines the PADRE[®] family of molecules can be referred to as an HLA Class II supermotif. A PADRE[®] molecule binds to HLA-DR molecules and stimulates *in vitro* and *in vivo* human helper T lymphocyte (HTL) responses. For a further definition of the PADRE[®] family, see copending application US serial Nos. 09/709,774, filed November 11, 2000; and 09/707,738, filed November 6, 2000; PCT publication Nos WO 95/07707, and WO 97/26784; U.S. Patent Nos. 5,736,142 issued April 7, 1998; 5,679,640, issued October 21, 1997; and 6,413,935, issued July 2, 2002.

[0092] "Pharmaceutically acceptable" refers to a generally non-toxic, inert, and/or physiologically compatible composition or component of a composition.

- [0093] A "pharmaceutical excipient" or "excipient" comprises a material such as an adjuvant, a carrier, pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like. A "pharmaceutical excipient" is an excipient which is pharmaceutically acceptable.
- [0094] A "primary anchor residue" is an amino acid at a specific position along a peptide sequence which is understood to provide a contact point between the immunogenic peptide and the HLA molecule. One, two or three, primary anchor residues within a peptide of defined length generally defines a "motif" for an immunogenic peptide. These residues are understood to fit in close contact with peptide binding grooves of an HLA molecule, with their side chains buried in specific pockets of the binding grooves themselves. In one embodiment of an HLA class I motif, the primary anchor residues are located at position 2 (from the amino terminal position) and at the carboxyl terminal position of a peptide epitope in accordance with the invention. The primary anchor positions for each motif and supermotif of HLA Class I are set forth in Tables 1-2. For example, analog peptides can be created by altering the presence or absence of particular residues in these anchor positions. Such analogs are used to modulate the binding affinity of an epitope comprising a particular motif or supermotif. A "preferred primary anchor residue" is an anchor residue of a motif or supermotif that is associated with optimal binding. Preferred primary anchor residues are indicated in bold-face in Tables 1-2. A "tolerated primary anchor residue" is an anchor residue of a motif or supermotif that is associated with binding to a lesser extent than a preferred residue. Tolerated primary anchor residues are indicated in italicized text in Tables 1-2.
- [0095] "Promiscuous recognition" by a TCR is where a distinct peptide is recognized by the various T cell clones in the context of various HLA molecules. Promiscuous binding by an HLA molecule is synonymous with cross-reactive binding.
- [0096] A "protective immune response" or "therapeutic immune response" refers to a CTL and/or an HTL response to an antigen derived from an antigen of an infectious agent, which in some way prevents or at least partially arrests disease symptoms, side effects or progression. The immune response may also include an antibody response which has been facilitated by the stimulation of helper T cells.
- [0097] By "ranking" the variants in a population of peptide epitopes is meant ordering each variant by its frequency of occurrence relative to the other variants.
- [0098] The term "residue" refers to an amino acid or amino acid mimetic incorporated into a peptide or protein by an amide bond or amide bond mimetic.
- [0099] A "secondary anchor residue" is an amino acid at a position other than a primary anchor position in a peptide which may influence peptide binding. A secondary anchor residue occurs at a significantly higher frequency amongst HLA-bound peptides than would be expected by random

distribution of amino acids at a given position. A secondary anchor residue can be identified as a residue which is present at a higher frequency among high or intermediate affinity binding peptides, or a residue otherwise associated with high or intermediate affinity binding. The secondary anchor residues are said to occur at "secondary anchor positions." For example, analog peptides can be created by altering the presence or absence of particular residues in these secondary anchor positions. Such analogs are used to finely modulate the binding affinity of an epitope comprising a particular motif or supermotif. The terminology "fixed peptide" is generally used to refer to an analog peptide that has changes in primary anchor position; not secondary.

[00100] A "subdominant epitope" is an epitope which evokes little or no response upon immunization with a whole antigen or a fragment of the whole antigen comprising a subdominant epitope and a dominant epitope, which comprise the epitope, but for which a response can be obtained by immunization with an isolated peptide, and this response (unlike the case of cryptic epitopes) is detected when whole antigen or a fragment of the whole antigen comprising a subdominant epitope and a dominant epitope is used to recall the response *in vitro* or *in vivo*.

[00101] A "supermotif" is a peptide binding specificity shared by HLA molecules encoded by two or more HLA alleles. Preferably, a supermotif-bearing peptide is recognized with high or intermediate affinity (as defined herein) by two or more HLA antigens.

[00102] "Synthetic peptide" refers to a peptide that is obtained from a non-natural source, *e.g.*, is man-made. Such peptides may be produced using such methods as chemical synthesis or recombinant DNA technology. "Synthetic peptides" include "fusion proteins."

[00103] As used herein, a "vaccine" is a composition used for vaccination, *e.g.*, for prophylaxis or therapy, that comprises one or more peptides of the invention. There are numerous embodiments of vaccines in accordance with the invention, such as by a cocktail of one or more peptides; one or more peptides of the invention comprised by a polyepitopic peptide; or nucleic acids that encode such peptides or polypeptides, *e.g.*, a minigene that encodes a polyepitopic peptide. The "one or more peptides" can include any whole unit integer from 1-150, *e.g.*, at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 or more peptides of the invention. The peptides or polypeptides can optionally be modified, such as by lipidation, addition of targeting or other sequences. HLA class I-binding peptides of the invention can be linked to HLA class II-binding peptides, *e.g.*, a PADRE[®] universal HTL-binding peptide, to facilitate activation of both cytotoxic T lymphocytes and helper T lymphocytes. Vaccines can comprise peptide pulsed antigen presenting cells, *e.g.*, dendritic cells.

[00104] A "variant of a peptide epitope" refers to a peptide that is identified from a different viral strain at the same position in an aligned sequence, and that varies by one or

more amino acids from the parent peptide epitope. Examples of peptide epitope variants include those shown in Tables 6-9 and Figures 1A-4. A "variant of an antigen" refers to an antigen that comprises at least one variant of a peptide epitope. Examples of antigen variants include those listed by sequence and/or accession number in Tables 10-22. A "variant of an infectious agent" refers to an infectious agent whose genome encodes at least one variant of an antigen. Variants of infectious agents are related viral, bacterial, fungal, or protozoan strains or isolates that vary in sequence but cause the same disease symptoms. Examples of infectious agent variants include HIV Clade A, B, and C subtypes, HBV subtypes adr, ayr, adw, and ayw, HCV types 1, 2, 3, 4, 5, and 6, HPV strains 1-92 (preferably strains 16, 18, 31, 33, 45, 52, 56, and 58) (see Table 10, listing accession numbers for the complete genome sequences of 167 HIV variants; Table 22, showing an alignment of the complete polyprotein sequences of 50 HCV variants) (see also, *Human Retroviruses and AIDS 2000: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences*, Kuiken CL, et al., Eds. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM).

[00105] The nomenclature used to describe peptides/proteins follows the conventional practice wherein the amino group is presented to the left (the N-terminus) and the carboxyl group to the right (the C-terminus) of each amino acid residue. When amino acid residue positions are referred to in a peptide epitope they are numbered in an amino to carboxyl direction with position one being the position closest to the amino terminal end of the epitope, or the peptide or protein of which it may be a part. In the formulae representing selected specific embodiments of the present invention, the amino- and carboxyl-terminal groups, although not specifically shown, are in the form they would assume at physiologic pH values, unless otherwise specified. In the amino acid structure formulae, each residue is generally represented by standard three letter or single letter designations. The L-form of an amino acid residue is represented by a capital single letter or a capital first letter of a three-letter symbol, and the D-form for those amino acids having D-forms is represented by a lower case single letter or a lower case three letter symbol. However, when three letter symbols or full names are used without capitals, they may refer to L amino acids. Glycine has no asymmetric carbon atom and is simply referred to as "Gly" or "G". The amino acid sequences of peptides set forth herein are generally designated using the standard single letter symbol. (A, Alanine; C, Cysteine; D, Aspartic Acid; E, Glutamic Acid; F, Phenylalanine; G, Glycine; H, Histidine; I, Isoleucine; K, Lysine; L, Leucine; M, Methionine; N, Asparagine; P, Proline; Q, Glutamine; R, Arginine; S, Serine; T, Threonine; V, Valine; W, Tryptophan; and Y, Tyrosine.) In addition to these symbols, "B" in the single letter abbreviations used herein

designates α -amino butyric acid. In some embodiments, α -amino butyric acid may be replaced with cysteine.

Acronyms used herein are as follows:

APC:	Antigen presenting cell
CD3:	Pan T cell marker
CD4:	Helper T lymphocyte marker
CD8:	Cytotoxic T lymphocyte marker
CEA:	Carcinoembryonic antigen (see, e.g., SEQ ID NO: 363)
CTL:	Cytotoxic T lymphocyte
DC:	Dendritic cells. DC functioned as potent antigen presenting cells by stimulating cytokine release from CTL lines that were specific for a model peptide derived from hepatitis B virus. <i>In vivo</i> experiments using DC pulsed <i>ex vivo</i> with an HBV peptide epitope have stimulated CTL immune responses <i>in vivo</i> following delivery to naïve mice.
DLT:	Dose-limiting toxicity, an adverse event related to therapy.
DMSO:	Dimethylsulfoxide
ELISA:	Enzyme-linked immunosorbant assay
E:T:	Effector:Target ratio
G-CSF:	Granulocyte colony-stimulating factor
GM-CSF:	Granulocyte-macrophage (monocyte)-colony stimulating factor
HBV:	Hepatitis B virus
HER2/neu:	A tumor associated antigen; c-erbB-2 is a synonym (see, e.g., SEQ ID NO: 364)
HLA:	Human leukocyte antigen
HLA-DR:	Human leukocyte antigen class II
HPLC:	High Performance Liquid Chromatography
HTC:	Helper T Cell
HTL:	Helper T Lymphocyte. A synonym for HTC.
ID:	Identity
IFN γ :	Interferon gamma
IL-4:	Interleukin-4
IV:	Intravenous
LU _{30%} :	Cytotoxic activity for 10 ⁶ effector cells required to achieve 30% lysis of a target cell population, at a 100:1 (E:T) ratio.
MAb:	Monoclonal antibody
MAGE:	Melanoma antigen (see, e.g., SEQ ID NO: 365 and 366 for MAGE2 and MAGE3)
MLR:	Mixed lymphocyte reaction
MNC:	Mononuclear cells
PB:	Peripheral blood
PBMC:	Peripheral blood mononuclear cell
ProGP TM :	Progenipoietin TM product (Searle, St. Louis, MO), a chimeric flt3/G-CSF receptor agonist.
SC:	Subcutaneous
S.E.M.:	Standard error of the mean
QD:	Once a day dosing
TAA:	Tumor Associated Antigen
TNF:	Tumor necrosis factor
WBC:	White blood cells

[00106] The following describes the peptides, nucleic acid molecules, compositions, and methods of the invention in more detail.

Methods of Identifying Candidate Peptide Epitopes

[00107] The present invention is directed to methods for selecting a variant of a peptide epitope which induces a CTL response against another variant(s) of the peptide epitope, by determining whether the variant comprises only conserved residues, as defined herein, at non-anchor positions in comparison to the other variant(s).

[00108] In some embodiments, antigen sequences from a population of an infectious agent, said antigens comprising variants of a peptide epitope, are optionally aligned (manually or by computer) along their length, preferably their full length. Variant(s) of a peptide epitope (preferably naturally occurring variants), each 8-11 amino acids in length and comprising the same MHC class I supermotif or motif, are identified manually or with the aid of a computer. In some embodiments, a variant is optionally chosen which comprises preferred anchor residues of said motif and/or which occurs with high frequency within the population of variants. In other embodiments, a variant is randomly chosen. The randomly or otherwise chosen variant is compared to from one to all the remaining variant(s) to determine whether it comprises only conserved residues in the non-anchor positions relative to from one to all the remaining variant(s).

[00109] The present invention is also directed to variants identified by the methods above; peptides comprising such variants; nucleic acids encoding such variants and peptides; cells comprising such variants, and/or peptides, and/or nucleic acids; compositions comprising such variants, and/or peptides, and/or nucleic acids, and/or cells; as well as therapeutic and diagnostic methods for using such variants, peptides, nucleic acids, cells, and compositions.

[00110] In some embodiments, the invention is directed to a method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising

- a) identifying, from a particular antigen of an infectious agent, variants of a peptide epitope 8-11 amino acids in length, each variant comprising primary anchor residues of the same HLA class I binding motif; and

- b) determining whether one of said variants comprises only conserved non-anchor residues in comparison to at least one remaining variant, thereby identifying a candidate peptide epitope.

[00111] In some embodiments, (b) comprises identifying a variant which comprises only conserved non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

[00112] In some embodiments, the invention is directed to a method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising

- a) identifying, from a particular antigen of an infectious agent, variants of a peptide epitope 8-11 amino acids in length, each variant comprising primary anchor residues of the same HLA class I binding motif;
- b) determining whether each of said variants comprises conserved, semi-conserved or non-conserved non-anchor residues in comparison to each of the remaining variants; and
- c) identifying a variant which comprises only conserved non-anchor residues in comparison to at least one remaining variant.

[00113] In some embodiments, (c) comprises identifying a variant which comprises only conservative non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

[00114] In some embodiments, the invention is directed to a method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising

- a) identifying, from a particular antigen of an infectious agent, a population of variants of a peptide epitope 8-11 amino acids in length, each peptide epitope comprising primary anchor residues of the same HLA class I binding motif;
- b) choosing a variant selected from the group consisting of:
 - i) a variant which comprises preferred primary anchor residues of said motif; and

- ii) a variant which occurs with high frequency within the population of variants; and
- c) determining whether the variant of (b) comprises only conserved non-anchor residues in comparison to at least one remaining variant, thereby identifying a candidate peptide epitope.

[00115] In some embodiments, (c) comprises identifying a variant which comprises only conservative non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

[00116] In some embodiments, the invention is directed to method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising

- a) identifying, from a particular antigen of an infectious agent, a population of variants of a peptide epitope 8-11 amino acids in length, each peptide epitope comprising primary anchor residues of the same HLA class I binding motif;
- b) choosing a variant selected from the group consisting of:
 - i) a variant which comprises preferred primary anchor residues of said motif; and
 - ii) a variant which occurs with high frequency within the population of variants; and
- c) determining whether the variant of (b) comprises conserved, semi-conserved or non-conserved non-anchor residues in comparison to each of the remaining variants; and
- d) identifying a variant which comprises only conserved non-anchor residues in comparison to at least one remaining variant.

[00117] In some embodiments, (d) comprises identifying a variant which comprises only conservative non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

[00118] In some embodiments, (a) comprises aligning the sequences of said antigens.

- [00119] In some embodiments, (b) comprises comprising choosing a variant which comprises preferred primary anchor residues of said motif.
- [00120] In some embodiments, (b) comprises comprising choosing a variant which occurs with high frequency within said population.
- [00121] In some embodiments, (b) comprises ranking said variants by frequency of occurrence within said population.
- [00122] In some embodiments, (b) comprises choosing a variant which comprises preferred primary anchor residues of said motif and which occurs with high frequency within said population.
- [00123] In some embodiments, (b) comprises ranking said variants by frequency of occurrence within said population.
- [00124] In some embodiments, the identified variant comprises the fewest conserved anchor residues in comparison to each of the remaining variants.
- [00125] In some embodiments, the remaining variants comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 27, 28, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 220, 240, 260, 280, or 300 variants.
- [00126] In some embodiments, the infectious agent is selected from the group consisting of: HIV, HBV, HCV, HPV, *Plasmodium falciparum*, Influenza virus, and Dengue virus, Epstein-Barr virus, *Mycobacterium tuberculosis*, *Chlamydia*, *Candida albicans*, *Cryptococcus neoformans*, *Coccidioides spp.*, *Histoplasma spp.*, *Aspergillus fumigatis*, *Plasmodium spp.*, *Trypanosoma spp.*, *Schistosoma spp.*, and *Leishmania spp.*
- [00127] In some embodiments, the infectious agent is selected from the group consisting of: HIV, HBV, HCV, HPV, *Plasmodium falciparum*, Influenza virus, and Dengue virus.
- [00128] In some embodiments, the infectious agent is HIV and the antigen is selected from the group consisting of: Gag, Env, Pol, Nef, Rev, Tat, Vif, Vpr, and Vpu.
- [00129] In some embodiments, the infectious agent is HBV and the antigen is selected from the group consisting of: Pol, Env, Core, and NS1/Env2.
- [00130] In some embodiments, the infectious agent is HCV and the antigen is selected from the group consisting of: Core, E1, E2, NS1, NS2, NS3, NS4, and NS5.
- [00131] In some embodiments, the infectious agent is HPV and the antigen is selected from the group consisting of: E1, E2, E3, E4, E5, E6, E7, L1, and L2.

- [00132] In some embodiments, the infectious agent is *Plasmodium falciparum* and the antigen is selected from the group consisting of: CSP, SSP2, EXP1, LSA1.
- [00133] In some embodiments, the selected variant and the at least one remaining variant comprise different primary anchor residues of the same motif or supermotif.
- [00134] In some embodiments, the motif or supermotif is selected from the group consisting of those in Tables 1-2.
- [00135] In some embodiments, the conserved non-anchor residues are at any of positions 3-7 of said variant.
- [00136] In some embodiments, the variant comprises only 1-3 conserved non-anchor residues compared to at least one remaining variant.
- [00137] In some embodiments, the variant comprises only 1-2 conserved non-anchor residues compared to at least one remaining variant.
- [00138] In some embodiments, the variant comprises only 1 conserved non-anchor residue compared to at least one remaining variant.
- [00139] In some embodiments, the infectious agent is HPV, and further wherein, the HPV infectious agent is selected from the group consisting of HPV strains 16, 18, 31, 33, 45, 52, 56, and 58.
- [00140] In some embodiments, the variants are a population of naturally occurring variants.
- [00141] **Optional Alignment.** Optionally, antigen sequences, either full-length or partial, may be aligned manually or by computer. Convenient computer programs for aligning multiple sequences include Omega, Oxford software, version 1.1.3, using ClustalW alignment, using an open gap penalty of 10.0, extend gap penalty of 0.05, and delay divergent sequences of 40.0 (See, e.g., Table 21); and BLASTP 2.2.5 (Nov-16-2002) (Altschul, S.F., et al., Nucleic Acids Res. 25:3389-3402 (1997)) using a cutoff = 3e-88 (to select human sequences) (see, e.g., Table 20). Alternatively, alignments may be obtained through publicly available sources such as published journal articles and published patent documents or as disclosed herein (see, e.g., Tables 10-22).
- [00142] **HLA Class I Motifs Indicative of CTL Inducing Peptide Epitopes.** A large fraction of HLA class I and class II molecules can be classified into a relatively few supertypes, each respective supertype characterized by largely overlapping peptide binding repertoires, and consensus structures of the main peptide binding pockets. Thus,

peptides of the present invention are preferably identified by the primary residues of any one of several HLA-specific amino acid motifs, or if the presence of the motif corresponds to the ability to bind several allele-specific HLA antigens, a supermotif (*see, e.g.*, Tables 1-2). The preferred primary residues are indicated in bold, while the tolerated primary residues are indicated by italics.

[00143] The primary anchor residues of the HLA class I peptide epitope supermotifs and motifs are summarized in Tables 1-2. Preferred primary anchors are shown in bold, while tolerated primary anchors are shown in italics. Primary and secondary anchor positions for HLA Class I are summarized in Table 3. Allele-specific HLA molecules that fall within the various HLA class I supertypes are listed in Table 4. In some cases, patterns of amino acid residues are present in both a motif and a supermotif. The relationship of a particular motif and any related supermotif is indicated in the description of the individual motifs.

[00144] Thus, the peptide motifs and supermotifs described below, and summarized in Tables 1-2, provide guidance for the identification and use of peptide epitopes comprising primary anchor residues of motifs or supermotifs in accordance with the invention.

[00145] Allele-specific HLA molecules that comprise HLA class I supertype families are listed in Table 4.

[00146] **HLA-A1 supermotif.** The HLA-A1 supermotif is characterized by the presence in peptide ligands of a small (T or S) or hydrophobic (L, I, V, or M) primary anchor residue in position 2, and an aromatic (Y, F, or W) primary anchor residue at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind to the A1 supermotif (*i.e.*, the HLA-A1 supertype) is comprised of at least A*0101, A*2601, A*2602, A*2501, and A*3201 (*see, e.g.*, DiBrino, M. *et al.*, *J. Immunol.* 151:5930, 1993; DiBrino, M. *et al.*, *J. Immunol.* 152:620, 1994; Kondo, A. *et al.*, *Immunogenetics* 45:249, 1997). Other allele-specific HLA molecules predicted to be members of the A1 superfamily are shown in Table 4. Peptides binding to each of the individual HLA proteins can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00147] **HLA-A2 supermotif.** Primary anchor specificities for allele-specific HLA-A2.1 molecules (*see, e.g.*, Falk *et al.*, *Nature* 351:290-296, 1991; Hunt *et al.*, *Science* 255:1261-1263, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992; Ruppert *et al.*, *Cell* 74:929-937, 1993) and cross-reactive binding among HLA-A2 and -A28 molecules have been

described. (See, e.g., Fruci *et al.*, *Human Immunol.* 38:187-192, 1993; Tanigaki *et al.*, *Human Immunol.* 39:155-162, 1994; Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994 for reviews of relevant data.) These primary anchor residues define the HLA-A2 supermotif; which presence in peptide ligands corresponds to the ability to bind several different HLA-A2 and -A28 molecules. The HLA-A2 supermotif comprises peptide ligands with L, I, V, M, A, T, or Q as a primary anchor residue at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope.

[00148] The corresponding family of HLA molecules (*i.e.*, the HLA-A2 supertype that binds these peptides) is comprised of at least: A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, and A*6901. Other allele-specific HLA molecules predicted to be members of the A2 superfamily are shown in Table 4. As explained in detail below, binding to each of the individual allele-specific HLA molecules can be modulated by substitutions at the primary anchor and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00149] The motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

[00150] **HLA-A3 supermotif.** The HLA-A3 supermotif is characterized by the presence in peptide ligands of A, L, I, V, M, S, or, T as a primary anchor at position 2, and a positively charged residue, R or K, at the C-terminal position of the epitope, *e.g.*, in position 9 of 9-mers (see, *e.g.*, Sidney *et al.*, *Hum. Immunol.* 45:79, 1996). Exemplary members of the corresponding family of HLA molecules (the HLA-A3 supertype) that bind the A3 supermotif include at least A*0301, A*1101, A*3101, A*3301, and A*6801. Other allele-specific HLA molecules predicted to be members of the A3 supertype are shown in Table 4. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions of amino acids at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

[00151] **HLA-A24 supermotif.** The HLA-A24 supermotif is characterized by the presence in peptide ligands of an aromatic (F, W, or Y) or hydrophobic aliphatic (L, I, V, M, or T) residue as a primary anchor in position 2, and Y, F, W, L, I, or M as primary anchor at the C-terminal position of the epitope (see, *e.g.*, Sette and Sidney, *Immunogenetics*, in press,

1999). The corresponding family of HLA molecules that bind to the A24 supermotif (*i.e.*, the A24 supertype) includes at least A*2402, A*3001, and A*2301. Other allele-specific HLA molecules predicted to be members of the A24 supertype are shown in Table 4. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00152] HLA-B7 supermotif. The HLA-B7 supermotif is characterized by peptides bearing proline in position 2 as a primary anchor, and a hydrophobic or aliphatic amino acid (L, I, V, M, A, F, W, or Y) as the primary anchor at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind the B7 supermotif (*i.e.*, the HLA-B7 supertype) is comprised of at least twenty six HLA-B proteins including: B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, and B*7801 (*see, e.g.*, Sidney, *et al.*, *J. Immunol.* 154:247, 1995; Barber, *et al.*, *Curr. Biol.* 5:179, 1995; Hill, *et al.*, *Nature* 360:434, 1992; Rammensee, *et al.*, *Immunogenetics* 41:178, 1995 for reviews of relevant data). Other allele-specific HLA molecules predicted to be members of the B7 supertype are shown in Table 4. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

[00153] HLA-B27 supermotif. The HLA-B27 supermotif is characterized by the presence in peptide ligands of a positively charged (R, H, or K) residue as a primary anchor at position 2, and a hydrophobic (F, Y, L, W, M, I, A, or V) residue as a primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sidney and Sette, *Immunogenetics*, in press, 1999). Exemplary members of the corresponding family of HLA molecules that bind to the B27 supermotif (*i.e.*, the B27 supertype) include at least B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, and B*7301. Other allele-specific HLA molecules predicted to be members of the B27 supertype are shown in Table 4. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

- [00154] **HLA-B44 supermotif.** The HLA-B44 supermotif is characterized by the presence in peptide ligands of negatively charged (D or E) residues as a primary anchor in position 2, and hydrophobic residues (F, W, Y, L, I, M, V, or A) as a primary anchor at the C-terminal position of the epitope (*see, e.g.,* Sidney et al., *Immunol. Today* 17:261, 1996). Exemplary members of the corresponding family of HLA molecules that bind to the B44 supermotif (*i.e.,* the B44 supertype) include at least: B*1801, B*1802, B*3701, B*4001, B*4002, B*4006, B*4402, B*4403, and B*4006. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the supermotif.
- [00155] **HLA-B58 supermotif.** The HLA-B58 supermotif is characterized by the presence in peptide ligands of a small aliphatic residue (A, S, or T) as a primary anchor residue at position 2, and an aromatic or hydrophobic residue (F, W, Y, L, I, V, M, or A) as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.,* Sidney and Sette, *Immunogenetics*, in press, 1999 for reviews of relevant data). Exemplary members of the corresponding family of HLA molecules that bind to the B58 supermotif (*i.e.,* the B58 supertype) include at least: B*1516, B*1517, B*5701, B*5702, and B*5801. Other allele-specific HLA molecules predicted to be members of the B58 supertype are shown in Table 4. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.
- [00156] **HLA-B62 supermotif.** The HLA-B62 supermotif is characterized by the presence in peptide ligands of the polar aliphatic residue Q or a hydrophobic aliphatic residue (L, V, M, I, or P) as a primary anchor in position 2, and a hydrophobic residue (F, W, Y, M, I, V, L, or A) as a primary anchor at the C-terminal position of the epitope (*see, e.g.,* Sidney and Sette, *Immunogenetics*, in press, 1999). Exemplary members of the corresponding family of HLA molecules that bind to the B62 supermotif (*i.e.,* the B62 supertype) include at least: B*1501, B*1502, B*1513, and B5201. Other allele-specific HLA molecules predicted to be members of the B62 supertype are shown in Table 4. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.
- [00157] **HLA-A1 motif.** The HLA-A1 motif is characterized by the presence in peptide ligands of T, S, or M as a primary anchor residue at position 2 and the presence of Y as a

primary anchor residue at the C-terminal position of the epitope. An alternative allele-specific A1 motif is characterized by a primary anchor residue at position 3 rather than position 2. This motif is characterized by the presence of D, E, A, or S as a primary anchor residue in position 3, and a Y as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.,* DiBrino *et al.*, *J. Immunol.*, 152:620, 1994; Kondo *et al.*, *Immunogenetics* 45:249, 1997; and Kubo *et al.*, *J. Immunol.* 152:3913, 1994 for reviews of relevant data). Peptide binding to HLA A1 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

[00158] Those epitopes comprising T, S, or M at position 2 and Y at the C-terminal position are also HLA-A1 supermotif-bearing peptide epitopes, as these residues are a subset of the A1 supermotif primary anchors.

[00159] **HLA-A*0201 motif.** An HLA-A2*0201 motif was determined to be characterized by the presence in peptide ligands of L or M as a primary anchor residue in position 2, and L or V as a primary anchor residue at the C-terminal position of a 9-residue peptide (*see, e.g.,* Falk *et al.*, *Nature* 351:290-296, 1991) and was further found to comprise an I at position 2 and I or A at the C-terminal position of a nine amino acid peptide (*see, e.g.,* Hunt *et al.*, *Science* 255:1261-1263, March 6, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992). The A*0201 allele-specific motif has also been defined by the present inventors to additionally comprise V, A, T, or Q as a primary anchor residue at position 2, and M or T as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.,* Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994). Thus, the HLA-A*0201 motif comprises peptide ligands with L, I, V, M, A, T, or Q as primary anchor residues at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope. The preferred and tolerated residues that characterize the primary anchor positions of the HLA-A*0201 motif are identical to the residues describing the A2 supermotif. (For reviews of relevant data, *see, e.g.,* Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Ruppert *et al.*, *Cell* 74:929-937, 1993; Sidney *et al.*, *Immunol. Today* 17:261-266, 1996; Sette and Sidney, *Curr. Opin. in Immunol.* 10:478-482, 1998). Secondary anchor residues that characterize the A*0201 motif have additionally been defined (*see, e.g.,* Ruppert *et al.*, *Cell* 74:929-937, 1993). These are shown in Table 3. Peptide binding to HLA-A*0201 molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

- [00160] **HLA-A3 motif.** The HLA-A3 motif is characterized by the presence in peptide ligands of L, M, V, I, S, A, T, F, C, G, or D as a primary anchor residue at position 2, and the presence of K, Y, R, H, F, or A as a primary anchor residue at the C-terminal position of the epitope (*see, e.g., DiBrino et al., Proc. Natl. Acad. Sci USA* 90:1508, 1993; and Kubo *et al., J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.
- [00161] The A3 supermotif primary anchor residues comprise a subset of the A3- and A11-allele specific motif primary anchor residues.
- [00162] **HLA-A11 motif.** The HLA-A11 motif is characterized by the presence in peptide ligands of V, T, M, L, I, S, A, G, N, C, D, or F as a primary anchor residue in position 2, and K, R, Y, or H as a primary anchor residue at the C-terminal position of the epitope (*see, e.g., Zhang et al., Proc. Natl. Acad. Sci USA* 90:2217-2221, 1993; and Kubo *et al., J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A11 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.
- [00163] There is extensive overlap between the A3 and A11 motif primary anchor specificities.
- [00164] **HLA-A24 motif.** The HLA-A24 motif is characterized by the presence in peptide ligands of Y, F, W, or M as a primary anchor residue in position 2, and F, L, I, or W as a primary anchor residue at the C-terminal position of the epitope (*see, e.g., Kondo et al., J. Immunol.* 155:4307-4312, 1995; and Kubo *et al., J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A24 molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the motif.
- [00165] The primary anchor residues characterizing the A24 allele-specific motif comprise a subset of the A24 supermotif primary anchor residues.
- [00166] **Computer or Manual Screening.** Peptides bearing HLA Class I or Class II supermotifs or motifs may be identified by computer searches or manually, e.g., as follows. In utilizing computer screening to identify peptide epitopes, a protein sequence or translated sequence may be analyzed using software developed to search for motifs, for example the "FINDPATTERNS" program (Devereux, *et al. Nucl. Acids Res.* 12:387-395,

1984) or MotifSearch 1.4 software program (D. Brown, San Diego, CA) to identify potential peptide sequences containing appropriate HLA binding motifs. The identified peptides can be scored using customized polynomial algorithms to predict their capacity to bind specific HLA class I or class II alleles. As appreciated by one of ordinary skill in the art, a large array of computer programming software and hardware options are available in the relevant art which can be employed to implement the motifs in order to evaluate (*e.g.*, without limitation, to identify epitopes, identify epitope concentration per peptide length, or to generate analogs) known or unknown peptide sequences.

[00167] Translated antigen protein sequences may be analyzed using a text string search software program, *e.g.*, MotifSearch 1.4 (D. Brown, San Diego) to identify potential peptide sequences containing appropriate HLA binding motifs; alternative programs are readily produced in accordance with information in the art in view of the motif/supermotif disclosure herein. Furthermore, such calculations can be made mentally.

[00168] Identified supermotif or motif sequences may be scored using polynomial algorithms to predict their capacity to bind to specific HLA-Class I or Class II molecules. These polynomial algorithms take into account both extended and refined motifs (that is, to account for the impact of different amino acids at different positions), and are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA molecule interactions can be approximated as a linear polynomial function of the type:

$$"\Delta G" = a_{1i} \times a_{2i} \times a_{3i} \dots \times a_{ni}$$

where a_{ji} is a coefficient which represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. The crucial assumption of this method is that the effects at each position are essentially independent of each other (*i.e.*, independent binding of individual side-chains). When residue j occurs at position i in the peptide, it is assumed to contribute a constant amount j_i to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide. This assumption is justified by studies from our laboratories that demonstrated that peptides are bound to MHC and recognized by T cells in essentially an extended conformation (data omitted herein).

[00169] The method of derivation of specific algorithm coefficients has been described in Gulukota *et al.*, *J. Mol. Biol.* 267:1258-126, 1997; (see also Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; and Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). Briefly, for all i positions, anchor and non-anchor alike, the geometric mean of the average

relative binding (ARB) of all peptides carrying j is calculated relative to the remainder of the group, and used as the estimate of j_i . For Class II peptides, if multiple alignments are possible, only the highest scoring alignment is utilized, following an iterative procedure. To calculate an algorithm score of a given peptide in a test set, the ARB values corresponding to the sequence of the peptide are multiplied. If this product exceeds a chosen threshold, the peptide is predicted to bind. Appropriate thresholds are chosen as a function of the degree of stringency of prediction desired.

[00170] Additional methods to identify preferred peptide sequences, which also make use of specific motifs, include the use of neural networks and molecular modeling programs (see, e.g., Milik *et al.*, *Nature Biotechnology* 16:753, 1998; Altuvia *et al.*, *Hum. Immunol.* 58:1, 1997; Altuvia *et al.*, *J. Mol. Biol.* 249:244, 1995; Buus, S. *Curr. Opin. Immunol.* 11:209-213, 1999; Brusic, V. *et al.*, *Bioinformatics* 14:121-130, 1998; Parker *et al.*, *J. Immunol.* 152:163, 1993; Meister *et al.*, *Vaccine* 13:581, 1995; Hammer *et al.*, *J. Exp. Med.* 180:2353, 1994; Sturniolo *et al.*, *Nature Biotechnol.* 17:555 1999).

[00171] **Conserved, Semi-conserved, and Non-conserved Non-anchor Residues.** The determination of non-anchor residues as being conserved (conservative) or semi-conserved (semi-conservative) or non-conserved (non-conservative) in comparison to the non-anchor positions of from one to all of the remaining variant(s) is defined by as follows, the results of which are summarized in Table 5.

[00172] Table 5 shows the similarity assignments between any given amino acid pair so that a given amino acid substitution could be characterized as being a (conservative) or semi-conserved (semi-conservative) or non-conserved (non-conservative) residue.

[00173] The degree of similarity between amino acid pairs was quantified by averaging, for each amino acid pair, the rank coefficient scores for PAM250, hydrophobicity, and side chain volume as described below. Based on the average values of these composite rankings, Table 5 shows each pair to be conserved, semi-conserved or non-conserved.

[00174] The Dayhoff PAM250 score (Dayhoff, M.O., *et al.*, *Atlas of Protein Sequence and Structure*, Vol. 5, suppl.3. (1978) M.O. Dayhoff, ed. National Biomedical Research Foundation, Washington DC, p. 345; Creighton, T.E., *Proteins: structures and molecular properties* (1993) (2nd edition) W.H. Freeman and Company, NY; <http://prowl.rockefeller.edu/aainfo/pam250.html>) is a commonly utilized protein

alignment scoring matrix which measures the percentage of acceptable point mutations (PAM) within a defined time frame. The frequencies of these mutations are different from what would be expected from the probability of random mutations, and presumably reflect a bias due to the degree of physical and chemical similarity of the amino acid pair involved in the substitution. To obtain a score of amino acid similarity that could be standardized with other measures of similarity, the PAM250 scores were converted to a rank value, where 1 indicates the highest probability of being an accepted mutation.

[00175] The most commonly utilized scales to represent the relative hydrophobicity of the 20 naturally occurring amino acids (Cornette, J., *et al.*, *J. Mol. Biol.* (1987) 195:659) are those developed on the basis of experimental data by Kyte and Doolittle (Kyte, J. and R.F. Doolittle, *J. Mol. Biol.* (1982) 157:105), and by Fauchere and Pliska (Fauchere, J. and V. Pliska, *Eur. J. Med. Chem.* (1983) 18:369). The Kyte/Doolittle scale measures the H₂O/organic solvent partition of individual amino acids. Because it considers the position of amino acids in folded proteins, it may most accurately reflect native hydrophobicity in the context of proteins. The Fauchere/Pliska scale measures the octanol/H₂O partitioning of N-acetyl amino acid amides, and most accurately reflects hydrophobicity in the context of denatured proteins and/or small synthetic peptides. To obtain scores for hydrophobicity, each amino acid residue was ranked on both the Kyte/Doolittle and Fauchere/Pliska hydrophobicity scales. An average rank between the two scales was calculated and the average difference in hydrophobicity for each pair was calculated.

[00176] Finally, for calculating amino acid side-chain volume, the partial volume in solution obtained by noting the increase in volume of water after adding either one molecule or one gram of amino acid residue was considered (Zamyatnin, A.A., *Ann. Rev. Biophys. Bioeng.* (1984) 13:145; Zamyatnin, A.A., *Prog. Biophys. Mol. Biol.* (1972) 24:107). The absolute difference in the partial volume of each possible pairing of the 20 naturally occurring amino acids was calculated and ranked, where 1 indicated residues with the most similar volumes, and 20 the most dissimilar.

[00177] Thus, by consulting Table 5, one can determine whether a residue in a variant is considered to be conserved, semi-conserved, or non-conserved in comparison to a residue in another variant(s). The residue of the parent variant (randomly or otherwise chosen variant) is shown across the top of Table 5, and the residue of the variant(s) it is compared with is shown below the parent residue.

[00178] As shown in Table 5, each of the amino acids shown across the top of the table bears a numerically defined relationship to the remaining 19 genetically encoded amino acids. The lower the index, the higher the conservation; the same amino acid will have a similarity assignment of 1.0; maximally different amino acids will have similarity assignments approaching 20. Using the method set forth above, amino acids which are not gene-encoded can also be assigned similarity indices and can be classified with respect to any natively occurring amino acid as conserved (conservative) or semi-conserved (semi-conservative) or non-conserved (non-conservative).

Variant Peptide Epitopes

[00179] In some embodiments, the invention is directed to an isolated peptide comprising or consisting of a variant. In some embodiments, the invention is directed to an isolated polynucleotide encoding such a peptide.

[00180] The isolated variants of the invention are all class I binding peptides, i.e., CTL peptides. In particular, the variants of the invention comprise a motif or supermotif, as described above. Variants of the invention are those set forth in Tables 6-9 and Figures 1A-4 (SEQ ID Nos: __). Variants of the invention may be referred to herein as “variants” and “variant peptide epitopes” or referred to by Table or referred to by SEQ ID NO. Other peptide epitopes are referred to herein as CTL epitopes or CTL peptides and HTL epitopes or HTL peptides.

[00181] **Peptides and Polynucleotides.** In some embodiments, the invention is directed to an isolated peptide comprising or consisting of a variant, wherein the variant consists of a sequence selected from those in Tables 6-9 and Figures 1A-4 (SEQ ID Nos: __).

[00182] Peptides of the invention may be fusion proteins of variant(s) to CTL epitope(s), and/or HTL epitope(s), and/or linker(s), and/or spacer(s), and/or carrier(s), and/or additional amino acid(s), and/or may comprise or consist of homopolymers of a variant or heteropolymers of more than one variant, as is described in detail below.

[00183] Peptides which comprise a variant of the invention may comprise or consist of a fragment of an antigen (“fragment” or “antigenic fragment”), wherein the fragment comprises a variant. The fragment may be a portion of any antigen of an infectious agent, e.g., the sequences in Tables 11-22 (SEQ ID Nos: __, respectively). The variant of the invention may be within the fragment or may be linked, directly or indirectly, to the fragment.

[00184] The fragment may comprise or consist of a region of a native antigen that contains a high concentration of class I and/or class II epitopes, preferably it contains the greatest number of epitopes per amino acid length. Such epitopes can be present in a frame-shifted manner, e.g. a 10

amino acid long peptide could contain two 9 amino acid long epitopes and one 10 amino acid long epitope.

[00185] The fragment may be less than or equal to 600 amino acids, less than or equal to 500 amino acids, less than or equal to 400 amino acids, less than or equal to 250 amino acids, less than or equal to 100 amino acids, less than or equal to 85 amino acids, less than or equal to 75 amino acids, less than or equal to 65 amino acids, or less than or equal to 50 amino acids in length. In certain embodiments, a fragment is less than 101 amino acids in length, in any increment down to 5 amino acids in length. For example, the fragment may be 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 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, or 100 amino acids in length. Fragments of full length antigens may be fragments from about residue 1-20, 21-40, 41-60, 61-80, 81-100, 101-120, 121-140, 141-160, 161-180, 181-200, 201-220, 221-240, 241-260, 261-280, 281-300, 301-320, 321-340, 341-360, 361-380, 381-400, 401-420, 421-440, 441-460, 461-480, 481-500, 501-520, 521-540, 541-560, 561-580, 581-600, 601-620, 621-680, 681-700, 701-720, 721-740, 741-780, 781-800, 801-820, 821-840, 841-860, 861-880, 881-900, 901-920, 921-940, 941-960, 961-980, 981 to the C-terminus of the antigen.

[00186] Peptides which comprise a variant of the invention may be a fusion protein comprising one or more amino acid residues in addition to the variant or fragment. Fusion proteins include homopolymers and heteropolymers, as described below.

[00187] In some embodiments, the peptide comprises or consists of multiple variants, *e.g.*, 2, 3, 4, 5, 6, 7, 8, or 9 variants of the invention. In some embodiments, the peptide comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 variants of the invention.

[00188] The peptide may also be a homopolymer of one variant or the peptide may be a heteropolymer which contains at least two different variants. Polymers have the advantage of increased probability for immunological reaction and, where different variants are used to make up the polymer, the ability to induce antibodies and/or T cells that react with different antigenic determinants of the antigen(s) targeted for an immune response.

[00189] A homopolymer may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 copies of the same variant.

[00190] A heteropolymer may comprise one or more copies of an individual variant and one or more copies of one or more different variants of the invention. The variants that form a heteropolymer may all be from the same antigen, *e.g.*, may be from any of those in Tables 11-22

(SEQ ID NOS:_) or other antigens herein or known in the art, or may be from different antigens, preferably from infectious agents. Combinations of variants that may form a heteropolymer include, for example, Gag 545 variants EPLTSLKSLF (SEQ ID NO:_) and YPLASLKSLF (SEQ ID NO:_), or combinations of peptides from different tables in Tables 6-9 and/or Figures 1A-4 or those combinations in Tables 23-28. Heteropolymers may contain multiple copies of one or more variants.

[00191] Thus, peptides of the invention such as heteropolymers may comprise a first variant and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 other (different) variants.

[00192] In some embodiments, the peptide comprising a variant may also comprise a number of CTL and/or HTL epitopes, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 CTL and/or HTL epitopes.

[00193] The CTL and/or HTL epitope and the variant of the invention may be from the same antigen of an infectious agent or from different antigens. Thus, for example, if the variant is from HIV pol, the CTL peptide and/or HTL peptide may also be from HIV pol. Alternatively, if the variant is from HIV pol, the CTL peptide and/or HTL peptide may be from another antigen such as HIV env or HIV vpr. As another example, if the variant is from HBV E6, the CTL peptide and/or HTL peptide may be from HBV E7. The CTL and/or HTL epitope and the variant of the invention may be from the same infectious agent or different infectious agents. Thus, for example, the variant may be from HIV, and the CTL and/or HTL epitope may be from HIV or may be from another infectious agent such as HBV, HCV, HPV, or *Plasmodium falciparum*.

[00194] The CTL peptide and/or HTL peptide may be from other antigens including hepatitis B core and surface antigens (HBVc, HBVs), hepatitis C antigens, Epstein-Barr virus antigens, human immunodeficiency virus (HIV) antigens and human papilloma virus (HPV) antigens (in particular antigens from HPV-16, HPV-18, HPV-31, HPV-33, HPV-45, HPV-52, HPV-56 and HPV-58, *Mycobacterium tuberculosis* and *Chlamydia*. Examples of suitable fungal antigens include those derived from *Candida albicans*, *Cryptococcus neoformans*, *Coccidioides spp.*, *Histoplasma spp.*, and *Aspergillus fumigatis*. Examples of suitable protozoan parasitic antigens include those derived from *Plasmodium spp.*, including *P. falciparum*, *Trypanosoma spp.*, *Schistosoma spp.*, *Leishmania spp.* and the like.

[00195] Alternatively, the CTL peptide and/or HTL peptide may be from tumor-associated antigens such as but not limited to, melanoma antigens MAGE-1, MAGE-2, MAGE-3, MAGE-11, MAGE-A10, as well as BAGE, GAGE, RAGE, MAGE-C1, LAGE-1, CAG-3, DAM, MUC1, MUC2, MUC18, NY-ESO-1, MUM-1, CDK4, BRCA2, NY-LU-1, NY-LU-7, NY-LU-12,

CASP8, RAS, KIAA-2-5, SCCs, p53, p73, CEA, HER2/neu, Melan-A, gp100, tyrosinase, TRP2, gp75/TRP1, kallikrein, prostate-specific membrane antigen (PSM), prostatic acid phosphatase (PAP), prostate-specific antigen (PSA), PT1-1, β -catenin, PRAME, Telomerase, FAK, cyclin D1 protein, NOEY2, EGF-R, SART-1, CAPB, HPVE7, p15, Folate receptor CDC27, PAGE-1, and PAGE-4.

[00196] Examples of CTL peptides and HTL peptides are disclosed in WO 01/42270, published 14 June 2001; WO 01/41788, published 14 June 2001; WO 01/42270, published 14 June 2001; WO 01/45728, published 28 June 2001; and WO 01/41787, published 14 June 2001.

[00197] The HTL peptide may comprise a "loosely HLA-restricted" or "promiscuous" sequence. Examples of amino acid sequences that are promiscuous include sequences from antigens such as tetanus toxoid at positions 830-843 (QYIKANSKFIGITE; SEQ ID NO: 627), *Plasmodium falciparum* CS protein at positions 378-398 (DIEKKIAKMEKASSVFNVVNS; SEQ ID NO: 628), and Streptococcus 18kD protein at positions 116-131 (GAVDSILGGVATYGAA; SEQ ID NO: 629). Other examples include peptides bearing a DR 1-4-7 supermotif, or either of the DR3 motifs.

[00198] The HTL peptide may comprise a synthetic peptide such as a Pan-DR-binding epitope (e.g., a PADRE[®] peptide, Epimmune Inc., San Diego, CA, described, for example, in U.S. Patent Number 5,736,142), for example, having the formula aKXVAAZTLKAAa, where "X" is either cyclohexylalanine, phenylalanine, or tyrosine; "Z" is either tryptophan, tyrosine, histidine or asparagine; and "a" is either D-alanine or L-alanine (SEQ ID NO: 746). Certain pan-DR binding epitopes comprise all "L" natural amino acids; these molecules can be provided as peptides or in the form of nucleic acids that encode the peptide. See also, U.S. Patent Nos. 5,679,640 and 6,413,935.

[00199] The peptide comprising a variant may comprise additional amino acid(s). Such additional amino acids may be Ala, Arg, Asn, Asp, Cys, Gln, Gly, Glu, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Tyr, Trp, Val, amino acid mimetics, and other unnatural amino acids such as those described below. Additional amino acids may provide for ease of linking peptides one to another, for linking variants to one another, for linking variants to CTL and/or HTL epitopes, for coupling to a carrier support or larger peptide, for modifying the physical or chemical properties of the peptide or oligopeptide, or the like. Amino acids such as Ala, Arg, Asn, Asp, Cys, Gln, Gly, Glu, His, Ile, Leu, Lys, Met, Phe,

Pro, Ser, Thr, Tyr, Trp, or Val, or the like, can be introduced at the C- and/or N-terminus of the peptide and/or can be introduced internally.

[00200] The peptide comprising a variant may comprise an amino acid spacer(s), which may be joined to the variants, CTL epitopes, HTL epitopes, carriers, etc. within a peptide or may be joined to the peptide at the N-and/or C-terminus. Thus, spacers may be at the N-terminus or C-terminus of peptide, or may be internal such that they link or join variants, CTL epitopes, HTL epitopes, carriers, additional amino acids, and/or antigenic fragments one to the other.

[00201] The spacer is typically comprised of one or more relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, *e.g.*, Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer may be composed of the same residues or may be composed of one or more different residues and thus may be a homo- or hetero-oligomer of spacer residues. Thus, the spacer may contain more than one Ala residue (poly-alanine) or more than one Gly residue (poly-glycine), or may contain both Ala and Gly residues, *e.g.*, Gly, Gly-Gly-, Ser, Ser-Ser-, Gly-Ser-, Ser-Gly-, etc. When present, the spacer will usually be at least one or two residues, more usually three to six residues and sometimes 10 or more residues, *e.g.*, 3, 4, 5, 6, 7, 8, 9, or 10, or even more residues. (Livingston, B.D. *et al. Vaccine* 19:4652-4660 (2000)).

[00202] Peptides comprising a variant may comprise carrier(s) such as those well known in the art, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza virus proteins, hepatitis B virus core protein, and the like. (See Table 29).

[00203] In addition, the peptide comprising or consisting of a variant may be modified by terminal-NH₂ acylation, *e.g.*, by alkanoyl (C₁-C₂₀) or thioglycolyl acetylation, terminal-carboxyl amidation, *e.g.*, ammonia, methylamine, *etc.* In some instances these modifications may provide sites for linking to a support or other molecule.

[00204] The peptides in accordance with the invention can contain modifications such as but not limited to glycosylation, side chain oxidation, biotinylation, phosphorylation, addition of a surface active material, *e.g.* a lipid, or can be chemically modified, *e.g.*,

acetylation, *etc.* Moreover, bonds in the peptide can be other than peptide bonds, *e.g.*, covalent bonds, ester or ether bonds, disulfide bonds, hydrogen bonds, ionic bonds, *etc.*

[00205] Peptides of the present invention may contain substitutions to modify a physical property (*e.g.*, stability or solubility) of the resulting peptide. For example, peptides may be modified by the substitution of a cysteine (C) with α -amino butyric acid ("B"). Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Substituting α -amino butyric acid for C not only alleviates this problem, but actually improves binding and crossbinding capability in certain instances. Substitution of cysteine with α -amino butyric acid may occur at any residue of a peptide, *e.g.*, at either anchor or non-anchor positions of a variant within a peptide, or at other positions of a peptide.

[00206] The peptides comprising a variant can comprise amino acid mimetics or unnatural amino acids, *e.g.* D- or L-naphylalanine; D- or L-phenylglycine; D- or L-2-thieneylalanine; D- or L-1, -2, 3, or 4-pyreneylalanine; D- or L-3 thieneylalanine; D- or L-(2-pyridinyl)-alanine; D- or L-(3-pyridinyl)-alanine; D- or L-(2-pyrazinyl)-alanine; D- or L-(4-isopropyl)-phenylglycine; D-(trifluoromethyl)-phenylglycine; D-(trifluoromethyl)-phenylalanine; D-*p*-fluorophenylalanine; D- or L-*p*-biphenylphenylalanine; D- or L-*p*-methoxybiphenylphenylalanine; D- or L-2-indole(alkyl)alanines; and, D- or L-alkylalanines, where the alkyl group can be a substituted or unsubstituted methyl, ethyl, propyl, hexyl, butyl, pentyl, isopropyl, iso-butyl, sec-isotyl, iso-pentyl, or a non-acidic amino acids. Aromatic rings of a non-natural amino acid include, *e.g.*, thiazolyl, thiophenyl, pyrazolyl, benzimidazolyl, naphthyl, furanyl, pyrrolyl, and pyridyl aromatic rings. Modified peptides that have various amino acid mimetics or unnatural amino acids are particularly useful, as they tend to manifest increased stability *in vivo*. Such peptides may also possess improved shelf-life or manufacturing properties.

[00207] Peptide stability can be assayed in a number of ways. For instance, peptidases and various biological media, such as human plasma and serum, have been used to test stability. *See, e.g.*, Verhoef, *et al.*, *Eur. J. Drug Metab. Pharmacokinetics* 11:291 (1986). Half-life of the peptides of the present invention is conveniently determined using a 25% human serum (v/v) assay. The protocol is generally as follows: Pooled human serum (Type AB, non-heat inactivated) is delipidated by centrifugation before use. The serum is then diluted to 25% with RPMI-1640 or another suitable tissue culture medium. At predetermined time intervals, a small amount of reaction solution is removed and added to

either 6% aqueous trichloroacetic acid (TCA) or ethanol. The cloudy reaction sample is cooled (4°C) for 15 minutes and then spun to pellet the precipitated serum proteins. The presence of the peptides is then determined by reversed-phase HPLC using stability-specific chromatography conditions.

[00208] As indicated above, the peptides in accordance with the invention can be a variety of lengths, and either in their neutral (uncharged) forms or in forms which are salts. The peptides in accordance with the invention can contain modifications such as glycosylation, side chain oxidation, or phosphorylation, generally subject to the condition that modifications do not destroy the biological activity of the peptides.

[00209] The peptides of the invention may be lyophilized, or may be in crystal form.

[00210] It is generally preferable that the variant peptide epitope be as small as possible while still maintaining substantially all of the immunologic activity of the native protein. When possible, it may be desirable to optimize HLA class I binding epitopes of the invention to a length of about 8 to about 13 amino acid residues, for example, 8, 9, 10, 11, 12 or 13, preferably 8 to 11 or 9 to 10. It is to be appreciated that one or more epitopes in this size range can be comprised by a longer peptide (see the Definition Section for the term "epitope" for further discussion of peptide length). HLA class II binding epitopes are preferably optimized to a length of about 6 to about 30 amino acids in length, *e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30, preferably to between about 13 and about 20 residues, *e.g.*, 13, 14, 15, 16, 17, 18, 19 or 20. Preferably, the epitopes are commensurate in size with endogenously processed pathogen-derived peptides or tumor cell peptides that are bound to the relevant HLA molecules. The identification and preparation of peptides of various lengths can be carried out using the techniques described herein.

[00211] Peptides in accordance with the invention can be prepared synthetically, by recombinant DNA technology or chemical synthesis, or can be isolated from natural sources such as native tumors or pathogenic organisms. Epitopes may be synthesized individually or joined directly or indirectly in a peptide. Although the peptide will preferably be substantially free of other naturally occurring host cell proteins and fragments thereof, in some embodiments the peptides may be synthetically conjugated to be joined to native fragments or particles.

[00212] The peptides of the invention can be prepared in a wide variety of ways. For relatively short sizes, the peptides can be synthesized in solution or on a solid support in

accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. (See, for example, Stewart & Young, SOLID PHASE PEPTIDE SYNTHESIS, 2D. ED., Pierce Chemical Co., 1984). Further, individual peptides can be joined using chemical ligation to produce larger peptides that are still within the bounds of the invention.

[00213] Alternatively, recombinant DNA technology can be employed wherein a nucleotide sequence which encodes a peptide inserted into an expression vector, transformed or transfected into an appropriate host cell and cultivated under conditions suitable for expression. These procedures are generally known in the art, as described generally in Sambrook *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989). Thus, recombinant peptides, which comprise or consist of one or more epitopes of the invention, can be used to present the appropriate T cell epitope.

[00214] Polynucleotides encoding each of the peptides above are also part of the invention. As appreciated by one of ordinary skill in the art, various nucleic acids will encode the same peptide due to the redundancy of the genetic code. Each of these nucleic acids falls within the scope of the present invention. This embodiment of the invention comprises DNA and RNA, and in certain embodiments a combination of DNA and RNA. It is to be appreciated that any polynucleotide that encodes a peptide in accordance with the invention falls within the scope of this invention.

[00215] The polynucleotides encoding peptides contemplated herein can be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, *et al.*, *J. Am. Chem. Soc.* 103:3185 (1981). Polynucleotides encoding peptides comprising or consisting of a variant can be made simply by substituting the appropriate and desired nucleic acid base(s) for those that encode a related (e.g., analogous) epitope.

[00216] The polynucleotide, e.g. minigene (see below), may be produced by assembling oligonucleotides that encode the plus and minus strands of the polynucleotide, e.g. minigene. Overlapping oligonucleotides (15-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. A polynucleotide, e.g. minigene, encoding the peptide of the invention, can be cloned into a desired vector such as an expression vector. The coding sequence can then be provided with appropriate linkers and ligated into expression vectors commonly

available in the art, and the vectors used to transform suitable hosts to produce the desired peptide such as a fusion protein.

[00217] A large number of such vectors and suitable host systems are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, phagescript, psiX174, pBluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); pCR (Invitrogen). Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia); p75.6 (valentis); pCEP (Invitrogen); pCEI (Epimmune). However, any other plasmid or vector can be used as long as it is replicable and viable in the host.

[00218] As representative examples of appropriate hosts, there can be mentioned: bacterial cells, such as *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*; fungal cells, such as yeast; insect cells such as *Drosophila* and Sf9; animal cells such as COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell* 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines or Bowes melanoma; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

[00219] Thus, the present invention is also directed to vectors, preferably expression vectors useful for the production of the peptides of the present invention, and to host cells comprising such vectors.

[00220] Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which can be, for example, a cloning vector or an expression vector. The vector can be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the polynucleotides. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

[00221] For expression of the peptides, the coding sequence will be provided with operably linked start and stop codons, promoter and terminator regions and usually a replication system to provide an expression vector for expression in the desired cellular host. For

example, promoter sequences compatible with bacterial hosts are provided in plasmids containing convenient restriction sites for insertion of the desired coding sequence. The resulting expression vectors are transformed into suitable bacterial hosts.

[00222] Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), ∇ -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product.

[00223] Yeast, insect or mammalian cell hosts may also be used, employing suitable vectors and control sequences. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell* 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. Such promoters may also be derived from viral sources, such as, *e.g.*, human cytomegalovirus (CMV-IE promoter) or herpes simplex virus type-1 (HSV TK promoter). Nucleic acid sequences derived from the SV40 splice, and polyadenylation sites can be used to provide the required nontranscribed genetic elements.

[00224] Polynucleotides encoding peptides of the invention may also comprise a ubiquitination signal sequence, and/or a targeting sequence such as an endoplasmic reticulum (ER) signal sequence to facilitate movement of the resulting peptide into the endoplasmic reticulum.

[00225] Polynucleotides of the invention, *e.g.*, minigenes, may be expressed in human cells. A human codon usage table can be used to guide the codon choice for each amino

acid. Such polynucleotides preferably comprise spacer amino acid residues between variants, such as those described above, or may comprise naturally-occurring flanking sequences adjacent to the variants (and/or CTL and HTL epitopes).

[00226] The peptides of the invention can also be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. As an example of this approach, vaccinia virus is used as a vector to express nucleotide sequences that encode the peptides of the invention. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the polypeptides of the invention, *e.g.* adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein. A preferred vector is Modified Vaccinia Ankara (MVA) (*e.g.*, Bavarian Noridic (MVA-BN)).

[00227] Standard regulatory sequences well known to those of skill in the art are preferably included in the vector to ensure expression in the human target cells. Several vector elements are desirable: a promoter with a downstream cloning site for polynucleotide, *e.g.*, minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (*e.g.* ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.*, the human cytomegalovirus (hCMV) promoter. See, *e.g.*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences. A preferred promoter is the CMV-IE promoter.

[00228] Polynucleotides, *e.g.* minigenes, may comprise one or more synthetic or naturally-occurring introns in the transcribed region. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing polynucleotide, *e.g.* minigene, expression.

[00229] In addition, the polynucleotide, *e.g.* minigene, may comprise immunostimulatory sequences (ISSs or CpGs). These sequences may be included in the vector, outside the polynucleotide (*e.g.* minigene) coding sequence to enhance immunogenicity.

[00230] In some embodiments, a bi-cistronic expression vector which allows production of both the polynucleotide- (*e.g.* minigene-) encoded peptides of the invention and a second protein (*e.g.*, one that modulates immunogenicity) can be used. Examples of proteins or

polypeptides that, if co-expressed with peptides of the invention, can enhance an immune response include cytokines (*e.g.*, IL-2, IL-12, GM-CSF), cytokine-inducing molecules (*e.g.*, LeIF), costimulatory molecules, or pan-DR binding proteins (PADRE[®] molecules, Epimmune, San Diego, CA). Helper T cell (HTL) epitopes such as PADRE[®] molecules can be joined to intracellular targeting signals and expressed separately from expressed peptides of the invention. Specifically decreasing the immune response by co-expression of immunosuppressive molecules (*e.g.* TGF- β) may be beneficial in certain diseases.

[00231] Once an expression vector is selected, the polynucleotide, *e.g.* minigene, is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate bacterial strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the polynucleotide, *e.g.* minigene, as well as all other elements included in the vector, are confirmed using restriction mapping, DNA sequence analysis, and/or PCR analysis. Bacterial cells harboring the correct plasmid can be stored as cell banks.

[00232] Therapeutic/prophylactic quantities of DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and are grown to saturation in shaker flasks or a bioreactor according to well known techniques. Plasmid DNA is purified using standard bioseparation technologies such as solid phase anion-exchange resins available, *e.g.*, from QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

[00233] Purified polynucleotides, *e.g.* minigenes, can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized polynucleotide, *e.g.* DNA, in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of polynucleotide vaccines, alternative methods of formulating purified plasmid DNA may be used. A variety of such methods have been described, and new techniques may become available. Cationic lipids, glycolipids, and fusogenic liposomes can also be used in the formulation (see, *e.g.*, WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Patent No. 5,279,833; WO 91/06309; and Felgner, *et al.*, *Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) can also be complexed to purified plasmid DNA to

influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

[00234] Known methods in the art can be used to enhance delivery and uptake of a polynucleotide *in vivo*. For example, the polynucleotide can be complexed to polyvinylpyrrolidone (PVP), to prolong the localized bioavailability of the polynucleotide, thereby enhancing uptake of the polynucleotide by the organism (*see e.g.*, U.S. Patent No. 6,040,295; EP 0 465 529; WO 98/17814). PVP is a polyamide that is known to form complexes with a wide variety of substances, and is chemically and physiologically inert.

[00235] Target cell sensitization can be used as a functional assay of the expression and HLA class I presentation of polynucleotide- (e.g. minigene-) encoded peptides. For example, the polynucleotide, e.g. plasmid DNA, is introduced into a mammalian cell line that is a suitable target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. For example, electroporation can be used for "naked" DNA, whereas cationic lipids or PVP-formulated DNA allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). The transfected cells are then chromium-51 (^{51}Cr) labeled and used as targets for epitope-specific CTLs. Cytolysis of the target cells, detected by ^{51}Cr release, indicates both production and HLA presentation of, polynucleotide-, e.g. minigene-, encoded variants of the invention, or peptides comprising them. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL activity.

[00236] *In vivo* immunogenicity is a second approach for functional testing of polynucleotides, e.g. minigenes. Transgenic mice expressing appropriate human HLA proteins are immunized with the polynucleotide, e.g. DNA, product. The dose and route of administration are formulation dependent (*e.g.*, IM for polynucleotide (*e.g.*, naked DNA or PVP-formulated DNA) in PBS, intraperitoneal (IP) for lipid-complexed polynucleotide (*e.g.*, DNA)). Eleven to twenty-one days after immunization, splenocytes are harvested and restimulated for one week in the presence of polynucleotides encoding each peptide being tested. Thereafter, for peptides comprising or consisting of variants, standard assays are conducted to determine if there is cytolysis of peptide-loaded, ^{51}Cr -labeled target cells. Once again, lysis of target cells that were exposed to variants corresponding to those encoded by the polynucleotide (e.g. minigene) demonstrates polynucleotide (e.g., DNA)

vaccine function and induction of CTLs. Immunogenicity of HTL epitopes is evaluated in transgenic mice in an analogous manner.

[00237] Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles comprised solely of a polynucleotide such as DNA are administered. In a further alternative embodiment for ballistic delivery, polynucleotides such as DNA can be adhered to particles, such as gold particles.

[00238] The use of polynucleotides such as multi-epitope minigenes is described herein and in, *e.g.* co-pending application U.S.S.N. 09/311,784; Ishioka *et al.*, *J. Immunol.* 162:3915-3925, 1999; An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997; Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a polynucleotide such as a multi-epitope DNA plasmid can be engineered which encodes an epitope derived from multiple regions of a infectious agent (*e.g.*, p53, HER2/nev, MAGE-2/3, or CEA), a pan-DR binding peptide such as the PADRE[®] universal helper T cell epitope, and an endoplasmic reticulum-translocating signal sequence. As described in the sections above, a peptide/polynucleotide may also comprise/encode epitopes that are derived from other infectious agents.

[00239] Thus, the invention includes peptides as described herein, polynucleotides encoding each of said peptides, as well as compositions comprising the peptides and polynucleotides, and includes methods for producing and methods of using the peptides, polynucleotides, and compositions, as further described below.

[00240] **Compositions.** In other embodiments, the invention is directed to a composition comprising one or more peptides and/or polynucleotides of the invention and optionally another component(s).

[00241] In some embodiments, the composition comprises or consists of multiple peptides, *e.g.*, 2, 3, 4, 5, 6, 7, 8, or 9 peptides of the invention. In some embodiments, the composition comprises at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 peptides of the invention. Combinations of peptides include, for example, a peptide comprising or alternatively consisting of the Gag 545 variant EPLTSLKSLF (SEQ ID NO:_) and a peptide comprising or alternatively consisting of the Gag 545 variant YPLASLKSLF (SEQ ID NO:_), or combinations of peptides from different tables in Tables 6-9 and/or Figures 1A-4.

[00242] Compositions of the invention may comprise polynucleotides encoding the above peptides and/or combinations of peptides.

[00243] The composition can comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 peptides and/or polynucleotides selected from those described above or below. At least one of the one or more peptides can be a heteropolymer or a homopolymer. Additionally, the composition can comprise a CTL and/or HTL epitope, which can be derived from a tumor-associated antigen. The additional epitope can also be a PanDR binding molecule, (*e.g.*, a PADRE[®] universal helper T cell epitope).

[00244] Optional components include excipients, diluents, proteins such as peptides comprising a CTL epitope, and/or an HTL epitope such as a pan-DR binding peptide (*e.g.*, a PADRE[®] universal helper T cell epitope), and/or a carrier, polynucleotides encoding such proteins, lipids, or liposomes, as well as other components described herein. There are numerous embodiments of compositions in accordance with the invention, such as a cocktail of one or more peptides and/or polynucleotides (*e.g.*, minigenes); a cocktail of one or more peptides and/or polynucleotides (*e.g.*, minigenes) and one or more CTL and/or HTL epitopes.

[00245] Compositions may comprise one or more peptides (and/or polynucleotides such as minigenes) of the invention, along with one or more other components as described above and herein. "One or more" refers to any whole unit integer from 1-150, *e.g.*, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 peptides, polynucleotides, or other components.

[00246] Compositions of the invention may be, for example, polynucleotides or polypeptides of the invention combined with or complexed to cationic lipid formulations; lipopeptides (*e.g.*, Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), encapsulated *e.g.*, in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g.*, Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991; Alonso *et al.*, *Vaccine* 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995); peptide compositions contained in immune stimulating complexes (ISCOMS) (*see, e.g.*, Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998); multiple antigen peptide systems (MAPs) (*see e.g.*, Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996); viral, bacterial, or, fungal delivery vectors (Perkus, M. E. *et*

al., In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990); particles of viral or synthetic origin (e.g., Kofler, N. *et al.*, *J. Immunol. Methods* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995); adjuvants (e.g., incomplete Freund's adjuvant) (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993); liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996); or, particle-absorbed cDNA or other polynucleotides of the invention (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993), etc. Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) or attached to a stress protein, e.g., HSP 96 (Stressgen Biotechnologies Corp., Victoria, BC, Canada) can also be used.

[00247] Compositions of the invention comprise polynucleotide-mediated modalities. DNA or RNA encoding one or more of the peptides of the invention can be administered to a patient. This approach is described, for instance, in Wolff *et al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; and, WO 98/04720. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivacaine, polymers (e.g., PVP, PINC, etc.), peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (*see, e.g.*, U.S. Patent No. 5,922,687). Accordingly, peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as Modified Vaccinia Ankara (MVA) (e.g., Bavarian Noridic), vaccinia or fowlpox. For example, vaccinia virus is used as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into an acutely or chronically infected host or into a non-infected host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin).

BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, *e.g.* adeno and adeno-associated virus vectors, alpha virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, are apparent to those skilled in the art from the description herein.

[00248] In certain embodiments, components that induce T cell responses are combined with components that induce antibody responses to the target antigen of interest. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. Alternatively, a composition comprises a class I and/or class II epitope in accordance with the invention, along with a PADRE[®] molecule (Epimmune, San Diego, CA).

[00249] Compositions of the invention can comprise antigen presenting cells, such as dendritic cells. Antigen presenting cells, *e.g.*, dendritic cells, may be transfected, *e.g.*, with a polynucleotide such as a minigene construct in accordance with the invention, in order to elicit immune responses. The peptide can be bound to an HLA molecule on the antigen-presenting cell, whereby when an HLA-restricted cytotoxic T lymphocyte (CTL) is present, a receptor of the CTL binds to a complex of the HLA molecule and the peptide.

[00250] The compositions of the invention may also comprise antiviral drugs such as interferon- α , or immune adjuvants such as IL-12, GM-CSF, etc.

[00251] Compositions may comprise an HLA heavy chain, β_2 -microglobulin, streptavidin, and/or biotin. The streptavidin may be fluorescently labeled. Compositions may comprise tetramers (*see e.g.*, U.S. Pat. No. 5,635,363; *Science* 274:94-96 (1996)). A tetramer composition comprising an HLA heavy chain, β_2 -microglobulin, streptavidin, and biotin. The streptavidin may be fluorescently labeled. Compositions may also comprise dimers. A dimer composition comprises as MHC molecule and an Ig molecule (*see e.g.*, *PNAS* 95:7568-73 (1998)).

[00252] In some embodiments it may be desirable to include in the compositions of the invention at least one component which primes cytotoxic T lymphocytes. Lipids have been identified as agents capable of priming CTL *in vivo* against viral antigens. For example, palmitic acid residues can be attached to the ϵ - and α - amino groups of a lysine residue and then linked, *e.g.*, via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic peptide. The lipidated peptide can then be administered either directly in a micelle or particle, incorporated into a liposome, or

emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. A preferred composition comprises palmitic acid attached to ϵ - and α - amino groups of Lys, which is attached via linkage, *e.g.*, Ser-Ser, to the amino terminus of the peptide.

[00253] As another example of lipid priming of CTL responses, *E. coli* lipoproteins, such as tripalmitoyl-S-glycerylcysteinylserine (P₃CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide (*see, e.g.*, Deres, *et al.*, *Nature* 342:561, 1989). Peptides of the invention can be coupled to P₃CSS, for example, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be primed with P₃CSS-conjugated epitopes, two such compositions can be combined to more effectively elicit both humoral and cell-mediated responses.

[00254] Another preferred embodiment is a composition comprising one or more peptides of the invention emulsified in IFA.

[00255] Compositions of the invention may also comprise CTL and/or HTL peptides. Such CTL and HTL peptides can be modified by the addition of amino acids to the termini of a peptide to provide for ease of linking peptides one to another, for coupling to a carrier support or larger peptide, for modifying the physical or chemical properties of the peptide or oligopeptide, or the like. Amino acids such as tyrosine, cysteine, lysine, glutamic or aspartic acid, or naturally or unnaturally occurring amino acid residues, can be introduced at the carboxyl- or amino-terminus of the peptide or oligopeptide, particularly class I peptides. However, it is to be noted that modification at the carboxyl terminus of a CTL epitope may, in some cases, alter binding characteristics of the peptide. In addition, the peptide or oligopeptide sequences can differ from the natural sequence by being modified by terminal-NH₂ acylation, *e.g.*, by alkanoyl (C₁-C₂₀) or thioglycolyl acetylation, terminal-carboxyl amidation, *e.g.*, ammonia, methylamine, *etc.* In some instances these modifications may provide sites for linking to a support or other molecule. CTL and HTL epitopes may comprise additional amino acids, such as those described above including spacers.

[00256] A further embodiment of a composition in accordance with the invention is an antigen presenting cell that comprises one or more peptides in accordance with the invention. The antigen presenting cell can be a "professional" antigen presenting cell, such as a dendritic cell. The antigen presenting cell can comprise the peptide of the invention by any means known or to be determined in the art. Such means include pulsing

of dendritic cells with one or more individual peptides, by nucleic acid administration such as ballistic nucleic acid delivery or by other techniques in the art for administration of nucleic acids, including vector-based, *e.g.* viral vector, delivery of nucleic acids.

[00257] Compositions may comprise carriers. Carriers that can be used with compositions of the invention are well known in the art, and include, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza virus proteins, hepatitis B virus core protein, and the like.

[00258] The compositions (*e.g.* pharmaceutical compositions) can contain a physiologically tolerable diluent such as water, or a saline solution, preferably phosphate buffered saline. Additionally, as disclosed herein, CTL responses can be primed by conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glycerol-cysteinyl-serine (P₃CSS).

[00259] Compositions of the invention may be pharmaceutically acceptable compositions. Pharmaceutical compositions preferably contain an immunologically effective amount of one or more peptides and/or polynucleotides of the invention, and optionally one or more other components which are pharmaceutically acceptable. A preferred composition comprises one or more peptides of the invention and IFA. A more preferred composition of the invention comprises one or more peptides of the invention, one or more peptides, and IFA.

[00260] Upon immunization with a peptide and/or polynucleotide and/or composition in accordance with the invention, via injection (*e.g.*, SC, ID, IM), aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by an immune response comprising the production of antibodies, CTLs and/or HTLs specific for the desired antigen(s). Consequently, the host becomes at least partially immune to subsequent exposure to the infectious agent(s), or at least partially resistant to further development of infectious agent-bearing cells and thereby derives a prophylactic or therapeutic benefit.

[00261] Furthermore, the peptides, primers, and epitopes of the invention can be used in any desired immunization or administration regimen; *e.g.*, as part of periodic vaccinations such as annual vaccinations as in the veterinary arts or as in periodic vaccinations as in the human medical arts, or as in a prime-boost regime wherein an inventive vector or recombinant is administered either before or after the administration of the same or of a different epitope of interest or recombinant or vector expressing such as a same or different epitope of interest (including an inventive recombinant or vector expressing such

as a same or different epitope of interest), see, e.g., U.S. Pat. Nos. 5,997,878; 6,130,066; 6,180,398; 6,267,965; and 6,348,450. An useful viral vector of the present invention is Modified Vaccinia Ankara (MVA) (e.g., Bavarian Noridic (MVA-BN)).

[00262] Recent studies have indicated that a prime-boost protocol, whereby immunization with a poxvirus recombinant expressing a foreign gene product is followed by a boost using a purified subunit preparation form of that gene product, elicits an enhanced immune response relative to the response elicited with either product alone. Human volunteers immunized with a vaccinia recombinant expressing the HIV-1 envelope glycoprotein and boosted with purified HIV-1 envelope glycoprotein subunit preparation exhibit higher HIV-1 neutralizing antibody titers than individuals immunized with just the vaccinia recombinant or purified envelope glycoprotein alone (Graham *et al.*, *J. Infect. Dis.*, 167:533-537 (1993); Cooney *et al.*, *Proc. Natl. Acad. Sci. USA*, 90:1882-1886 (1993)). Humans immunized with two injections of an ALVAC-HIV-1 env recombinant (vCP125) failed to develop HIV specific antibodies. Boosting with purified rgp160 from a vaccinia virus recombinant resulted in detectable HIV-1 neutralizing antibodies. Furthermore, specific lymphocyte T cell proliferation to rgp160 was clearly increased by the boost with rgp160. Envelope specific cytotoxic lymphocyte activity was also detected with this vaccination regimen (Pialoux *et al.*, *AIDS Res. and Hum. Retroviruses*, 11:272-381 (1995)). Macaques immunized with a vaccinia recombinant expressing the simian immunodeficiency virus (SIV) envelope glycoprotein and boosted with SIV envelope glycoprotein from a baculovirus recombinant are protected against SIV challenge (Hu *et al.*, *AID Res. and Hum. Retroviruses*, 3:615-620 (1991); Hu *et al.*, *Science* 255:456-459 (1992)). In the same fashion, purified HCMVgB protein can be used in prime-boost protocols with NYVAC or ALVAC-gB recombinants.

[00263] In certain embodiments, the polynucleotides are complexed in a liposome preparation. Liposomal preparations for use in the instant invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. However, cationic liposomes are particularly preferred because a tight charge complex can be formed between the cationic liposome and the polyanionic nucleic acid. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner *et al.*, *Proc. Natl. Acad. Sci. USA* 84:74137416 (1987), which is herein incorporated by reference); mRNA (Malone *et al.*, *Proc. Natl. Acad. Sci. USA* 86:60776081 (1989), which is herein incorporated by reference); and purified transcription factors (Debs *et al.*, *J. Biol.*

Chem. 265:1018910192 (1990), which is herein incorporated by reference), in functional form.

[00264] Cationic liposomes are readily available. For example, N-[12,3-dioleyloxy]-propyl]-N,N,N-triethylammonium (DOTMA) liposomes are particularly useful and are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. (See, also, Felgner *et al.*, *Proc. Natl Acad. Sci. USA* 84:74137416 (1987)). Other commercially available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boehringer).

[00265] Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g. PCT Publication No. WO 90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes. Preparation of DOTMA liposomes is explained in the literature, see, e.g., P. Felgner *et al.*, *Proc. Natl. Acad. Sci. USA* 84:74137417. Similar methods can be used to prepare liposomes from other cationic lipid materials.

[00266] Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, Ala.), or can be easily prepared using readily available materials. Such materials include phosphatidyl, choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

[00267] For example, commercially available dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), and dioleoylphosphatidyl ethanolamine (DOPE) can be used in various combinations to make conventional liposomes, with or without the addition of cholesterol. Thus, for example, DOPG/DOPC vesicles can be prepared by drying 50 mg each of DOPG and DOPC under a stream of nitrogen gas into a sonication vial. The sample is placed under a vacuum pump overnight and is hydrated the following day with deionized water. The sample is then sonicated for 2 hours in a capped vial, using a Heat Systems model 350 sonicator equipped with an inverted cup (bath type) probe at the maximum setting while the bath is circulated at 15EC. Alternatively, negatively charged vesicles can be prepared without sonication to produce multilamellar vesicles or by extrusion through nucleopore membranes to produce unilamellar vesicles of discrete size. Other methods are known and available to those of skill in the art.

[00268] The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs), with SUVs being preferred. The various liposome nucleic acid complexes are prepared using methods well known in the art. See, e.g., Straubinger *et al.*, *Methods of Immunology* 101:512527 (1983). For example, MLVs containing nucleic acid can be prepared by depositing a thin film of phospholipid on the walls of a glass tube and subsequently hydrating with a solution of the material to be encapsulated. SUVs are prepared by extended sonication of MLVs to produce a homogeneous population of unilamellar liposomes. The material to be entrapped is added to a suspension of preformed MLVs and then sonicated. When using liposomes containing cationic lipids, the dried lipid film is resuspended in an appropriate solution such as sterile water or an isotonic buffer solution such as 10 mM Tris/NaCl, sonicated, and then the preformed liposomes are mixed directly with the DNA. The liposome and DNA form a very stable complex due to binding of the positively charged liposomes to the cationic DNA. SUVs find use with small nucleic acid fragments. LUVs are prepared by a number of methods, well known in the art. Commonly used methods include Ca^{2+} -EDTA chelation (Papahadjopoulos *et al.*, *Biochim. Biophys. Acta* 394:483 (1975); Wilson *et al.*, *Cell* 17:77 (1979)); ether injection (Deamer, D. and Bangham, A., *Biochim. Biophys. Acta* 443:629 (1976); Ostro *et al.*, *Biochem. Biophys. Res. Commun.* 76:836 (1977); Fraley *et al.*, *Proc. Natl. Acad. Sci. USA* 76:3348 (1979)); detergent dialysis (Enoch, H. and Strittmatter, P., *Proc. Natl. Acad. Sci. USA* 76:145 (1979)); and reversephase evaporation (REV) (Fraley *et al.*, *J. Biol. Chem.* 255:10431 (1980); Szoka, F. and Papahadjopoulos, D., *Proc. Natl. Acad. Sci. USA* 75:145 (1978); SchaeferRidder *et al.*, *Science* 215:166 (1982)).

[00269] Generally, the ratio of DNA to liposomes will be from about 10:1 to about 1:10. Preferably, the ration will be from about 5:1 to about 1:5. More preferably, the ration will be about 3:1 to about 1:3. Still more preferably, the ratio will be about 1:1.

[00270] U.S. Patent No. 5,676,954 reports on the injection of genetic material, complexed with cationic liposome carriers, into mice. U.S. Patent Nos. 4,897,355, 4,946,787, 5,049,386, 5,459,127, 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication no. WO 94/9469 provide cationic lipids for use in transfecting DNA into cells and mammals. U.S. Patent Nos. 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication no. WO 94/9469 provide methods for delivering DNA-cationic lipid complexes to mammals.

Binding Affinity of Variants for HLA Molecules

[00271] As indicated herein, the large degree of HLA polymorphism is an important factor to be taken into account with the epitope-based approach to developing therapeutics and diagnostics. To address this factor, epitope selection encompassing identification of peptides capable of binding at high or intermediate affinity to multiple HLA molecules is preferably utilized, most preferably these epitopes bind at high or intermediate affinity to two or more allele-specific HLA molecules. However, in some embodiments, it is preferred that all epitopes in a given composition bind to the alleles of a single HLA supertype or a single HLA molecule.

[00272] Variants of the invention preferably include those that have an IC_{50} or binding affinity value for a class I HLA molecule(s) of 500 nM or better (*i.e.*, the value is ≤ 500 nM). In certain embodiments of the invention, peptides of interest have an IC_{50} or binding affinity value for a class I HLA molecule(s) of 200 nM or better. In certain embodiments of the invention, peptides of interest, such as A1 and A24 peptides, have an IC_{50} or binding affinity value for a class I HLA molecule(s) of 100 nM or better. If HTL epitopes are included, they preferably are HTL epitopes that have an IC_{50} or binding affinity value for class II HLA molecules of 1000 nM or better, (*i.e.*, the value is $\leq 1,000$ nM). For example, peptide binding is assessed by testing the capacity of a candidate peptide to bind to a purified HLA molecule *in vitro*. Peptides exhibiting high or intermediate affinity are then considered for further analysis. Selected peptides are generally tested on other members of the supertype family. In preferred embodiments, peptides that exhibit cross-reactive binding are then used in cellular screening analyses or vaccines.

[00273] The relationship between binding affinity for HLA class I molecules and immunogenicity of discrete peptide epitopes on bound antigens was determined for the first time by inventors at Epimmune. As disclosed in greater detail herein, higher HLA binding affinity is correlated with greater immunogenicity.

[00274] Greater immunogenicity can be manifested in several different ways. Immunogenicity corresponds to whether an immune response is elicited at all, and to the vigor of any particular response, as well as to the extent of a population in which a response is elicited. For example, a peptide might elicit an immune response in a diverse array of the population, yet in no instance produce a vigorous response. In accordance

with these principles, close to 90% of high binding peptides have been found to elicit a response and thus be “immunogenic,” as contrasted with about 50% of the peptides that bind with intermediate affinity. (See, e.g., Schaeffer *et al.* PNAS (1988)) High affinity-binding class I peptides generally have an affinity of less than or equal to 100 nM. Moreover, not only did peptides with higher binding affinity have an enhanced probability of generating an immune response, the generated response tended to be more vigorous than the response seen with weaker binding peptides. As a result, less peptide is required to elicit a similar biological effect if a high affinity binding peptide is used rather than a lower affinity one. Thus, in some preferred embodiments of the invention, high affinity binding epitopes are used.

[00275] The correlation between binding affinity and immunogenicity was analyzed by the present inventors by two different experimental approaches (*see, e.g., Sette, et al., J. Immunol.* 153:5586-5592 (1994)). In the first approach, the immunogenicity of potential epitopes ranging in HLA binding affinity over a 10,000-fold range was analyzed in HLA-A*0201 transgenic mice. In the second approach, the antigenicity of approximately 100 different hepatitis B virus (HBV)-derived potential epitopes, all carrying A*0201 binding motifs, was assessed by using PBL from acute hepatitis patients. Pursuant to these approaches, it was determined that an affinity threshold value of approximately 500 nM (preferably 50 nM or less) determines the capacity of a peptide epitope to elicit a CTL response. These data are true for class I binding affinity measurements for naturally processed peptides and for synthesized T cell epitopes. These data also indicate the important role of determinant selection in the shaping of T cell responses (*see, e.g., Schaeffer et al. Proc. Natl. Acad. Sci. USA* 86:4649-4653 (1989)).

[00276] An affinity threshold associated with immunogenicity in the context of HLA class II (*i.e.*, HLA DR) molecules has also been delineated (*see, e.g., Southwood et al. J. Immunology* 160:3363-3373 (1998), and U.S. Patent No. 6,413,527, issued July 2, 2002). In order to define a biologically significant threshold of HLA class II binding affinity, a database of the binding affinities of 32 DR-restricted epitopes for their restricting element (*i.e.*, the HLA molecule that binds the epitope) was compiled. In approximately half of the cases (15 of 32 epitopes), DR restriction was associated with high binding affinities, *i.e.* binding affinity values of 100 nM or less. In the other half of the cases (16 of 32), DR restriction was associated with intermediate affinity (binding affinity values in the 100-1000 nM range). In only one of 32 cases was DR restriction associated with an IC₅₀

of 1000 nM or greater. Thus, 1000 nM is defined as an affinity threshold associated with immunogenicity in the context of DR molecules.

[00277] The binding affinity of peptides for HLA molecules can be determined as described in Example 1, below.

Enhancing Population Coverage of the Vaccine

[00278] The primary anchor residues of the HLA class I peptide epitope supermotifs and motifs are summarized in Tables 1-2. Allele-specific HLA molecules that are comprised by the various HLA class I supertypes are listed in Table 4. In some cases, patterns of amino acid residues are present in both a motif and a supermotif. The relationship of a particular motif and any related supermotif is indicated in the description of the individual motifs.

[00279] By inclusion of one or more epitopes from several motifs or supermotifs in a vaccine composition, enhanced population coverage for major global ethnicities can be obtained.

Assays to Detect T-Cell Responses

[00280] Once HLA binding peptides are identified, they can be tested for the ability to elicit a T-cell response. The preparation and evaluation of motif-bearing peptides are described, *e.g.*, in PCT publications WO 94/20127 and WO 94/03205. Briefly, peptides comprising epitopes from a particular antigen are synthesized and tested for their ability to bind to relevant HLA proteins. These assays may involve evaluation of peptide binding to purified HLA class I molecules in relation to the binding of a radioiodinated reference peptide. Alternatively, cells expressing empty class I molecules (*i.e.* cell surface HLA molecules that lack any bound peptide) may be evaluated for peptide binding by immunofluorescent staining and flow microfluorimetry. Other assays that may be used to evaluate peptide binding include peptide-dependent class I assembly assays and/or the inhibition of CTL recognition by peptide competition. Those peptides that bind to an HLA class I molecule, typically with an affinity of 500 nM or less, are further evaluated for their ability to serve as targets for CTLs derived from infected or immunized individuals, as well as for their capacity to induce primary *in vitro* or *in vivo* CTL

responses that can give rise to CTL populations capable of reacting with selected target cells associated with pathology.

[00281] Analogous assays are used for evaluation of HLA class II binding peptides. HLA class II motif-bearing peptides that are shown to bind, typically at an affinity of 1000 nM or less, are further evaluated for the ability to stimulate HTL responses.

[00282] Conventional assays utilized to detect T cell responses include proliferation assays, lymphokine secretion assays, direct cytotoxicity assays, and limiting dilution assays. For example, antigen-presenting cells that have been incubated with a peptide can be assayed for the ability to induce CTL responses in responder cell populations. Antigen-presenting cells can be normal cells such as peripheral blood mononuclear cells or dendritic cells. Alternatively, mutant, non-human mammalian cell lines that have been transfected with a human class I MHC gene, and that are deficient in their ability to load class I molecules with internally processed peptides, are used to evaluate the capacity of the peptide to induce *in vitro* primary CTL responses. Peripheral blood mononuclear cells (PBMCs) can be used as the source of CTL precursors. Antigen presenting cells are incubated with peptide, after which the peptide-loaded antigen-presenting cells are then incubated with the responder cell population under optimized culture conditions. Positive CTL activation can be determined by assaying the culture for the presence of CTLs that lyse radio-labeled target cells, either specific peptide-pulsed targets or target cells that express endogenously processed antigen from which the specific peptide was derived. Alternatively, the presence of epitope-specific CTLs can be determined by IFN γ *in situ* ELISA.

[00283] In an embodiment of the invention, directed to diagnostics, a method has been devised which allows direct quantification of antigen-specific T cells by staining with fluorescein-labelled HLA tetrameric complexes (Altman, J. D. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:10330, 1993; Altman, J. D. *et al.*, *Science* 274:94, 1996). Other options include staining for intracellular lymphokines, and interferon release assays or ELISPOT assays. Tetramer staining, intracellular lymphokine staining and ELISPOT assays all appear to be at least 10-fold more sensitive than more conventional assays (Lalvani, A. *et al.*, *J. Exp. Med.* 186:859, 1997; Dunbar, P. R. *et al.*, *Curr. Biol.* 8:413, 1998; Murali-Krishna, K. *et al.*, *Immunity* 8:177, 1998). Additionally, DimerX technology can be used as a means of quantitation (*see, e.g.*, *Science* 274:94-99 (1996) and *Proc. Natl. Acad. Sci.* 95:7568-73 (1998)).

[00284] HTL activation may also be assessed using techniques known to those in the art, such as T cell proliferation or lymphokine secretion (*see, e.g. Alexander et al., Immunity* 1:751-761, 1994).

[00285] Alternatively, immunization of HLA transgenic mice can be used to determine immunogenicity of peptide epitopes. Several transgenic mouse strains, *e.g.*, mice with human A2.1, A11 (which can additionally be used to analyze HLA-A3 epitopes), and B7 alleles have been characterized. Other transgenic mice strains (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. Moreover, HLA-DR1 and HLA-DR3 mouse models have been developed. In accordance with principles in the art, additional transgenic mouse models with other HLA alleles are generated as necessary.

[00286] Such mice can be immunized with peptides emulsified in Incomplete Freund's Adjuvant; thereafter any resulting T cells can be tested for their capacity to recognize target cells that have been peptide-pulsed or transfected with genes encoding the peptide of interest. CTL responses can be analyzed using cytotoxicity assays described above. Similarly, HTL responses can be analyzed using, *e.g.*, T cell proliferation or lymphokine secretion assays.

Minigenes

[00287] A number of different approaches are available which allow simultaneous delivery of multiple epitopes. Nucleic acids encoding multiple epitopes are a useful embodiment of the invention; discrete peptide epitopes or polyepitopic peptides can be encoded. The epitopes to be included in a minigene are preferably selected according to the guidelines set forth in the previous section. Examples of amino acid sequences that can be included in a minigene include: HLA class I epitopes, HLA class II epitopes, a ubiquitination signal sequence, and/or a targeting sequence such as an endoplasmic reticulum (ER) signal sequence to facilitate movement of the resulting peptide into the endoplasmic reticulum. Examples of minigene constructs are shown in Tables 23-28.

[00288] The use of multi-epitope minigenes is also described in, *e.g.*, co-pending applications U.S.S.N. 09/311,784, 09/894,018, 60/419,973, 60/415,463; Ishioka *et al.*, *J. Immunol.* 162:3915-3925, 1999; An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997; Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a multi-epitope DNA plasmid

encoding nine dominant HLA-A*0201- and A11-restricted CTL epitopes derived from the polymerase, envelope, and core proteins of HBV and human immunodeficiency virus (HIV), a PADRE[®] universal helper T cell (HTL) epitope, and an endoplasmic reticulum-translocating signal sequence has been engineered. Immunization of HLA transgenic mice with this plasmid construct resulted in strong CTL induction responses against the nine CTL epitopes tested. This CTL response was similar to that observed with a lipopeptide of known immunogenicity in humans, and significantly greater than immunization using peptides in oil-based adjuvants. Moreover, the immunogenicity of DNA-encoded epitopes *in vitro* was also correlated with the *in vitro* responses of specific CTL lines against target cells transfected with the DNA plasmid. These data show that the minigene served: 1.) to generate a CTL response and 2.) to generate CTLs that recognized cells expressing the encoded epitopes. A similar approach can be used to develop minigenes encoding epitopes of an infectious agent.

[00289] For example, to create a DNA sequence encoding the selected epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes may be reverse translated. A human codon usage table can be used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences may be directly adjoined, so that when translated, a continuous peptide sequence is created. However, to optimize expression and/or immunogenicity, additional elements can be incorporated into the minigene design such as spacer amino acid residues between epitopes. HLA presentation of CTL and HTL epitopes may be improved by including synthetic (*e.g.* poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL or HTL epitopes; these larger peptides comprising the epitope(s) are within the scope of the invention. In one embodiment, spacer amino acid residues between one or more CTL and/or HTL epitopes are designed so as to minimize junctional epitopes that may result from the juxtaposition of 2 CTL and/or HTL epitopes.

[00290] The minigene sequence may be converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides (30-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. This synthetic minigene, encoding the epitope peptide, can then be cloned into a desired expression vector.

- [00291] Standard regulatory sequences well known to those of skill in the art are preferably included in the vector to ensure expression in the target cells. Several vector elements are desirable: a promoter with a downstream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (*e.g.* ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.*, the human cytomegalovirus (hCMV) CMV-IE promoter. See, *e.g.*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.
- [00292] Optimized peptide expression and immunogenicity can be achieved by certain modifications to a minigene construct. For example, in some cases introns facilitate efficient gene expression, thus one or more synthetic or naturally-occurring introns can be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing minigene expression.
- [00293] Once an expression vector is selected, the minigene is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate bacterial strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are confirmed using restriction mapping, PCR and/or DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as cell banks.
- [00294] In addition, immunostimulatory sequences (ISSs or CpGs) appear to play a role in the immunogenicity of DNA vaccines. These sequences may be included in the vector, outside the minigene coding sequence to enhance immunogenicity.
- [00295] In some embodiments, a bi-cistronic expression vector which allows production of both the minigene-encoded epitopes and a second protein (*e.g.*, one that modulates immunogenicity) can be used. Examples of proteins or polypeptides that, if co-expressed with epitopes, can enhance an immune response include cytokines (*e.g.*, IL-2, IL-12, GM-CSF), cytokine-inducing molecules (*e.g.*, LeIF), costimulatory molecules, or pan-DR binding proteins (PADRE[®], Epimmune, San Diego, CA). Helper T cell (HTL) epitopes such as PADRE[®] molecules can be joined to intracellular targeting signals and expressed separately from expressed CTL epitopes. This can be done in order to direct HTL epitopes to a cell compartment different than that of the CTL epitopes, one that provides for more efficient entry of HTL epitopes into the HLA class II pathway, thereby

improving HTL induction. In contrast to HTL or CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (*e.g.* TGF- β) may be beneficial in certain diseases.

[00296] Therapeutic quantities of plasmid DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and are grown to saturation in shaker flasks or a bioreactor according to well known techniques. Plasmid DNA is purified using standard bioseparation technologies such as solid phase anion-exchange resins available, *e.g.*, from QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

[00297] Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of minigene vaccines, alternative methods of formulating purified plasmid DNA may be used. A variety of such methods have been described, and new techniques may become available. Cationic lipids, glycolipids, and fusogenic liposomes can also be used in the formulation (see, *e.g.*, WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Patent No. 5,279,833; WO 91/06309; and Felgner, *et al.*, *Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) can also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

[00298] Known methods in the art can be used to enhance delivery and uptake of a polynucleotide *in vivo*. For example, the polynucleotide can be complexed to polyvinylpyrrolidone (PVP), to prolong the localized bioavailability of the polynucleotide, thereby enhancing uptake of the polynucleotide by the organism (*see e.g.*, U.S. Patent No. 6,040,295; EP 0 465 529; WO 98/17814). PVP is a polyamide that is known to form complexes with a wide variety of substances, and is chemically and physiologically inert.

[00299] Target cell sensitization can be used as a functional assay of the expression and HLA class I presentation of minigene-encoded epitopes. For example, the plasmid DNA is introduced into a mammalian cell line that is a suitable target for standard CTL chromium release assays. The transfection method used will be dependent on the final

formulation, electroporation can be used for "naked" DNA, whereas cationic lipids or DNA:PVP compositions allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). The transfected cells are then chromium-51 (^{51}Cr) labeled and used as targets for epitope-specific CTLs. Cytolysis of the target cells, detected by ^{51}Cr release, indicates both the production and HLA presentation of, minigene-encoded CTL epitopes. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL activity.

[00300] *In vivo* immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human HLA proteins are immunized with the DNA product. The dose and route of administration are formulation dependent (*e.g.*, IM for DNA in PBS, intraperitoneal (IP) for lipid-complexed DNA). Eleven to twenty-one days after immunization, splenocytes are harvested and restimulated for one week in the presence of peptides encoding each epitope being tested. Thereafter, for CTLs, standard assays are conducted to determine if there is cytolysis of peptide-loaded, ^{51}Cr -labeled target cells. Once again, lysis of target cells that were exposed to epitopes corresponding to those in the minigene, demonstrates DNA vaccine function and induction of CTLs. Immunogenicity of HTL epitopes is evaluated in transgenic mice in an analogous manner.

[00301] Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles comprised solely of DNA are administered. In a further alternative embodiment for ballistic delivery, DNA can be adhered to particles, such as gold particles.

Vaccine Compositions

[00302] Vaccines that contain an immunologically effective amount of one or more peptides or polynucleotides of the invention are a further embodiment of the invention. The peptides can be delivered by various means or formulations, all collectively referred to as "vaccine" compositions. Such vaccine compositions, and/or modes of administration, can include, for example, naked DNA, DNA formulated with PVP, DNA in cationic lipid formulations; lipopeptides (*e.g.*, Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), DNA or peptides, encapsulated *e.g.*, in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g.*, Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991; Alonso *et al.*,

Vaccine 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995); peptide compositions contained in immune stimulating complexes (ISCOMS) (*see, e.g.*, Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998); multiple antigen peptide systems (MAPs) (*see e.g.*, Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996); viral, bacterial, or, fungal delivery vectors (Perkus, M. E. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990); particles of viral or synthetic origin (*e.g.*, Kofler, N. *et al.*, *J. Immunol. Methods.* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995); adjuvants (*e.g.*, incomplete freund's adjuvant) (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993); liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996); or, particle-absorbed DNA (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993), etc. Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) or attached to a stress protein, *e.g.*, HSP 96 (Stressgen Biotechnologies Corp., Victoria, BC, Canada) can also be used.

[00303] Vaccines of the invention comprise nucleic acid mediated modalities. DNA or RNA encoding one or more of the peptides of the invention can be administered to a patient. This approach is described, for instance, in Wolff *et al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; and, WO 98/04720. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivacaine, polymers (*e.g.*, PVP), peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (*see, e.g.*, U.S. Patent No. 5,922,687). Accordingly, peptide vaccines of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. For example, vaccinia virus is used as a vector to express nucleotide sequences that encode the peptides of the invention

(*e.g.*, MVA). Upon introduction into an acutely or chronically infected host or into a non-infected host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, *e.g.* adeno and adeno-associated virus vectors, alpha virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, are apparent to those skilled in the art from the description herein.

[00304] Furthermore, vaccines in accordance with the invention can comprise one or more peptides of the invention. Accordingly, a peptide can be present in a vaccine individually; alternatively, the peptide can exist as a homopolymer comprising multiple copies of the same peptide, or as a heteropolymer of various peptides. Polymers have the advantage of increased probability for immunological reaction and, where different peptide epitopes are used to make up the polymer, the ability to induce antibodies and/or T cells that react with different antigenic determinants of the antigen targeted for an immune response. The composition may be a naturally occurring region of an antigen or can be prepared, *e.g.*, recombinantly or by chemical synthesis.

[00305] Carriers that can be used with vaccines of the invention are well known in the art, and include, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza virus proteins, hepatitis B virus core protein, and the like. The vaccines can contain a physiologically tolerable diluent such as water, or a saline solution, preferably phosphate buffered saline. Generally, the vaccines also include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are examples of materials well known in the art. Additionally, as disclosed herein, CTL responses can be primed by conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glycerol-cysteinyl-serine (P₃CSS).

[00306] Upon immunization with a peptide composition in accordance with the invention, via injection (*e.g.*, SC, ID, IM), aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by producing antibodies, CTLs and/or HTLs specific for the desired antigen.

Consequently, the host becomes at least partially immune to subsequent exposure to the infectious agent, and thereby derives a prophylactic or therapeutic benefit.

[00307] In certain embodiments, components that induce T cell responses are combined with components that induce antibody responses to the target antigen of interest. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. Alternatively, a composition comprises a class I and/or class II epitope in accordance with the invention, along with a PADRE[®] molecule (Epimmune, San Diego, CA).

[00308] Vaccines of the invention can comprise antigen presenting cells, such as dendritic cells, as a vehicle to present peptides of the invention. For example, dendritic cells are transfected, *e.g.*, with a minigene construct in accordance with the invention, in order to elicit immune responses. Minigenes are discussed in greater detail in a following section. Vaccine compositions can be created *in vitro*, following dendritic cell mobilization and harvesting, whereby loading of dendritic cells occurs *in vitro*.

[00309] The vaccine compositions of the invention may also be used in combination with antiviral drugs such as interferon- α , or immune adjuvants such as IL-12, GM-CSF, etc.

[00310] Preferably, the following principles are utilized when selecting epitope(s) and/or analogs for inclusion in a vaccine, either peptide-based or nucleic acid-based formulations. Exemplary variants that may be utilized in a vaccine to treat or prevent infectious agent-mediated disease are set out in Tables 6-9 and Figures 1A-4. Each of the following principles can be balanced in order to make the selection. When multiple epitopes are to be used in a vaccine, the epitopes may be, but need not be, contiguous in sequence in the native antigen from which the epitopes are derived. Such multiple epitopes can refer to the order of epitopes within a peptide, or to the selection of epitopes that come from the same region, for use in either individual peptides or in a multi-epitopic peptide.

1.) Variants are selected which, upon administration, mimic immune responses that have been observed to be correlated with prevention or clearance of infectious disease. For HLA Class I, this generally includes 3-7 variants from at least one infectious agent or antigen thereof.

2.) Variants are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an IC₅₀ of 500 nM or less, or for Class II an IC₅₀ of 1000 nM or less. For HLA Class I it is presently preferred to select a peptide having an IC₅₀ of 200 nM or less, as this is believed to better correlate not only to induction of an immune response, but to *in vitro* tumor cell killing as well. For HLA A1 and A24, it is especially preferred to select a peptide having an IC₅₀ of 100 nM or less.

3.) Supermotif bearing-variants, or a sufficient array of allele-specific motif-bearing variants, are selected to give broad population coverage. In general, it is preferable to have at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess the breadth of population coverage.

4.) Of particular relevance are “nested epitopes.” Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. For example, a nested epitope can be a fragment of an antigen from a region that contains multiple epitopes that are overlapping, or one epitope that is completely encompassed by another, *e.g.*, A2 peptides MAGE3.159 and MAGE3.160 are nested epitopes. A peptide comprising “transcendent nested epitopes” is a peptide that has both HLA class I and HLA class II epitopes in it. When providing nested epitopes, it is preferable to provide a sequence that has the greatest number of epitopes per provided sequence. Preferably, one avoids providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a sequence comprising nested epitopes, it is important to evaluate the sequence in order to insure that it does not have pathological or other deleterious biological properties; this is particularly relevant for vaccines directed to infectious organisms.

5.) If a protein with multiple epitopes or a polynucleotide (*e.g.*, minigene) is created, an objective is to generate the smallest peptide that encompasses the epitopes of interest. This principle is similar, if not the same as that employed when selecting a peptide comprising nested epitopes. However, with an artificial peptide comprising multiple epitopes, the size minimization objective is balanced against the need to integrate any spacer sequences between epitopes in the polyepitopic protein. Spacer amino acid residues can be introduced to avoid junctional epitopes (an epitope recognized by the immune system, not present in the target antigen, and only created by the man-made juxtaposition of epitopes), or to facilitate cleavage between epitopes and thereby enhance epitope presentation. Junctional epitopes are generally to be avoided because the recipient may generate an immune response to that non-native epitope. Of particular concern is a junctional epitope that is a “dominant epitope.” A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

[00311] The principles are the same, except junctional epitopes applies to the sequences surrounding the epitope. One must also take care with other sequences in construct to avoid immune response.

T CELL PRIMING MATERIALS

[00312] In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes cytotoxic T

lymphocytes. Lipids have been identified as agents capable of facilitating the priming *in vitro* CTL response against viral antigens. For example, palmitic acid residues can be attached to the ϵ - and α - amino groups of a lysine residue and then linked to an immunogenic peptide. One or more linking moieties can be used such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like. The lipidated peptide can then be administered directly in a micelle or particle, incorporated into a liposome, or emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. A preferred immunogenic composition comprises palmitic acid attached to ϵ - and α - amino groups of Lys via a linking moiety, *e.g.*, Ser-Ser, added to the amino terminus of an immunogenic peptide.

[00313] In another embodiment of lipid-facilitated priming of CTL responses, *E. coli* lipoproteins, such as tripalmitoyl-S-glycerol-cysteinyl-seryl-serine (P₃CSS) can be used to prime CTL when covalently attached to an appropriate peptide. (See, *e.g.*, Deres, *et al.*, *Nature* 342:561, 1989). Thus, peptides of the invention can be coupled to P₃CSS, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be primed with P₃CSS-conjugated epitopes, two such compositions can be combined to elicit both humoral and cell-mediated responses.

DENDRITIC CELLS PULSED WITH CTL AND/OR HTL PEPTIDES

[00314] An embodiment of a vaccine composition in accordance with the invention comprises *ex vivo* administration of a cocktail of epitope-bearing peptides to PBMC, or isolated DC therefrom, from the patient's blood. A pharmaceutical to facilitate harvesting of DC can be used, such as Progenipoinetin™ (Monsanto, St. Louis, MO) or GM-CSF/IL-4. After pulsing the DC with peptides and prior to reinfusion into patients, the DC are washed to remove unbound peptides. In this embodiment, a vaccine comprises peptide-pulsed DCs which present the pulsed peptide epitopes in HLA molecules on their surfaces.

[00315] The DC can be pulsed *ex vivo* with a cocktail of peptides, some of which stimulate CTL responses to one or more antigens of interest, *e.g.*, antigens from infectious agents such as HIV env, HIV pol, HIV gag, HIV vpu, HBV and/or the antigens in Tables 11-22, or otherwise described herein or known in the art. Optionally, a helper T cell (HTL) peptide such as PADRE®, can be included to facilitate the CTL response. Thus, a vaccine in accordance with the invention comprising epitopes from an infectious agent is used to

treat or prevent disease mediated by these agents in patients. A vaccine can be used prior to, during, or following other therapies including, for example, antibiotic therapy, anti-viral therapy (e.g., highly active antiretroviral therapy (HAART) in the case of HIV-AIDS), antibody therapy, cancer therapy, and adjunct therapy, whereupon the vaccine provides decreased morbidity, increased disease free survival and overall survival in recipients.

DIAGNOSTIC AND PROGNOSTIC USES

- [00316] In one embodiment of the invention, HLA class I and class II binding peptides can be used as reagents to evaluate an immune response. Preferably, the following principles are utilized when selecting a variant(s) for diagnostic, prognostic and similar uses. Potential principles include having the binding affinities described earlier, and/or matching the HLA-motif/supermotif of a peptide with the HLA-type of a patient.
- [00317] The evaluated immune response can be induced by any immunogen. For example, the immunogen may result in the production of antigen-specific CTLs or HTLs that recognize the peptide epitope(s) employed as the reagent. Thus, a peptide of the invention may or may not be used as the immunogen. Assay systems that can be used for such analyses include tetramer-based protocols (e.g., DimerX technology (*see, e.g., Science* 274:94-99 (1996) and *Proc. Natl. Acad. Sci.* 95:7568-73 (1998)), staining for intracellular lymphokines, interferon release assays, or ELISPOT assays.
- [00318] For example, following exposure to a putative immunogen, a peptide of the invention can be used in a tetramer staining assay to assess peripheral blood mononuclear cells for the presence of any antigen-specific CTLs. The HLA-tetrameric complex is used to directly visualize antigen-specific CTLs and thereby determine the frequency of such antigen-specific CTLs in a sample of peripheral blood mononuclear cells (*see, e.g., Ogg et al., Science* 279:2103-2106, 1998; and Altman *et al., Science* 174:94-96, 1996).
- [00319] A tetramer reagent comprising a peptide of the invention is generated as follows: A peptide that binds to an HLA molecule is refolded in the presence of the corresponding HLA heavy chain and β_2 -microglobulin to generate a trimolecular complex. The complex is biotinylated at the carboxyl terminal end of the HLA heavy chain, at a site that was previously engineered into the protein. Tetramer formation is then induced by adding streptavidin. When fluorescently labeled streptavidin is used, the tetrameric complex is

used to stain antigen-specific cells. The labeled cells are then readily identified, *e.g.*, by flow cytometry. Such procedures are used for diagnostic or prognostic purposes; the cells identified by the procedure can be used for therapeutic purposes.

[00320] Peptides of the invention are also used as reagents to evaluate immune recall responses. (*see, e.g.*, Bertoni *et al.*, *J. Clin. Invest.* 100:503-513, 1997 and Penna *et al.*, *J. Exp. Med.* 174:1565-1570, 1991.) For example, a PBMC sample from an individual expressing a disease-associated antigen (*e.g.* an antigen from an infectious agent) can be analyzed for the presence of antigen-specific CTLs or HTLs using specific peptides. A blood sample containing mononuclear cells may be evaluated by cultivating the PBMCs and stimulating the cells with a peptide of the invention. After an appropriate cultivation period, the expanded cell population may be analyzed, for example, for CTL or for HTL activity.

[00321] Thus, the peptides can be used to evaluate the efficacy of a vaccine. PBMCs obtained from a patient vaccinated with an immunogen may be analyzed by methods such as those described herein. The patient is HLA typed, and peptide epitopes that are bound by the HLA molecule(s) present in that patient are selected for analysis. The immunogenicity of the vaccine is indicated by the presence of CTLs and/or HTLs directed to epitopes present in the vaccine.

[00322] The peptides of the invention may also be used to make antibodies, using techniques well known in the art (*see, e.g.* *CURRENT PROTOCOLS IN IMMUNOLOGY*, Wiley/Greene, NY; and *Antibodies A Laboratory Manual* Harlow, Harlow and Lane, Cold Spring Harbor Laboratory Press, 1989). Such antibodies are useful as reagents to determine the presence of disease-associated antigens. Antibodies in this category include those that recognize a peptide when bound by an HLA molecule, *i.e.*, antibodies that bind to a peptide-MHC complex.

ADMINISTRATION FOR THERAPEUTIC OR PROPHYLACTIC PURPOSES

[00323] The peptides and polynucleotides of the present invention, including cells and compositions comprising them, are useful for administration to mammals, particularly humans, to treat and/or prevent infection by an infectious agent such as HIV, HBV, HCV, HPV, *Plasmodium falciparum* and other agents described herein or known in the art. Vaccine compositions containing the peptides of the invention are administered to a

patient infected with a particular infectious agent or to an individual susceptible to, or otherwise at risk for, infection with such an agent to elicit an immune response against antigens of that agent and thus enhance the patient's own immune response capabilities. Where susceptible individuals are identified prior to infection, the composition can be targeted to them, thus minimizing the need for administration to a larger population.

[00324] In therapeutic applications, peptide and/or nucleic acid compositions are administered to a patient in an amount sufficient to elicit an effective immune response to the infectious agent antigen and to thereby cure, arrest or slow symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, *e.g.*, the particular composition administered, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

[00325] The vaccine compositions of the invention can be used purely as prophylactic agents. Generally the dosage for an initial prophylactic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000 μg of peptide and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg of peptide. Dosage values for a human typically range from about 500 μg to about 50,000 μg of peptide per 70 kilogram patient. This is followed by boosting dosages of between about 1.0 μg to about 50,000 μg of peptide, administered at defined intervals from about four weeks to six months after the initial administration of vaccine. The immunogenicity of the vaccine may be assessed by measuring the specific activity of CTL and HTL obtained from a sample of the patient's blood.

[00326] As noted above, peptides comprising CTL and/or HTL epitopes of the invention induce immune responses when presented by HLA molecules and contacted with a CTL or HTL specific for an epitope comprised by the peptide. The manner in which the peptide is contacted with the CTL or HTL is not critical to the invention. For instance, the peptide can be contacted with the CTL or HTL either *in vitro* or *in vivo*. If the contacting occurs *in vivo*, peptide can be administered directly, or in other forms/vehicles, *e.g.*, DNA vectors encoding one or more peptides, viral vectors encoding the peptide(s), liposomes, antigen presenting cells such as dendritic cells, and the like.

[00327] Accordingly, for pharmaceutical compositions of the invention in the form of peptides or polypeptides, the peptides or polypeptides can be administered directly. Alternatively, the peptide/polypeptides can be administered indirectly presented on APCs, or as DNA encoding them. Furthermore, the peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences.

[00328] For therapeutic use, administration should generally begin at the first diagnosis of infectious agent-related disease. This is followed by boosting doses at least until symptoms are

substantially abated and for a period thereafter. In chronic disease states, loading doses followed by boosting doses may be required.

[00329] The dosage for an initial therapeutic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1,000 μg of peptide and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg of peptide. Dosage values for a human typically range from about 500 μg to about 50,000 μg of peptide per 70 kilogram patient. Boosting dosages of between about 1.0 μg to about 50,000 μg of peptide, administered pursuant to a boosting regimen over weeks to months, can be administered depending upon the patient's response and condition. Patient response can be determined by measuring the specific activity of CTL and HTL obtained from the patient's blood.

[00330] In certain embodiments, peptides and compositions of the present invention are used in serious disease states. In such cases, as a result of the minimal amounts of extraneous substances and the relative nontoxic nature of the peptides, it is possible and may be desirable to administer substantial excesses of these peptide compositions relative to these stated dosage amounts.

[00331] For treatment of chronic disease, a representative dose is in the range disclosed above, namely where the lower value is about 1, 5, 50, 500, or 1,000 μg of peptide and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg of peptide, preferably from about 500 μg to about 50,000 μg of peptide per 70 kilogram patient. Initial doses followed by boosting doses at established intervals, *e.g.*, from four weeks to six months, may be required, possibly for a prolonged period of time to effectively immunize an individual. In the case of chronic disease, administration should continue until at least clinical symptoms or laboratory tests indicate that the disease has been eliminated or substantially abated, and for a follow-up period thereafter. The dosages, routes of administration, and dose schedules are adjusted in accordance with methodologies known in the art.

[00332] The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral, intrathecal, or local administration. Preferably, the pharmaceutical compositions are administered parentally, *e.g.*, intravenously, subcutaneously, intradermally, or intramuscularly.

[00333] Thus, in a preferred embodiment the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, *e.g.*, water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances or pharmaceutical excipients as may be required to approximate physiological conditions, such as pH-adjusting and

buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, *etc.*

[00334] The concentration of peptides of the invention in the pharmaceutical formulations can vary widely, *i.e.*, from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, *etc.*, in accordance with the particular mode of administration selected.

[00335] A human unit dose form of the peptide composition is typically included in a pharmaceutical composition that also comprises a human unit dose of an acceptable carrier, preferably an aqueous carrier, and is administered in a volume of fluid that is known by those of skill in the art to be used for administration of such compositions to humans (*see, e.g., Remington's Pharmaceutical Sciences*, 17th Edition, A. Gennaro, Editor, Mack Publishing Co., Easton, Pennsylvania, 1985).

[00336] The peptides of the invention can also be administered via liposomes, which serve to target the peptides to a particular tissue, such as lymphoid tissue, or to target selectively to infected cells, as well as to increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to a receptor prevalent among lymphoid cells (such as monoclonal antibodies which bind to the CD45 antigen) or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the peptide compositions. Liposomes for use in accordance with the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, *e.g.*, liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, *e.g.*, Szoka, *et al., Ann. Rev. Biophys. Bioeng.* 9:467 (1980), and U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

[00337] For targeting compositions of the invention to cells of the immune system, a ligand can be incorporated into the liposome, *e.g.*, antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be administered intravenously, locally, topically, *etc.* in a dose which varies according to, *inter alia*, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

[00338] For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, often at a concentration of 25%-75%.

[00339] For aerosol administration, the immunogenic peptides are preferably supplied in finely divided form, along with a surfactant and propellant. Typical percentages of peptides are 0.01%-20% by weight, often 1%-10%. The surfactant must, of course, be pharmaceutically acceptable, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant, although an atomizer may be used in which no propellant is necessary and other percentages are adjusted accordingly. A carrier can also be included, *e.g.*, lecithin for intranasal delivery.

[00340] Antigenic peptides of the invention have been used to elicit a CTL and/or HTL response *ex vivo*, as well. The resulting CTLs or HTLs can be used to treat chronic infections, or tumors in patients that do not respond to other conventional forms of therapy, or who do not respond to a therapeutic peptide or nucleic acid vaccine in accordance with the invention. *Ex vivo* CTL or HTL responses to a particular antigen (infectious or tumor-associated) are induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptide. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cell (an infected cell or a tumor cell).

KITS

[00341] The peptide and nucleic acid compositions of this invention can be provided in kit form together with instructions for vaccine administration. Typically the kit would include desired composition(s) of the invention in a container, preferably in unit dosage form and instructions for administration. For example, a kit would include an APC, such

as a dendritic cell, previously exposed to and now presenting peptides of the invention in a container, preferably in unit dosage form together with instructions for administration. An alternative kit would include a minigene construct with desired nucleic acids of the invention in a container, preferably in unit dosage form together with instructions for administration. Lymphokines such as IL-2 or IL-12 may also be included in the kit. Other kit components that may also be desirable include, for example, a sterile syringe, booster dosages, and other desired excipients.

[00342] The invention will be described in greater detail by way of specific examples. The following examples are offered for illustrative purposes, and are not intended to limit the invention in any manner. Those of skill in the art will readily recognize a variety of non-critical parameters that can be changed or modified to yield alternative embodiments in accordance with the invention.

EXAMPLES

EXAMPLE 1. HLA CLASS I AND CLASS II BINDING ASSAYS

[00343] The following example of peptide binding to HLA molecules demonstrates quantification of binding affinities of HLA class I and class II peptides. Binding assays can be performed with peptides that are either motif-bearing or not motif-bearing.

[00344] Cell lysates were prepared and HLA molecules purified in accordance with disclosed protocols (Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998); Sidney, *et al.*, *J. Immunol.* 154:247 (1995); Sette, *et al.*, *Mol. Immunol.* 31:813 (1994)). The cell lines used as sources of HLA molecules and the antibodies used for the extraction of the HLA molecules from the cell lysates are also described in these publications and are well known in the art.

[00345] Epstein-Barr virus (EBV)-transformed homozygous cell lines, fibroblasts, CIR, or 721.221-transfectants were used as sources of HLA class I molecules. These cells were cultured in RPMI 1640 medium supplemented with 2mM L-glutamine (GIBCO, Grand Island, NY), 50μM 2-ME, 100μg/ml of streptomycin, 100U/ml of penicillin (Irvine Scientific) and 10% heat-inactivated FCS (Irvine Scientific, Santa Ana, CA).

[00346] Cell lysates were prepared as follows. Briefly, cells were lysed at a concentration of 10^8 cells/ml in 50 mM Tris-HCl, pH 8.5, containing 1% Nonidet P-40 (Fluka

Biochemika, Buchs, Switzerland), 150 mM NaCl, 5 mM EDTA, and 2 mM PMSF. Lysates were cleared of debris and nuclei by centrifugation at 15,000 x g for 30min.

[00347] HLA molecules were purified from lysates by affinity chromatography. Lysates were passed twice through two pre-columns of inactivated Sepharose CL4-B and protein A-Sepharose. Next, the lysate was passed over a column of Sepharose CL-4B beads coupled to an appropriate antibody. The anti-HLA column was then washed with 10-column volumes of 10mM Tris-HCL, pH 8.0, in 1% NP-40, PBS, 2-column volumes of PBS, and 2-column volumes of PBS containing 0.4% n-octylglucoside. Finally, MHC molecules were eluted with 50mM diethylamine in 0.15M NaCl containing 0.4% n-octylglucoside, pH 11.5. A 1/25 volume of 2.0M Tris, pH 6.8, was added to the eluate to reduce the pH to ~8.0. Eluates were then concentrated by centrifugation in Centriprep 30 concentrators at 2000 rpm (Amicon, Beverly, MA). Protein content was evaluated by a BCA protein assay (Pierce Chemical Co., Rockford, IL) and confirmed by SDS-PAGE.

[00348] A detailed description of the protocol utilized to measure the binding of peptides to Class I and Class II MHC has been published (Sette *et al.*, *Mol. Immunol.* 31:813, 1994; Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998). Briefly, purified MHC molecules (5 to 500nM) were incubated with various unlabeled peptide inhibitors and 1-10nM ¹²⁵I-radiolabeled probe peptides for 48h in PBS containing 0.05% Nonidet P-40 (NP40) (or 20% w/v digitonin for H-2 IA assays) in the presence of a protease inhibitor cocktail. The final concentrations of protease inhibitors (each from CalBioChem, La Jolla, CA) were 1 mM PMSF, 1.3 nM 1.10 phenanthroline, 73 μM pepstatin A, 8mM EDTA, 6mM N-ethylmaleimide (for Class II assays), and 200 μM N alpha-p-tosyl-L-lysine chloromethyl ketone (TLCK). All assays were performed at pH 7.0 with the exception of DRB1*0301, which was performed at pH 4.5, and DRB1*1601 (DR2w21β₁) and DRB4*0101 (DRw53), which were performed at pH 5.0. pH was adjusted as described elsewhere (see Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998).

[00349] Following incubation, MHC-peptide complexes were separated from free peptide by gel filtration on 7.8 mm x 15 cm TSK200 columns (TosoHaas 16215, Montgomeryville, PA), eluted at 1.2 mls/min with PBS pH 6.5 containing 0.5% NP40 and 0.1% NaN₃. Because the large size of the radiolabeled peptide used for the DRB1*1501 (DR2w2β₁) assay makes separation of bound from unbound peaks more difficult under these conditions, all DRB1*1501 (DR2w2β₁) assays were performed using a 7.8mm x

30cm TSK2000 column eluted at 0.6 mls/min. The eluate from the TSK columns was passed through a Beckman 170 radioisotope detector, and radioactivity was plotted and integrated using a Hewlett-Packard 3396A integrator, and the fraction of peptide bound was determined.

[00350] Radiolabeled peptides were iodinated using the chloramine-T method. Representative radiolabeled probe peptides utilized in each assay, and its assay specific IC_{50} nM, are known in the art. Typically, in preliminary experiments, each MHC preparation was titrated in the presence of fixed amounts of radiolabeled peptides to determine the concentration of HLA molecules necessary to bind 10-20% of the total radioactivity. All subsequent inhibition and direct binding assays were performed using these HLA concentrations.

[00351] Since under these conditions $[label] < [HLA]$ and $IC_{50} \geq [HLA]$, the measured IC_{50} values are reasonable approximations of the true K_D values. Peptide inhibitors are typically tested at concentrations ranging from 120 μ g/ml to 1.2 ng/ml, and are tested in two to four completely independent experiments. To allow comparison of the data obtained in different experiments, a relative binding figure is calculated for each peptide by dividing the IC_{50} of a positive control for inhibition by the IC_{50} for each tested peptide (typically unlabeled versions of the radiolabeled probe peptide). For inter-experiment comparisons, relative binding values are compiled. These values can subsequently be converted back into IC_{50} nM values by dividing the IC_{50} nM of the positive controls for inhibition by the relative binding of the peptide of interest. This method of data compilation has proven to be the most accurate and consistent for comparing peptides that have been tested on different days, or with different lots of purified MHC.

[00352] Because the antibody used for HLA-DR purification (LB3.1) is α -chain specific, β_1 molecules are not separated from β_3 (and/or β_4 and β_5) molecules. The β_1 specificity of the binding assay is obvious in the cases of DRB1*0101 (DR1), DRB1*0802 (DR8w2), and DRB1*0803 (DR8w3), where no β_3 is expressed. It has also been demonstrated for DRB1*0301 (DR3) and DRB3*0101 (DR52a), DRB1*0401 (DR4w4), DRB1*0404 (DR4w14), DRB1*0405 (DR4w15), DRB1*1101 (DR5), DRB1*1201 (DR5w12), DRB1*1302 (DR6w19) and DRB1*0701 (DR7). The problem of β chain specificity for DRB1*1501 (DR2w2 β_1), DRB5*0101 (DR2w2 β_2), DRB1*1601 (DR2w21 β_1), DRB5*0201 (DR51Dw21), and DRB4*0101 (DRw53) assays is circumvented by the use

of fibroblasts. Development and validation of assays with regard to DR β molecule specificity have been described previously (*see, e.g., Southwood et al., J. Immunol.* 160:3363-3373, 1998).

[00353] Binding assays as outlined above may be used to analyze supermotif and/or motif-bearing epitopes.

EXAMPLE 2. RECOGNITION OF VARIANT PEPTIDES BY CTL DERIVED FROM DNA IMMUNIZATION

[00354] Variants corresponding to five HLA-A2 and -A3 restricted epitopes from 167 HIV variant were identified and synthesized. These represented all the complete sequences in the Los Alamos database at the time (116 strains), as well as 51 complete clade C sequences from Botswana, and included 22 subtype B and 62 subtype C sequences. These peptides were then characterized with regard to MHC binding, variant distribution, and immunogenicity. To measure immunogenicity, HLA-A2/K^b or HLA-A11/K^b transgenic mice were immunized with the epitopes encoded in a DNA based format (). Eleven days after immunization, splenocytes were restimulated with either the epitope corresponding to the epitope encoded by the DNA (parent) or each of the variant peptides. After 6 days in culture, IFN- γ secretion was measured in response to the peptide used to stimulate each culture.

[00355] The data for these epitopes are shown in Figure 1. The HLA-A2-restricted epitope corresponding to the Env 134 epitope (KLTPLCVTL; Figure 1A) used as the immunogen was the form observed most often (134/167). All single anchor variants were recognized to approximately the same extent as the parent peptide. Many of the single non-anchor variants (9/13) were also recognized within 10-fold of the parent peptide. Conservative substitutions (R and Q for K; see Table 4) at position 1 (P1) were tolerated, while the non-conservative substitution (E for K; see Table 4) lowered binding and eliminated recognition. Three P4 variants were observed. Two of these (F or S for P) were recognized within 10-fold of the recognition of the parent peptide, while one substitution (Q for P) completely eliminated recognition. The binding for these peptides was not significantly different from the parent peptide, indicating that this residue may be involved in TCR recognition. Both the conservative (F for L) and non-conservative (R for L) substitutions

seen at P5 completely abrogated recognition, indicating that this residue is important in TCR recognition. Finally, one substitution at P8 (I for V), and four substitutions at P9 show little effect on recognition. None of the variants with multiple substitutions were recognized, although this may be due to the poor binding of these peptides.

[00356] The Gag 386 sequence utilized as the immunogen was the second most common form (VLAEAMSQV), present in 54 strains (Figure 1B). The most prevalent variant, differing by a single tolerated C terminal anchor residue (V to A; 67 strains), was recognized equally to the parent epitope by CTL raised against the parent, as were the remaining single-anchor variants. Single substitutions were also tolerated at the non-anchor positions, P1 (I for V) and P8 (R, K, or H for Q). Only the P7 variant (G for S), probably a TCR contact residue, was not recognized.

[00357] Many of the multiple variants for Gag 386 were also recognized by CTL raised against the parent peptide. All the variants with multiple changes combined a change of V to A or T at the C terminus with 1-3 additional substitutions. Two variants with N terminal changes (V to A or I) were observed. The non-conservative A substitution was not recognized, while the conservative I substitution was. A double variant with a conservative substitution at P3 (A to G) was not recognized, implicating P3 in TCR recognition. Double variants with conservative changes at position 8 (Q to R, K, or H) were not well recognized, although the variants with single changes at the same positions were recognized. The variant combining a non-conservative A residue at position 8 with A at the C terminus was recognized as well as the parent. Equally surprising was the observation that all the variants with 3 or 4 substitutions were recognized within 10-fold of the parent peptide.

[00358] The parent form of the HLA-A2-restricted epitope, Vpr 62 (RILQQLLFI; Figure 1C) was the most common form observed (86/167). Seven well-tolerated single anchor substitutions, 4 P2 and 3 C terminal, were also observed, accounting for most of the remaining variants (47/167). Single substitutions were, in general, also well tolerated. The single exception was the non-conservative substitution (P for L) at P6, while an M for L substitution at the same site was well tolerated. Binding was not affected for either variant, indicating that the reduction in activity is due to a change in a contact residue. Most variants with multiple changes also showed recognition to approximately the same extent as the parent. Several variants however did show reduced recognition. The variant with changes at both anchors (I to T at P2 and I to T at P9) had reduced binding (IC_{50} of 9700),

and recognition of the peptide was reduced, although not lost completely. Two variants with Q to H changes at P5, in combination with anchor residue changes (I to M at P2 and I to A at P9), exhibited greatly reduced recognition although binding was not affected. Other changes at P5 (Q to R or L at P5) reduced recognition only slightly.

[00359] The HLA-A3/11-restricted epitope, Pol 98 (Figure 1D), represented the most diverse epitope in terms of the number of variant epitopes identified. The peptide encoded in the DNA was represented in only 18 out of 167 strains. Approximately a third of the peptides identified at that position (49 out of 167) did not have recognizable A3/A11 motifs. The most common variant (30 strains) differed from the parent peptide at 3 residues (VSIKVGQIK), but was recognized within 10-fold of the parent peptide. Two variants with conservative changes at anchor residues were both recognized, although the T to A substitution at P2 resulted in a 10-fold reduction in recognition of the variant peptide. All peptides with single changes in non-anchor positions were also recognized, although the P5 variant (G to E) exhibited a decrease in recognition. As the binding was not affected, this probably indicates involvement in T cell recognition.

[00360] Peptides with two changes showed mixed results. In general, peptides with a V substitution at position 3, in combination with another substitution were recognized to the same extent as the corresponding single substitution, indicating the V substitution was tolerated well and is not a TCR contact residue. Combinations including the P2 anchor residue (T to A or N) were not recognized, although the binding of these peptides was also low. Variants with 3 substitutions were generally not recognized well. Two exceptions with very conservative substitutions were noted (Figure 1D). CTL were unable to recognize peptides with four or more substitutions.

[00361] The HLA-A3/11- restricted Env 47 epitope (Figure 1E; VTVYYGVPVWK) was highly conserved, with only 9 variants identified. The most common form observed was the parent peptide (99 strains), while the second most common form, a single anchor substitution observed in 40 strains, was recognized to the same extent as the parent. All the variants were recognized within 10-fold of the parent epitope.

[00362] Taken together, these data show trends towards promiscuous recognition of variant peptides by CTL generated from immunization with a single peptide. In general, changes that disrupted binding also decreased recognition. Recognition was also affected by the position of the change, with potential TCR contact residues (P3-7) exerting a greater effect on recognition than other residues. In general, conservative residue changes were more

widely tolerated than were non-conservative changes. Recognition was also dependent on the number of changes, with progressively lower recognition with a greater number of changes.

[00363] Recognition after multiple restimulations The observed recognition of variant peptides by CTL raised against the parent peptide might be due to either promiscuous recognition at the level of a single TCR or simply a mixture of TCRs against the immunizing peptide which are each able to recognize subtly different peptides. To distinguish between these two possibilities, Env 134- or Gag 386-specific T cell lines were generated by stimulating five times with the immunizing peptide, and then tested for recognition of a partial panel of variant peptides. These T cell lines were also characterized for V β TCR usage against a panel of antibodies predicted to react with the TCR of the mouse strains utilized for these experiments.

[00364] The data for these peptide-specific lines are shown in Table 5. Because the SU is a measure of the number of cells needed to secrete a defined amount of IFN- γ , a higher SU value would correspond to an enrichment of IFN- γ producing cells. A comparison of one and five peptide stimulations indeed shows an enrichment of CTL specific for the immunizing peptide for both of the peptide lines generated (Table 5A and 5B, first line). The Gag 386 line (Table 5A) also demonstrated increased recognition of all the variant peptides measured except one peptide (ILAEAMSKA) that was never recognized. The Env 134 line also demonstrated enrichment for CTL able to recognize several of the variant peptides (Table 5B).

[00365] To further characterize these lines, we examined them for V β usage, utilizing a panel of commercially available antibodies available for mouse TCR V β 2-14. To determine background levels for the various TCR V β molecules, primary splenocytes from mice that had been immunized with EP HIV-1090 were also examined. The results for the Gag 386 line are shown in Figure 2A. After a single stimulation with the parent peptide, the Gag 386 line showed a mixture of TCR positive populations, including V β 3, 5, and 14. After 5 stimulations, those populations had been reduced to background levels, and approximately 50% of the CD8⁺ cells expressed the V β 6 TCR. The Env 134 line showed a similar pattern of multiple TCR positive populations after a single round of stimulation with reduction to background levels after 5 stimulations (data not shown). However, no

single V β usage significantly above background could be demonstrated, probably due to lack of the relevant TCR V β antibody.

[00366] Both lines were also characterized with regard to the affinity of certain of the variant peptides by titrating the variant peptides examined above (Table 5A and 5B). The data for both the Gag 386 and Env 134 lines are shown in Figure 2B. For the Gag 386 line, the parent peptide along with two single anchor variants (VLAEAMSQI and VLAEAMSQA) showed the highest affinity. Four other peptides demonstrated lower affinity, but still produced IFN- γ in response to higher peptide concentrations. A single peptide (ILAEAMSKA) was not recognized.

[00367] As expected, the parent peptide, which was used to generate the Env 134 line, showed the highest affinity for the TCR. The other 2 variant peptides, KITPLCVTL and QLTPLCVTL, also demonstrated higher affinity, but reduced from the parent peptide by approximately 10-fold and 100-fold, respectively. It was notable that only at the highest peptide concentration examined (1 μ g/ml) was any IFN- γ secretion detected for five of the peptides (QITPLCVTL, ELTPLCVTL, KLTPFCVTL, KLTPLCVIL, and KLTPLCVPL). These five peptides showed little or no enrichment of CTL able to recognize them, and exhibited the lowest activity as measured by SU after five restimulations (see Table 5B).

[00368] In summary, these cell lines seem to consist of a narrow, possibly single, TCR population. This TCR population recognizes the parent peptide with the highest affinity, but is also able to recognize a number of other variant peptides with equal or lesser affinity.

[00369] **Recognition of variant peptides by CTL derived from an HIV infected patient.**

[00370] To determine if the same immunological conservation was observed in natural infections, we identified an HIV-infected individual expressing the HLA-A3 allele. The HIV strain and subtype with which this patient was infected is unknown. We had previously shown that T cells from this individual responded to the HLA-A3 restricted epitopes Pol 98 and Env 47. PBL from this patient were examined in an ELISPOT assay to determine if they also showed the capacity for broad cross-reactivity. The data are shown in Figure 3. Although the actual peptide represented in the HIV strain with which this individual is infected is unknown, we observed recognition of a large number of the variant peptides for both Pol 98 (Figure 3A) and Env 47 (Figure 3B). The recognition patterns were remarkably similar for the mouse and patient data (compare Figure 1 and

Figure 3), although the mouse expressed a transgene for HLA-A11 and the patient was HLA-A3.

[00371] **Prediction of Immunological Conservation.** We had observed that the variant peptides that were recognized by CTL raised against the parent epitope had amino acid substitutions that followed previous observations. For example, the anchor residue changes that were tolerated in the variant peptides were also described as anchors that to define the respective HLA supertypes (). In general, conservative substitutions were tolerated at non-anchor residues, while non-conservative substitutions were less well tolerated. These followed closely the prediction model used to identify heteroclitic analogs (Tangri et al).

[00372] Based on these observations, we designed a computer program to predict immunological conservation. For anchor positions, this program utilized the conserved anchor residues described for the A2, A3, and B7 supertypes. For non-anchor positions only conservative substitutions, as defined in Tangri et.al. (), were allowed. All substitutions at non-anchor positions were analyzed independently and all conservative substitutions were allowed regardless of the number of substitutions. Finally, the position of the substitution was not factored into analysis. Each variant was compared with the parent epitope, and its ability to be recognized was predicted as either positive or negative.

[00373] The first sets of epitopes to be evaluated by this program were the five HIV epitopes and variants previously described. For the Env 134 epitope, the program predicted that 13 of the variant peptides should be immunologically conserved, while 6 should not be recognized. Comparison of the observed immunological data with the prediction showed that the program predicted correctly for 14 of the peptides and incorrectly for 5. Of the incorrect predictions, in two cases the program predicted negative results for peptides that were recognized, while in 3 cases the program predicted positive results for peptides that were not recognized. A similar analysis was performed for all five peptides. Of 101 total variant peptides, 68 were correctly identified (67%). The discordant data were fairly evenly split between peptides incorrectly predicted negative (15) and those incorrectly predicted positive (18).

[00374] As noted previously, the more substitutions present in a variant peptide, the lower the likelihood of its immunogenicity. Since the prediction program treated all substitutions independently, and did not take into account the number of substitutions, we hypothesized

that prediction of single substitutions would be more accurate. Indeed, the immunogenicity of 38 of 47 single substitution variants (80%) was correctly predicted.

[00375] With the limitations of the program in mind, it is useful to predict the recognition of the variants for a package of HLA-A2, -A3, and -B7 supertype epitopes. These epitopes had been identified as being well conserved in Clade B variants. When comparing the conservation of this group of epitopes based on sequence identity versus immunological conservation, it is interesting to note that the predicted recognition gains taking into account immunological conservation are significant (Table 6).

[00376] This particular group of 21 epitopes was selected based on their identity conservation in Clade B HIV sequences, with conservation across HIV clades as a secondary consideration. Because of this criteria, the form of epitope chosen as the parent peptide was not the most common variant (e.g. Gag 386, Gag 271, Pol 98). In some cases (e.g., see Gag 386 data), the "parent" epitope and the most common variant were recognized to the same extent. However, in some cases the selection of epitope to include as the "parent" epitope was predicted to make a difference in the immunological conservation. An example of this was the Gag 271 epitope (Figure 4). The variant most commonly seen in clade B sequences was the MTNNPPIP form, while the most common form of the epitope was MTSNPPIP. Not all amino acids are considered equal to each other in their ability to substitute (Tangri). For example, asparagine (N) is considered a conservative substitution for serine (S), while the opposite substitution is only considered semi-conserved. When the program calculated immunological conservation using the MTNNPPIP peptide as the parent peptide, only two variants were predicted to be immunogenic. However, when the immunological conservation was predicted using the MTSNPPIP peptide, most of the variants were predicted to be recognized (Figure 4). This prediction was tested using HLA-A2 transgenic mice. The results show that if the MTSNPPIP form of the peptide was utilized in vaccines, approximately 152 of 167 variants would be recognized, while if the MTNNPPIP form of the epitope was utilized, only 39 of 167 variants would be recognized. This has important implications in epitope selection for vaccine development, and epitope performance can be predicted.

EXAMPLE 3. A PADRE[®] MOLECULE AS A HELPER EPTOPE FOR ENHANCEMENT OF CTL INDUCTION

[00377] There is increasing evidence that HTL activity is critical for the induction of long lasting CTL responses (Livingston *et al. J. Immunol* 162:3088-3095 (1999); Walter *et al., New Engl. J. Med.* 333:1038-1044 (1995); Hu *et al., J. Exp. Med.* 177:1681-1690 (1993)). Therefore, one or more peptides that bind to HLA class II molecules and stimulate HTLs can be used in accordance with the invention. Accordingly, a preferred embodiment of a vaccine includes a molecule from the PADRE[®] family of universal T helper cell epitopes (HTL) that target most DR molecules in a manner designed to stimulate helper T cells. For instance, a pan-DR-binding epitope peptide having the formula: aKXVAAZTLKAAa, where "X" is either cyclohexylalanine, phenylalanine, or tyrosine; "Z" is either tryptophan, tyrosine, histidine or asparagine; and "a" is either D-alanine or L-alanine (SEQ ID NO:29), has been found to bind to most HLA-DR alleles, and to stimulate the response of T helper lymphocytes from most individuals, regardless of their HLA type.

[00378] A particularly preferred PADRE[®] molecule is a synthetic peptide, aKXVAAWTLKAAa (a = D-alanine, X = cyclohexylalanine), containing non-natural amino acids, specifically engineered to maximize both HLA-DR binding capacity and induction of T cell immune responses.

[00379] Alternative preferred PADRE[®] molecules are the peptides, aKFVAAWTLKAAa, aKYVAAWTLKAAa, aKFVAAZTLKAAa, aKXVAAZTLKAAa, aKYVAAZTLKAAa, aKFVAAHTLKAAa, aKXVAAHTLKAAa, aKYVAAHTLKAAa, aKFVAAANTLKAAa, aKXVAAANTLKAAa, aKYVAAANTLKAAa, AKXVAAWTLKAAA (SEQ ID NO:30), AKFVAAWTLKAAA (SEQ ID NO:31), AKYVAAWTLKAAA (SEQ ID NO:32), AKFVAAZTLKAAA (SEQ ID NO:33), AKXVAAZTLKAAA (SEQ ID NO:34), AKYVAAZTLKAAA (SEQ ID NO:35), AKFVAAHTLKAAA (SEQ ID NO:36), AKXVAAHTLKAAA (SEQ ID NO:37), AKYVAAHTLKAAA (SEQ ID NO:38), AKFVAAANTLKAAA (SEQ ID NO:39), AKXVAAANTLKAAA (SEQ ID NO:40), AKYVAAANTLKAAA (SEQ ID NO:41) (a = D-alanine, X = cyclohexylalanine).

[00380] In a preferred embodiment, the PADRE[®] peptide is amidated. For example, a particularly preferred amidated embodiment of a PADRE[®] molecule is conventionally written aKXVAAWTLKAAa-NH₂.

[00381] Competitive inhibition assays with purified HLA-DR molecules demonstrated that the PADRE[®] molecule aKXVAAWTLKAAa-NH₂ binds with high or intermediate affinity (IC₅₀ ≤1,000 nM) to 15 out of 16 of the most prevalent HLA-DR molecules ((Kawashima *et al.*, *Human Immunology* 59:1-14 (1998); Alexander *et al.*, *Immunity* 1:751-761 (1994)). A comparison of the DR binding capacity of PADRE[®] and tetanus toxoid (TT) peptide 830-843, a "universal" epitope has been published (Panina-Bordignon *et al.*, *Eur. J. Immunology* 19:2237-2242 (1989)). The TT 830-843 peptide bound to only seven of 16 DR molecules tested, while PADRE[®] bound 15 of 16. At least 1 of the 15 DR molecules that bind PADRE[®] is predicted to be present in >95% of all humans. Therefore, this PADRE[®] molecule is anticipated to induce an HTL response in virtually all patients, despite the extensive polymorphism of HLA-DR molecules in the human population.

[00382] PADRE[®] has been specifically engineered for optimal immunogenicity for human T cells. Representative data from *in vitro* primary immunizations of normal human T cells with TT 830-843 antigen and the PADRE[®] molecule aKXVAAWTLKAAa-NH₂ are shown in Figure 1. Peripheral blood mononuclear cells (PBMC) from three normal donors were stimulated with the peptides *in vitro*. Following the third round of stimulation, it was observed that PADRE[®] generated significant primary T cell responses for all three donors as measured in a standard T cell proliferation assay. With the PADRE[®] peptide, the 10,000 cpm proliferation level was generally reached with 10 to 100 ng/ml of antigen. In contrast, TT 830-843 antigen generated responses for only 2 out of 3 of the individuals tested. Responses approaching the 10,000 cpm range were reached with about 10,000 ng/ml of antigen. In this respect, it was noted that PADRE[®] was, on a molar basis, about 100-fold more potent than TT 830-843 antigen for activation of T cell responses.

[00383] Early data from a phase I/II investigator-sponsored trial, conducted at the University of Leiden (C.J.M. Melief), support the principle that the PADRE[®] molecule aKXVAAWTLKAAa, possibly the amidated aKXVAAWTLKAAa -NH₂, is highly immunogenic in humans (Ressing *et al.*, *J. Immunother.* 23(2):255-66 (2000)). In this trial, a PADRE[®] molecule was co-emulsified with various human papilloma virus (HPV)-derived CTL epitopes and was injected into patients with recurrent or residual cervical carcinoma. However, because of the late stage of carcinoma with the study patients, it was expected that these patients were immunocompromised. The patients' immunocompromised status was demonstrated by their low frequency of influenza virus-

specific CTL, reduced levels of CD3 expression, and low incidence of proliferative recall responses after *in vitro* stimulation with conventional antigens. Thus, no efficacy was anticipated in the University of Leiden trial, rather the goal of that trial was essentially to evaluate safety. Safety was, in fact, demonstrated. In addition to a favorable safety profile, PADRE[®] T cell reactivity was detected in four of 12 patients (Figure 2) in spite of the reduced immune competence of these patients.

[00384] Thus, the PADRE[®] peptide component(s) of the vaccine bind with broad specificity to multiple allelic forms of HLA-DR molecules. Moreover, PADRE[®] peptide component(s) bind with high affinity ($IC_{50} \leq 1000$ nM), i.e., at a level of affinity correlated with being immunogenic for HLA Class II restricted T cells. The *in vivo* administration of PADRE[®] peptide(s) stimulates the proliferation of HTL in normal humans as well as patient populations.

[00385] One or more PADRE[®] peptide(s) may be included in a composition, e.g., a vaccine, comprising one or more peptides, either as an individual peptide(s), fused to one or more variant peptides, or both.

EXAMPLE 4. CTL RECOGNITION OF ENDOGENOUS PROCESSED ANTIGENS AFTER PRIMING

[00386] This example determines that CTL induced by native or analoged peptide epitopes recognize endogenously synthesized, *i.e.*, native antigens.

[00387] Effector cells isolated from transgenic mice that are immunized with peptide epitopes are re-stimulated *in vitro* using peptide-coated stimulator cells. Six days later, effector cells are assayed for cytotoxicity and the cell lines that contain peptide-specific cytotoxic activity are further re-stimulated. An additional six days later, these cell lines are tested for cytotoxic activity on ⁵¹Cr labeled Jurkat-A2.1/K^b target cells in the absence or presence of peptide, and also tested on ⁵¹Cr labeled target cells bearing the endogenously synthesized antigen, *i.e.* cells that are stably transfected with HIV expression vectors.

[00388] The result will demonstrate that CTL lines obtained from animals primed with peptide epitope recognize endogenously synthesized HIV antigen. The choice of transgenic mouse model to be used for such an analysis depends upon the epitope(s) that is being evaluated. In addition to HLA-A*0201/K^b transgenic mice, several other transgenic

mouse models including mice with human A11, which may also be used to evaluate A3 epitopes, and B7 alleles have been characterized and others (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed, which may be used to evaluate HTL epitopes.

EXAMPLE 5. ACTIVITY OF CTL-HTL CONJUGATED EPITOPES IN TRANSGENIC MICE

[00389] This example illustrates the induction of CTLs and HTLs in transgenic mice by use of a HIV CTL/HTL peptide conjugate whereby the vaccine composition comprises peptides administered to an HIV-infected patient or an individual at risk for HIV. The peptide composition can comprise multiple CTL and/or HTL epitopes. This analysis demonstrates enhanced immunogenicity that can be achieved by inclusion of one or more HTL epitopes in a vaccine composition. Such a peptide composition can comprise an HTL epitope conjugated to a preferred CTL epitope containing, for example, at least one CTL epitope, or an analog of that epitope. The peptides may be lipidated, if desired.

[00390] Immunization procedures: Immunization of transgenic mice is performed as described (Alexander *et al.*, *J. Immunol.* 159:4753-4761, 1997). For example, A2/K^b mice, which are transgenic for the human HLA A2.1 allele and are useful for the assessment of the immunogenicity of HLA-A*0201 motif- or HLA-A2 supermotif-bearing epitopes, are primed subcutaneously (base of the tail) with a 0.1 ml of peptide in Incomplete Freund's Adjuvant, or if the peptide composition is a lipidated CTL/HTL conjugate, in DMSO/saline or if the peptide composition is a polypeptide, in PBS or Incomplete Freund's Adjuvant. Seven days after priming, splenocytes obtained from these animals are restimulated with syngenic irradiated LPS-activated lymphoblasts coated with peptide.

[00391] Cell lines: Target cells for peptide-specific cytotoxicity assays are Jurkat cells transfected with the HLA-A2.1/K^b chimeric gene (*e.g.*, Vitiello *et al.*, *J. Exp. Med.* 173:1007, 1991).

[00392] *In vitro* CTL activation: One week after priming, spleen cells (30x10⁶ cells/flask) are co-cultured at 37°C with syngeneic, irradiated (3000 rads), peptide coated

lymphoblasts (10×10^6 cells/flask) in 10 ml of culture medium/T25 flask. After six days, effector cells are harvested and assayed for cytotoxic activity.

[00393] Assay for cytotoxic activity: Target cells (1.0 to 1.5×10^6) are incubated at 37°C in the presence of $200\ \mu\text{l}$ of ^{51}Cr . After 60 minutes, cells are washed three times and resuspended in R10 medium. Peptide is added where required at a concentration of $1\ \mu\text{g/ml}$. For the assay, $10^4\ ^{51}\text{Cr}$ -labeled target cells are added to different concentrations of effector cells (final volume of $200\ \mu\text{l}$) in U-bottom 96-well plates. After a 6 hour incubation period at 37°C , a $0.1\ \text{ml}$ aliquot of supernatant is removed from each well and radioactivity is determined in a Micromedic automatic gamma counter. The percent specific lysis is determined by the formula: percent specific release = $100 \times (\text{experimental release} - \text{spontaneous release}) / (\text{maximum release} - \text{spontaneous release})$. To facilitate comparison between separate CTL assays run under the same conditions, % ^{51}Cr release data is expressed as lytic units/ 10^6 cells. One lytic unit is arbitrarily defined as the number of effector cells required to achieve 30% lysis of 10,000 target cells in a 6 hour ^{51}Cr release assay. To obtain specific lytic units/ 10^6 , the lytic units/ 10^6 obtained in the absence of peptide is subtracted from the lytic units/ 10^6 obtained in the presence of peptide. For example, if 30% ^{51}Cr release is obtained at the effector (E): target (T) ratio of 50:1 (i.e., 5×10^5 effector cells for 10,000 targets) in the absence of peptide and 5:1 (i.e., 5×10^4 effector cells for 10,000 targets) in the presence of peptide, the specific lytic units would be: $[(1/50,000) - (1/500,000)] \times 10^6 = 18\ \text{LU}$.

[00394] The results are analyzed to assess the magnitude of the CTL responses of animals injected with the immunogenic CTL/HTL conjugate vaccine preparation and are compared to the magnitude of the CTL response achieved using the CTL epitope as outlined in above. Analyses similar to this may be performed to evaluate the immunogenicity of peptide conjugates containing multiple CTL epitopes and/or multiple HTL epitopes. In accordance with these procedures it is found that a CTL response is induced, and concomitantly that an HTL response is induced upon administration of such compositions.

EXAMPLE 6. SELECTION OF CTL AND HTL EPITOPES FOR INCLUSION IN AN HIV-SPECIFIC VACCINE.

[00395] This example illustrates the procedure for the selection of peptide epitopes for vaccine compositions of the invention. The peptides in the composition can be in the form of a nucleic acid sequence, either single or one or more sequences (*i.e.*, minigene) that encodes peptide(s), or can be single and/or polyepitopic peptides.

[00396] The following principles are utilized when selecting an array of epitopes for inclusion in a vaccine composition. Each of the following principles is balanced in order to make the selection.

[00397] Epitopes are selected which, upon administration, mimic immune responses that correlate with virus clearance. For example, if it has been observed that patients who clear HIV generate an immune response to at least 3 epitopes on at least one HIV antigen, then 3-4 epitopes should be included for HLA class I. A similar rationale is used to determine HLA class II epitopes.

[00398] When selecting an array of HIV epitopes, it is preferred that at least some of the epitopes are derived from early and late proteins. The early proteins of HIV are expressed when the virus is replicating, either following acute or dormant infection. Therefore, it is particularly preferred to use epitopes from early stage proteins to alleviate disease manifestations at the earliest stage possible.

[00399] Epitopes are often selected that have a binding affinity of an IC_{50} of 500 nM or less for an HLA class I molecule, or for class II, an IC_{50} of 1000 nM or less.

[00400] Sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. For example, epitopes are selected to provide at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess breadth, or redundancy, of population coverage.

[00401] When creating a polyepitopic compositions, *e.g.* a minigene, it is typically desirable to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same, as those employed when selecting a peptide comprising nested epitopes.

[00402] In cases where the sequences of multiple variants of the same target protein are available, potential peptide epitopes can also be selected on the basis of their conservancy.

For example, a criterion for conservancy may define that the entire sequence of an HLA class I binding peptide or the entire 9-mer core of a class II binding peptide be conserved in a designated percentage of the sequences evaluated for a specific protein antigen.

[00403] Peptide epitopes for inclusion in vaccine compositions are, for example, selected from those listed in Tables 6-9 or Figures 1A-4. A vaccine composition comprised of selected peptides, when administered, is safe, efficacious, and elicits an immune response similar in magnitude of an immune response that clears an acute HIV infection.

EXAMPLE 7. CONSTRUCTION OF MINIGENE MULTI-EPITOPE DNA PLASMIDS

[00404] This example provides general guidance for the construction of a minigene expression plasmid. Minigene plasmids may, of course, contain various configurations of CTL and/or HTL epitopes or epitope analogs as described herein. Expression plasmids have been constructed and evaluated as described, for example, in co-pending U.S.S.N. 09/311,784 filed 5/13/99 and in Ishioka *et al.*, *J. Immunol.* 162:3915-3925, 1999. An example of such a plasmid for the expression of HIV epitopes is shown in Figure 2, which illustrates the orientation of HIV peptide epitopes in a minigene construct.

[00405] A minigene expression plasmid typically includes multiple CTL and HTL peptide epitopes. In the present example, HLA-A2, -A3, -B7 supermotif-bearing peptide epitopes and HLA-A1 and -A24 motif-bearing peptide epitopes are used in conjunction with DR supermotif-bearing epitopes and/or DR3 epitopes (Figure 2). Preferred epitopes are identified, for example, in Tables 6-9 and Figures 1A-4. HLA class I supermotif or motif-bearing peptide epitopes derived from multiple HIV antigens, are selected such that multiple supermotifs/motifs are represented to ensure broad population coverage. Similarly, HLA class II epitopes are selected from multiple HIV antigens to provide broad population coverage, *i.e.* both HLA DR-1-4-7 supermotif-bearing epitopes and HLA DR-3 motif-bearing epitopes are selected for inclusion in the minigene construct. The selected CTL and HTL epitopes are then incorporated into a minigene for expression in an expression vector.

[00406] Such a construct may additionally include sequences that direct the HTL epitopes to the endoplasmic reticulum. For example, the Ii protein may be fused to one or more HTL epitopes as described in co-pending application U.S.S.N. 09/311,784 filed 5/13/99,

wherein the CLIP sequence of the Ii protein is removed and replaced with an HLA class II epitope sequence so that HLA class II epitope is directed to the endoplasmic reticulum, where the epitope binds to an HLA class II molecules.

[00407] This example illustrates the methods to be used for construction of a minigene-bearing expression plasmid. Other expression vectors that may be used for minigene compositions are available and known to those of skill in the art.

[00408] The minigene DNA plasmid contains a consensus Kozak sequence and a consensus murine kappa Ig-light chain signal sequence followed by CTL and/or HTL epitopes selected in accordance with principles disclosed herein. The construct can also include, for example, The sequence encodes an open reading frame fused to the Myc and His antibody epitope tag coded for by the pcDNA 3.1 Myc-His vector.

[00409] Overlapping oligonucleotides, for example eight oligonucleotides, averaging approximately 70 nucleotides in length with 15 nucleotide overlaps, are synthesized and HPLC-purified. The oligonucleotides encode the selected peptide epitopes as well as appropriate linker nucleotides, Kozak sequence, and signal sequence. The final multiepitope minigene is assembled by extending the overlapping oligonucleotides in three sets of reactions using PCR. A Perkin/Elmer 9600 PCR machine is used and a total of 30 cycles are performed using the following conditions: 95°C for 15 sec, annealing temperature (5° below the lowest calculated T_m of each primer pair) for 30 sec, and 72°C for 1 min.

[00410] For the first PCR reaction, 5 µg of each of two oligonucleotides are annealed and extended: Oligonucleotides 1+2, 3+4, 5+6, and 7+8 are combined in 100 µl reactions containing *Pfu* polymerase buffer (1x= 10 mM KCL, 10 mM (NH₄)₂SO₄, 20 mM Tris-chloride, pH 8.75, 2 mM MgSO₄, 0.1% Triton X-100, 100 µg/ml BSA), 0.25 mM each dNTP, and 2.5 U of *Pfu* polymerase. The full-length dimer products are gel-purified, and two reactions containing the product of 1+2 and 3+4, and the product of 5+6 and 7+8 are mixed, annealed, and extended for 10 cycles. Half of the two reactions are then mixed, and 5 cycles of annealing and extension carried out before flanking primers are added to amplify the full length product for 25 additional cycles. The full-length product is gel-purified and cloned into pCR-blunt (Invitrogen) and individual clones are screened by sequencing.

EXAMPLE 8. THE PLASMID CONSTRUCT AND THE DEGREE TO WHICH IT
INDUCES IMMUNOGENICITY.

[00411] The degree to which a plasmid construct, for example a plasmid constructed in accordance as above is able to induce immunogenicity can be evaluated *in vitro* by testing for epitope presentation by APC following transduction or transfection of the APC with an epitope-expressing nucleic acid construct. Such a study determines “antigenicity” and allows the use of human APC. The assay determines the ability of the epitope to be presented by the APC in a context that is recognized by a T cell by quantifying the density of epitope-HLA class I complexes on the cell surface. Quantitation can be performed by directly measuring the amount of peptide eluted from the APC (*see, e.g.,* Sijts *et al.*, *J. Immunol.* 156:683-692, 1996; Demotz *et al.*, *Nature* 342:682-684, 1989); or the number of peptide-HLA class I complexes can be estimated by measuring the amount of lysis or lymphokine release induced by infected or transfected target cells, and then determining the concentration of peptide necessary to obtained equivalent levels of lysis or lymphokine release (*see, e.g.,* Kageyama *et al.*, *J. Immunol.* 154:567-576, 1995).

[00412] Alternatively, immunogenicity can be evaluated through *in vivo* injections into mice and subsequent *in vitro* assessment of CTL and HTL activity, which are analysed using cytotoxicity and proliferation assays, respectively, as detailed *e.g.,* in copending U.S.S.N. 09/311,784 filed 5/13/99 and Alexander *et al.*, *Immunity* 1:751-761, 1994.

[00413] For example, to assess the capacity of a DNA minigene construct (*e.g.,* a pMin minigene construct generated as described in U.S.S.N. 09/311,784) containing at least one HLA-A2 supermotif peptide to induce CTLs *in vivo*, HLA-A2.1/K^b transgenic mice, for example, are immunized intramuscularly with 100 µg of naked cDNA. As a means of comparing the level of CTLs induced by cDNA immunization, a control group of animals is also immunized with an actual peptide composition that comprises multiple epitopes synthesized as a single polypeptide as they would be encoded by the minigene.

[00414] Splenocytes from immunized animals are stimulated twice with each of the respective compositions (peptide epitopes encoded in the minigene or the polyepitopic peptide), then assayed for peptide-specific cytotoxic activity in a ⁵¹Cr release assay. The results indicate the magnitude of the CTL response directed against the A2-restricted epitope, thus indicating the *in vivo* immunogenicity of the minigene vaccine and polyepitopic vaccine. It is, therefore, found that the minigene elicits immune responses

directed toward the HLA-A2 supermotif peptide epitopes as does the polyepitopic peptide vaccine. A similar analysis is also performed using other HLA-A3 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 and HLA-B7 motif or supermotif epitopes.

[00415] To assess the capacity of a class II epitope encoding minigene to induce HTLs *in vivo*, DR transgenic mice, or for those epitope that cross react with the appropriate mouse MHC molecule, I-A^b-restricted mice, for example, are immunized intramuscularly with 100 µg of plasmid DNA. As a means of comparing the level of HTLs induced by DNA immunization, a group of control animals is also immunized with an actual peptide composition emulsified in complete Freund's adjuvant. CD4⁺ T cells, *i.e.* HTLs, are purified from splenocytes of immunized animals and stimulated with each of the respective compositions (peptides encoded in the minigene). The HTL response is measured using a ³H-thymidine incorporation proliferation assay, (*see, e.g.*, Alexander et al. *Immunity* 1:751-761, 1994). The results indicate the magnitude of the HTL response, thus demonstrating the *in vivo* immunogenicity of the minigene.

[00416] DNA minigenes, constructed as described above or below, may also be evaluated as a vaccine in combination with a boosting agent using a prime boost protocol. The boosting agent can consist of recombinant protein (*e.g.*, Barnett *et al.*, *Aids Res. and Human Retroviruses* 14, Supplement 3:S299-S309, 1998) or recombinant vaccinia, for example, expressing a minigene or DNA encoding the complete protein of interest (*see, e.g.*, Hanke *et al.*, *Vaccine* 16:439-445, 1998; Sedegah *et al.*, *Proc. Natl. Acad. Sci USA* 95:7648-53, 1998; Hanke and McMichael, *Immunol. Letters* 66:177-181, 1999; and Robinson *et al.*, *Nature Med.* 5:526-34, 1999).

[00417] For example, the efficacy of the DNA minigene used in a prime boost protocol is initially evaluated in transgenic mice. In this example, A2.1/K^b transgenic mice are immunized IM with 100 µg of a DNA minigene encoding the immunogenic peptides including at least one HLA-A2 supermotif-bearing peptide. After an incubation period (ranging from 3-9 weeks), the mice are boosted IP with 10⁷ pfu/mouse of a recombinant vaccinia virus expressing the same sequence encoded by the DNA minigene. Control mice are immunized with 100 µg of DNA or recombinant vaccinia without the minigene sequence, or with DNA encoding the minigene, but without the vaccinia boost. After an additional incubation period of two weeks, splenocytes from the mice are immediately assayed for peptide-specific activity in an ELISPOT assay. Additionally, splenocytes are

stimulated *in vitro* with the A2-restricted peptide epitopes encoded in the minigene and recombinant vaccinia, then assayed for peptide-specific activity in an IFN- γ ELISA.

[00418] It is found that the minigene utilized in a prime-boost protocol elicits greater immune responses toward the HLA-A2 supermotif peptides than with DNA alone. Such an analysis can also be performed using HLA-A11 or HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 or HLA-B7 motif or supermotif epitopes.

[00419] The use of prime boost protocols in humans is described in below.

EXAMPLE 9. PEPTIDE COMPOSITION FOR PROPHYLACTIC USES

[00420] Vaccine compositions of the present invention can be used to prevent HIV infection in persons who are at risk for such infection. For example, a polyepitopic peptide epitope composition (or a nucleic acid comprising the same) containing multiple CTL and HTL epitopes, which are also selected to target greater than 80% of the population, is administered to individuals at risk for HIV infection.

[00421] For example, a peptide-based composition can be provided as a single polypeptide that encompasses multiple epitopes. The vaccine is typically administered in a physiological solution that comprises an adjuvant, such as Incomplete Freund's Adjuvant. The dose of peptide for the initial immunization is from about 1 to about 50,000 μg , generally 100-5,000 μg , for a 70 kg patient. The initial administration of vaccine is followed by booster dosages at 4 weeks followed by evaluation of the magnitude of the immune response in the patient, by techniques that determine the presence of epitope-specific CTL populations in a PBMC sample. Additional booster doses are administered as required. The composition is found to be both safe and efficacious as a prophylaxis against HIV infection.

[00422] Alternatively, a composition typically comprising transfecting agents can be used for the administration of a nucleic acid-based vaccine in accordance with methodologies known in the art and disclosed herein.

EXAMPLE 10. POLYEPITOPIC VACCINE COMPOSITIONS DERIVED FROM
NATIVE HIV SEQUENCES

[00423] A native HIV polyprotein sequence is screened, preferably using computer algorithms defined for each class I and/or class II supermotif or motif, to identify “relatively short” regions of the polyprotein that comprise multiple epitopes and is preferably less in length than an entire native antigen. This relatively short sequence that contains multiple distinct, even overlapping, epitopes is selected and used to generate a minigene construct. The construct is engineered to express the peptide, which corresponds to the native protein sequence. The “relatively short” peptide is generally less than 250 amino acids in length, often less than 100 amino acids in length, preferably less than 75 amino acids in length, and more preferably less than 50 amino acids in length. The protein sequence of the vaccine composition is selected because it has maximal number of epitopes contained within the sequence, *i.e.*, it has a high concentration of epitopes. As noted herein, epitope motifs may be nested or overlapping, for example, two 9-mer epitopes and one 10-mer epitope can be present in a 10 amino acid peptide. Such a vaccine composition is administered for therapeutic or prophylactic purposes.

[00424] The vaccine composition will preferably include, for example, three CTL epitopes and at least one HTL epitope from HIV. This polyepitopic native sequence is administered either as a peptide or as a nucleic acid sequence which encodes the peptide. Alternatively, an analog can be made of this native sequence, whereby one or more of the epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide.

[00425] The embodiment of this example provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (absent analogs) directs the immune response to multiple peptide sequences that are actually present in native HIV antigens thus avoiding the need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of scale when producing nucleic acid vaccine compositions.

[00426] Related to this embodiment, computer programs can be derived in accordance with principles in the art, which identify in a target sequence, the greatest number of epitopes per sequence length.

EXAMPLE 11. POLYEPITOPIC VACCINE COMPOSITIONS DIRECTED TO MULTIPLE DISEASES

[00427] The HIV peptide epitopes of the present invention are used in conjunction with peptide epitopes from target antigens related to one or more other diseases, to create a vaccine composition that is useful for the prevention or treatment of HIV as well as the one or more other disease(s). Examples of the other diseases include, but are not limited to, HCV and HBV.

[00428] For example, a polyepitopic peptide composition comprising multiple CTL and HTL epitopes that target greater than 98% of the population may be created for administration to individuals at risk for both HBV and HIV infection. The composition can be provided as a single polypeptide that incorporates the multiple epitopes from the various disease-associated sources, or can be administered as a composition comprising one or more discrete epitopes.

EXAMPLE 12. USE OF PEPTIDES TO EVALUATE AN IMMUNE RESPONSE

[00429] Peptides of the invention may be used to analyze an immune response for the presence of specific CTL or HTL populations directed to HIV. Such an analysis may be performed in a manner as that described by Ogg *et al.*, *Science* 279:2103-2106, 1998. In the following example, peptides in accordance with the invention are used as a reagent for diagnostic or prognostic purposes, not as an immunogen.

[00430] In this example highly sensitive human leukocyte antigen tetrameric complexes ("tetramers") are used for a cross-sectional analysis of, for example, HIV HLA-A*0201-specific CTL frequencies from HLA A*0201-positive individuals at different stages of infection or following immunization using an HIV peptide containing an A*0201 motif. Tetrameric complexes are synthesized as described (Musey *et al.*, *N. Engl. J. Med.* 337:1267, 1997). Briefly, purified HLA heavy chain (A*0201 in this example) and β 2-

microglobulin are synthesized by means of a prokaryotic expression system. The heavy chain is modified by deletion of the transmembrane-cytosolic tail and COOH-terminal addition of a sequence containing a BirA enzymatic biotinylation site. The heavy chain, β 2-microglobulin, and peptide are refolded by dilution. The 45-kD refolded product is isolated by fast protein liquid chromatography and then biotinylated by BirA in the presence of biotin (Sigma, St. Louis, Missouri), adenosine 5'triphosphate and magnesium. Streptavidin-phycoerythrin conjugate is added in a 1:4 molar ratio, and the tetrameric product is concentrated to 1 mg/ml. The resulting product is referred to as tetramer-phycoerythrin.

[00431] For the analysis of patient blood samples, approximately one million PBMCs are centrifuged at 300 x g for 5 minutes and resuspended in 50 μ l of cold phosphate-buffered saline. Tri-color analysis is performed with the tetramer-phycoerythrin, along with anti-CD8-Tricolor, and anti-CD38. The PBMCs are incubated with tetramer and antibodies on ice for 30 to 60 min and then washed twice before formaldehyde fixation. Gates are applied to contain >99.98% of control samples. Controls for the tetramers include both A*0201-negative individuals and A*0201-positive uninfected donors. The percentage of cells stained with the tetramer is then determined by flow cytometry. The results indicate the number of cells in the PBMC sample that contain epitope-restricted CTLs, thereby readily indicating the extent of immune response to the HIV epitope, and thus the stage of infection with HIV, the status of exposure to HIV, or exposure to a vaccine that elicits a protective or therapeutic response.

EXAMPLE 13. USE OF PEPTIDE EPITOPES TO EVALUATE RECALL RESPONSES

[00432] The peptide epitopes of the invention are used as reagents to evaluate T cell responses, such as acute or recall responses, in patients. Such an analysis may be performed on patients who have recovered from infection, who are chronically infected with HIV, or who have been vaccinated with an HIV vaccine.

[00433] For example, the class I restricted CTL response of persons who have been vaccinated may be analyzed. The vaccine may be any HIV vaccine. PBMC are collected from vaccinated individuals and HLA typed. Appropriate peptide epitopes of the invention that, optimally, bear supermotifs to provide cross-reactivity with multiple HLA

supertype family members, are then used for analysis of samples derived from individuals who bear that HLA type.

[00434] PBMC from vaccinated individuals are separated on Ficoll-Histopaque density gradients (Sigma Chemical Co., St. Louis, MO), washed three times in HBSS (GIBCO Laboratories), resuspended in RPMI-1640 (GIBCO Laboratories) supplemented with L-glutamine (2mM), penicillin (50U/ml), streptomycin (50 µg/ml), and Hepes (10mM) containing 10% heat-inactivated human AB serum (complete RPMI) and plated using microculture formats. A synthetic peptide comprising an epitope of the invention is added at 10 µg/ml to each well and HBV core 128-140 epitope is added at 1 µg/ml to each well as a source of T cell help during the first week of stimulation.

[00435] In the microculture format, 4×10^5 PBMC are stimulated with peptide in 8 replicate cultures in 96-well round bottom plate in 100 µl/well of complete RPMI. On days 3 and 10, 100 ml of complete RPMI and 20 U/ml final concentration of rIL-2 are added to each well. On day 7 the cultures are transferred into a 96-well flat-bottom plate and restimulated with peptide, rIL-2 and 10^5 irradiated (3,000 rad) autologous feeder cells. The cultures are tested for cytotoxic activity on day 14. A positive CTL response requires two or more of the eight replicate cultures to display greater than 10% specific ^{51}Cr release, based on comparison with uninfected control subjects as previously described (Rehermann, *et al.*, *Nature Med.* 2:1104,1108, 1996; Rehermann *et al.*, *J. Clin. Invest.* 97:1655-1665, 1996; and Rehermann *et al.* *J. Clin. Invest.* 98:1432-1440, 1996).

[00436] Target cell lines are autologous and allogeneic EBV-transformed B-LCL that are either purchased from the American Society for Histocompatibility and Immunogenetics (ASHI, Boston, MA) or established from the pool of patients as described (Guilhot, *et al.* *J. Virol.* 66:2670-2678, 1992).

[00437] Cytotoxicity assays are performed in the following manner. Target cells consist of either allogeneic HLA-matched or autologous EBV-transformed B lymphoblastoid cell line that are incubated overnight with the synthetic peptide epitope of the invention at 10 µM, and labeled with 100 µCi of ^{51}Cr (Amersham Corp., Arlington Heights, IL) for 1 hour after which they are washed four times with HBSS.

[00438] Cytolytic activity is determined in a standard 4-h, split well ^{51}Cr release assay using U-bottomed 96 well plates containing 3,000 targets/well. Stimulated PBMC are tested at effector/target (E/T) ratios of 20-50:1 on day 14. Percent cytotoxicity is determined from the formula: $100 \times [(\text{experimental release} - \text{spontaneous$

release)/maximum release-spontaneous release)]. Maximum release is determined by lysis of targets by detergent (2% Triton X-100; Sigma Chemical Co., St. Louis, MO). Spontaneous release is <25% of maximum release for all experiments.

[00439] The results of such an analysis indicate the extent to which HLA-restricted CTL populations have been stimulated by previous exposure to HIV or an HIV vaccine.

[00440] The class II restricted HTL responses may also be analyzed. Purified PBMC are cultured in a 96-well flat bottom plate at a density of 1.5×10^5 cells/well and are stimulated with 10 $\mu\text{g/ml}$ synthetic peptide, whole antigen, or PHA. Cells are routinely plated in replicates of 4-6 wells for each condition. After seven days of culture, the medium is removed and replaced with fresh medium containing 10U/ml IL-2. Two days later, 1 μCi ^3H -thymidine is added to each well and incubation is continued for an additional 18 hours. Cellular DNA is then harvested on glass fiber mats and analyzed for ^3H -thymidine incorporation. Antigen-specific T cell proliferation is calculated as the ratio of ^3H -thymidine incorporation in the presence of antigen divided by the ^3H -thymidine incorporation in the absence of antigen.

EXAMPLE 14. INDUCTION OF SPECIFIC CTL RESPONSE IN HUMANS

[00441] A human clinical trial for an immunogenic composition comprising CTL and HTL epitopes of the invention is set up as an IND Phase I, dose escalation study and carried out as a randomized, double-blind, placebo-controlled trial. Such a trial is designed, for example, as follows:

[00442] A total of about 27 subjects are enrolled and divided into 3 groups:

Group I: 3 subjects are injected with placebo and 6 subjects are injected with 5 μg of peptide composition;

Group II: 3 subjects are injected with placebo and 6 subjects are injected with 50 μg peptide composition;

Group III: 3 subjects are injected with placebo and 6 subjects are injected with 500 μg of peptide composition.

[00443] After 4 weeks following the first injection, all subjects receive a booster inoculation at the same dosage.

[00444] The endpoints measured in this study relate to the safety and tolerability of the peptide composition as well as its immunogenicity. Cellular immune responses to the peptide composition are an index of the intrinsic activity of this the peptide composition, and can therefore be viewed as a measure of biological efficacy. The following summarize the clinical and laboratory data that relate to safety and efficacy endpoints.

[00445] Safety: The incidence of adverse events is monitored in the placebo and drug treatment group and assessed in terms of degree and reversibility.

[00446] Evaluation of Vaccine Efficacy: For evaluation of vaccine efficacy, subjects are bled before and after injection. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

[00447] The vaccine is found to be both safe and efficacious.

EXAMPLE 15. PHASE II TRIALS IN PATIENTS INFECTED WITH HIV

[00448] Phase II trials are performed to study the effect of administering the CTL-HTL peptide compositions to HIV-infected patients. The main objectives of the trials are to determine an effective dose and regimen for inducing CTLs in chronically infected HIV patients, to establish the safety of inducing a CTL and HTL response in these patients, and to see to what extent activation of CTLs improves the clinical picture of chronically infected HIV patients, as manifested by a reduction in viral load and an increase in CD4⁺ cells counts. Such a study is designed, for example, as follows:

[00449] The studies are performed in multiple centers. The trial design is an open-label, uncontrolled, dose escalation protocol wherein the peptide composition is administered as a single dose followed six weeks later by a single booster shot of the same dose. The dosages are 50, 500 and 5,000 micrograms per injection. Drug-associated adverse effects (severity and reversibility) are recorded.

[00450] There are three patient groupings. The first group is injected with 50 micrograms of the peptide composition and the second and third groups with 500 and 5,000 micrograms of peptide composition, respectively. The patients within each group range in age from 21-65, include both males and females, and represent diverse ethnic

backgrounds. All of them are infected with HIV for over five years and are HCV, HBV and delta hepatitis virus (HDV) negative, but have positive levels of HIV antigen.

[00451] The viral load and CD4⁺ levels are monitored to assess the effects of administering the peptide compositions. The vaccine composition is found to be both safe and efficacious in the treatment of HIV infection.

EXAMPLE 16. INDUCTION OF CTL RESPONSES USING A PRIME BOOST PROTOCOL

[00452] A prime boost protocol can also be used for the administration of the vaccine to humans. Such a vaccine regimen can include an initial administration of, for example, naked DNA followed by a boost using recombinant virus encoding the vaccine, or recombinant protein/polypeptide or a peptide mixture administered in an adjuvant.

[00453] For example, the initial immunization is performed using an expression vector, such as that constructed above, in the form of naked nucleic acid administered IM (or SC or ID) in the amounts of 0.5-5 mg at multiple sites. The nucleic acid (0.1 to 1000 µg) can also be administered using a gene gun. Following an incubation period of 3-4 weeks, a booster dose is then administered. The booster is, for example, recombinant fowlpox virus administered at a dose of $5 \cdot 10^7$ to $5 \cdot 10^9$ pfu. An alternative recombinant virus, such as an MVA, canarypox, adenovirus, or adeno-associated virus, can also be used for the booster, or the polyepitopic protein or a mixture of the peptides can be administered. For evaluation of vaccine efficacy, patient blood samples are obtained before immunization as well as at intervals following administration of the initial vaccine and booster doses of the vaccine. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

[00454] Analysis of the results indicates that a magnitude of sufficient response to achieve protective immunity against HIV is generated.

EXAMPLE 17. ADMINISTRATION OF VACCINE COMPOSITIONS USING DENDRITIC CELLS

- [00455] Vaccines comprising peptide epitopes of the invention can be administered using APCs, or “professional” APCs such as DC. In this example, the peptide-pulsed DC are administered to a patient to stimulate a CTL response *in vivo*. In this method, dendritic cells are isolated, expanded, and pulsed with a vaccine comprising peptide CTL and HTL epitopes of the invention. The dendritic cells are infused back into the patient to elicit CTL and HTL responses *in vivo*. The induced CTL and HTL then destroy or facilitate destruction of the specific target cells that bear the proteins from which the epitopes in the vaccine are derived.
- [00456] For example, a cocktail of epitope-bearing peptides is administered *ex vivo* to PBMC, or isolated DC therefrom. A pharmaceutical to facilitate harvesting of DC can be used, such as Progenipoietin™ (Monsanto, St. Louis, MO) or GM-CSF/IL-4. After pulsing the DC with peptides and prior to reinfusion into patients, the DC are washed to remove unbound peptides.
- [00457] As appreciated clinically, and readily determined by one of skill based on clinical outcomes, the number of DC reinfused into the patient can vary (*see, e.g., Nature Med.* 4:328, 1998; *Nature Med.* 2:52, 1996 and *Prostate* 32:272, 1997). Although $2\text{--}50 \times 10^6$ DC per patient are typically administered, larger number of DC, such as 10^7 or 10^8 can also be provided. Such cell populations typically contain between 50-90% DC.
- [00458] In some embodiments, peptide-loaded PBMC are injected into patients without purification of the DC. For example, PBMC containing DC generated after treatment with an agent such as Progenipoietin™ are injected into patients without purification of the DC. The total number of PBMC that are administered often ranges from 10^8 to 10^{10} . Generally, the cell doses injected into patients is based on the percentage of DC in the blood of each patient, as determined, for example, by immunofluorescence analysis with specific anti-DC antibodies. Thus, for example, if Progenipoietin™ mobilizes 2% DC in the peripheral blood of a given patient, and that patient is to receive 5×10^6 DC, then the patient will be injected with a total of 2.5×10^8 peptide-loaded PBMC. The percent DC mobilized by an agent such as Progenipoietin™ is typically estimated to be between 2-10%, but can vary as appreciated by one of skill in the art.

Ex vivo activation of CTL/HTL responses

[00459] Alternatively, *ex vivo* CTL or HTL responses to HIV antigens can be induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of APC, such as DC, and the appropriate immunogenic peptides. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy or facilitate destruction of their specific target cells.

[00460] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, patent applications and sequence listings cited herein are hereby incorporated by reference in their entirety for all purposes.

TABLE 1

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	T, I, L, V, M, S		F, W, Y
A2	L, I, V, M, A, T, Q		I, V, M, A, T, L
A3	V, S, M, A, T, L, I		R, K
A24	Y, F, W, I, V, L, M, T		F, I, Y, W, L, M
B7	P		V, I, L, F, M, W, Y, A
B27	R, H, K		F, Y, L, W, M, I, V, A
B44	E, D		F, W, L, I, M, V, A
B58	A, T, S		F, W, Y, L, I, V, M, A
B62	Q, L, I, V, M, P		F, W, Y, M, I, V, L, A
MOTIFS			
A1	T, S, M		Y
A1		D, E, A, S	Y
A2.1	L, M, V, Q, I, A, T		V, L, I, M, A, T
A3	L, M, V, I, S, A, T, F, C, G, D		K, Y, R, H, F, A
A11	V, T, M, L, I, S, A, G, N, C, D, F		K, R, Y, H
A24	Y, F, W, M		F, L, I, W
A*3101	M, V, T, A, L, I, S		R, K
A*3301	M, V, A, L, F, I, S, T		R, K
A*6801	A, V, T, M, S, L, I		R, K
B*0702	P		L, M, F, W, Y, A, I, V
B*3501	P		L, M, F, W, Y, I, V, A
B51	P		L, I, V, F, W, Y, A, M
B*5301	P		I, M, F, W, Y, A, L, V
B*5401	P		A, T, I, V, L, M, F, W, Y

Bolded residues are preferred, italicized residues are tolerated: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE 2

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	T, I, L, V, M, S		F, W, Y
A2	V, Q, A, T		I, V, L, M, A, T
A3	V, S, M, A, T, L, I		R, K
A24	Y, F, W, I, V, L, M, T		F, I, Y, W, L, M
B7	P		V, I, L, F, M, W, Y, A
B27	R, H, K		F, Y, L, W, M, I, V, A
B58	A, T, S		F, W, Y, L, I, V, M, A
B62	Q, L, I, V, M, P		F, W, Y, M, I, V, L, A
MOTIFS			
A1	T, S, M		Y
A1		D, E, A, S	Y
A2.1	V, Q, A, T*		V, L, I, M, A, T
A3.2	L, M, V, I, S, A, T, F, C, G, D		K, Y, R, H, F, A
A11	V, T, M, L, I, S, A, G, N, C, D, F		K, R, H, Y
A24	Y, F, W		F, L, I, W

*If 2 is V, or Q, the C-term is not L

Bolded residues are preferred, italicized residues are tolerated: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE 3

		POSITION								C-terminus
		1	2	3	4	5	6	7	8	
<u>SUPERMOTIFS</u>										
A1			<u>1° Anchor</u> T,I,L,V,M,S							<u>1° Anchor</u> F,W,Y
A2			<u>1° Anchor</u> L,I,V,M,A, T,Q							<u>1° Anchor</u> L,I,V,M,A,T
A3	preferred		<u>1° Anchor</u> V,S,M,A,T, L,I	Y,F,W, (4/5)		Y,F,W, (3/5)	Y,F,W, (4/5)	P, (4/5)		<u>1° Anchor</u> R,K
	deleterious	D,E (3/5); P, (5/5)		D,E, (4/5)						
A24			<u>1° Anchor</u> Y,F,W,I,V, L,M,T							<u>1° Anchor</u> F,I,Y,W,L,M
B7	preferred	F,W,Y (5/5) L,I,V,M, (3/5)	<u>1° Anchor</u> P	F,W,Y (4/5)			F,W,Y, (3/5)			<u>1° Anchor</u> V,I,L,F,M,W,Y,A
	deleterious	D,E (3/5); P(5/5); G(4/5); A(3/5); Q,N, (3/5)					D,E, (3/5)	G, (4/5)	Q,N, (4/5)	D,E, (4/5)
B27			<u>1° Anchor</u> R,H,K							<u>1° Anchor</u> F,Y,L,W,M,V,A
B44			<u>1° Anchor</u> E,D							<u>1° Anchor</u> F,W,Y,L,I,M,V,A
			<u>1° Anchor</u>							<u>1° Anchor</u>

TABLE 3 (Continued)

		POSITION									
		1	2	3	4	5	6	7	8	9	C-terminus or C-terminus us
B58			A,T,S								F,W,Y,L,I,V,M,A
B62			<u>1° Anchor</u> Q,L,I,V,M, P							<u>1° Anchor</u> F,W,Y,M,I,V,L,A	
<u>MOTIFS</u>											
A1	preferred	G,F,Y,W,	<u>1°Anchor</u> S,T,M,	D,E,A,	Y,F,W,		P,	D,E,Q,N,	Y,F,W,	<u>1°Anchor</u> Y	
	deleterious	D,E,		R,H,K,L,I,V M,P,	A,	G,	A,				
A1	preferred	G,R,H,K	A,S,T,C,L,I V,M,	<u>1°Anchor</u> D,E,A,S	G,S,T,C,		A,S,T,C,	L,I,V,M,	D,E,	<u>1°Anchor</u> Y	
	deleterious	A	R,H,K,D,E, P,Y,F,W,		D,E,	P,Q,N,	R,H,K,	P,G,	G,P,		
A1	peferred	Y,F,W,	<u>1°Anchor</u> S,T,M	D,E,A,Q,N,	A,	Y,F,W,Q,N,		P,A,S,T,C,	G,D,E,	P,	<u>1°Anchor</u> Y
	deleterious	G,P,		R,H,K,G,L,I V,M,	D,E,	R,H,K,	Q,N,A	R,H,K,Y,F, W,	R,H,K,	A	

TABLE 3 (Continued)

		POSITION									C-terminus	C-terminus
		1	2	3	4	5	6	7	8	9	or	us
A1 10-mer	preferred	Y,F,W,	S,T,C,L,I,V M,	<u>1°Anchor</u> D,E,A,S	A,	Y,F,W,		P,G,	G,	Y,F,W,		<u>1°Anchor</u> Y
	deleterious	R,H,K,	R,H,K,D,E, P,Y,F,W,			P,	G,		P,R,H,K,	Q,N,		
A2.1 9-mer	preferred	Y,F,W,	<u>1°Anchor</u> L,M,I,V,Q, A,T	Y,F,W,	S,T,C,	Y,F,W,		A,	P	<u>1°Anchor</u> V,L,I,M,A,T		
	deleterious	D,E,P,		D,E,R,K,H			R,K,H		D,E,R,K,H			
A2.1 10-mer	preferred	A,Y,F,W,	<u>1°Anchor</u> L,M,I,V,Q, A,T	L,V,I,M,	G,		G,		F,Y,W,L, V,I,M,			<u>1°Anchor</u> V,L,I,M,A,T
	deleterious	D,E,P,		D,E,	R,K,H,A,	P,		R,K,H,	D,E,R,K, H,	R,K,H,		

	POSITION	C-terminus or C-terminal us
1		
2		
3		
4		
5		
6		
7		
8		
9		

TABLE 4 (Continued)

A3	preferred	R,H,K,	<u>1°Anchor</u> L,M,V,I,S, A,T,F,C,G	Y,F,W,	P,R,H,K,Y, A, F,W,	Y,F,W,	P,	<u>1°Anchor</u> K,Y,R,H,F,A
			D					

deleterious D,E,P, D,E

A11	preferred	A,	$\frac{1^{\circ}\text{Anchor}}{V,T,L,M,I}$	$Y,F,W,$	$Y,FW,$	$A,$	$Y,F,W,$	$Y,FW,$	P,	$\frac{1^{\circ}\text{Anchor}}{K,,RY,H}$
			S,A,G,N,C,							
			D,F							

deleterious D,E,P, A G,

A24	preferred	Y,F,W,R,H,K,	$\frac{1^\circ \text{Anchor}}{Y,F,W,M}$	S,T,C	Y,F,W,	Y,F,W,	$\frac{1^\circ \text{Anchor}}{F,L,I,W}$
9-mer							
	deleterious	D,E,G,	D,E,	G,	Q,N,P,	D,E,R,H,K,	A,Q,N,

A24	preferred	1°Anchor Y,F,W,M	P,	Y,F,W,P,	P,	1°Anchor F,L,I,W
10-mer						
	deleterious	G,D,E	Q,N	R,H,K	A	Q,N, D,E,A,

A3101	preferred	R,H,K,	$\frac{1^{\circ}\text{Anchor}}{M,V,T,A,L,}$ I,S	$Y,F,W,$	P,	$Y,F,W,$	$Y,F,W,$	$A,P,$	$\frac{1^{\circ}\text{Anchor}}{R,K}$
-------	-----------	--------	--	----------	----	----------	----------	--------	--------------------------------------

TABLE 3 (Continued)

		POSITION								C-terminus or C-terminus us
		1	2	3	4	5	6	7	8	
deleterious	D,E,P,			D,E,		A,D,E,	D,E,	D,E,		
preferred			<u>1°Anchor</u> M,V,A,L,F, I,S,T	Y,F,W				A,Y,F,W		<u>1°Anchor</u> R,K
deleterious	G,P			D,E						
A6801	preferred	Y,F,W,S,T,C,	<u>1°Anchor</u> A,V,T,M,S, L,I		Y,F,W,L,I, V,M			Y,F,W,	P,	<u>1°Anchor</u> R,K
deleterious	G,P,			D,E,G,		R,H,K,			A,	
B0702	preferred	R,H,K,F,W,Y,	<u>1°Anchor</u> P	R,H,K,	R,H,K,	R,H,K,		R,H,K,	P,A,	<u>1°Anchor</u> L,M,F,W,Y,A, I,V
deleterious	D,E,Q,N,P,			D,E,P,	D,E,	D,E,	G,D,E,	Q,N,	D,E,	
B3501	preferred	F,W,Y,L,I,V,M,	<u>1°Anchor</u> P	F,W,Y,				F,W,Y,		<u>1°Anchor</u> L,M,F,W,Y,I, V,A
deleterious	A,G,P,					G,	G,			

TABLE 3 (Continued)

		POSITION									C-terminus or C-terminus	us
		1	2	3	4	5	6	7	8	9		
B51	preferred	L,I,V,M,F,W,Y,	<u>1°Anchor</u> P	F,W,Y,	S,T,C,	F,W,Y,		G,	F,W,Y,	<u>1°Anchor</u> L,I,V,F,W, Y,A,M		
	deleterious	A,G,P,D,E,R,H,K, S,T,C,				D,E,	G,	D,E,Q,N,	G,D,E,			
B5301	preferred	L,I,V,M,F,W,Y,	<u>1°Anchor</u> P	F,W,Y,	S,T,C,	F,W,Y,		L,I,V,M,F, W,Y,	F,W,Y,	<u>1°Anchor</u> L,M,F,W,Y, A,L,V		
	deleterious	A,G,P,Q,N,					G,	R,H,K,Q,N,	D,E,			
B5401	preferred	F,W,Y,	<u>1°Anchor</u> P	F,W,Y,L,I,V M,		L,I,V,M,		A,L,I,V,M,	F,W,Y,A,P,	<u>1°Anchor</u> A,T,I,V,L, M,F,W,Y		
	deleterious	G,P,Q,N,D,E,		G,D,E,S,T,C,		R,H,K,D,E,	D,E,	Q,N,D,G,E,	D,E,			

Italicized residues indicate "tolerated" residues.

The information in Table II is specific for 9-mers unless otherwise specified.

Secondary anchor specificities are designated for each position independently.

Table 4

HLA-supertype	Allele-specific HLA-supertype members	
	Verified ^a	Predicted ^b
A1	A*0101, A*2501, A*2601, A*2602, A*3201	A*0102, A*2604, A*3601, A*4301, A*8001
A2	A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, A*6901	A*0208, A*0210, A*0211, A*0212, A*0213
A3	A*0301, A*1101, A*3101, A*3301, A*6801	A*0302, A*1102, A*2603, A*3302, A*3303, A*3401, A*3402, A*6601, A*6602, A*7401
A24	A*2301, A*2402, A*3001	A*2403, A*2404, A*3002, A*3003
B7	B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, B*7801	B*1511, B*4201, B*5901
B27	B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, B*7301	B*2701, B*2707, B*2708, B*3802, B*3903, B*3904, B*3905, B*4801, B*4802, B*1510, B*1518, B*1503
B44	B*1801, B*1802, B*3701, B*4402, B*4403, B*4404, B*4001, B*4002, B*4006	B*4101, B*4501, B*4701, B*4901, B*5001
B58	B*5701, B*5702, B*5801, B*5802, B*1516, B*1517	
B62	B*1501, B*1502, B*1513, B*5201	B*1301, B*1302, B*1504, B*1505, B*1506, B*1507, B*1515, B*1520, B*1521, B*1512, B*1514, B*1510

- a. Verified alleles include alleles whose specificity has been determined by pool sequencing analysis, peptide binding assays, or by analysis of the sequences of CTL epitopes.
- b. Predicted alleles are alleles whose specificity is predicted on the basis of B and F pocket structure to overlap with the supertype specificity.

Table 5. Compiled rankings and similarity assignments.

A	C	D	E	F	G	H	I	K	L
A 1.0	C 1.0	D 1.0	E 1.0	F 1.0	G 1.0	H 1.0	I 1.0	K 1.0	L 1.0
S 4.5	V 5.5	N 3.5	Q 3.3	L 3.8	S 2.8	Q 2.0	L 2.5	R 2.7	I 3.3
T 4.8	T 6.5	E 4.0	N 4.2	Y 4.3	A 4.2	E 5.8	M 4.5	Q 6.0	M 4.0
P 5.3	A 6.7	Q 6.2	D 4.7	I 4.8	T 4.7	N 6.2	F 5.2	H 6.8	F 4.5
G 5.7	S 7.3	T 7.2	H 5.3	M 6.2	D 6.3	R 8.2	V 5.2	N 7.2	V 5.5
C 9.3	P 8.0	S 7.7	T 8.3	V 6.7	P 7.0	K 8.7	Y 8.8	E 7.5	Y 8.2
V 9.5	I 9.0	H 7.8	K 8.7	W 7.2	N 7.3	P 9.0	T 10.3	D 8.5	H 9.7
D 10.0	Y 10.2	G 8.8	P 9.0	H 10.2	E 8.3	D 9.2	H 11.3	M 9.2	Q 10.5
M 10.0	N 10.7	P 9.2	R 10.5	C 11.2	Q 9.3	S 10.5	A 11.5	T 10.5	W 11.0
N 10.2	F 11.0	A 9.3	S 10.5	T 11.5	H 9.8	T 10.5	C 11.7	S 11.3	T 11.3
E 11.3	G 11.0	K 9.5	V 11.0	R 11.8	C 11.5	Y 11.0	K 12.0	P 11.5	A 11.8
Q 11.3	M 11.0	R 11.5	G 12.0	A 13.0	V 12.0	M 11.2	P 12.0	I 12.2	K 12.2
H 12.3	H 11.3	C 12.0	A 12.2	K 13.3	M 13.5	V 12.2	Q 12.2	L 12.7	P 12.2
Y 13.0	D 12.0	V 13.2	M 12.3	P 13.3	K 13.8	L 13.3	R 13.0	Y 12.8	F 13.2
I 13.7	L 13.2	M 14.3	I 15.0	S 13.8	Y 14.2	I 13.7	S 13.2	V 13.7	C 13.7
L 14.3	E 13.7	Y 15.8	C 15.2	Q 14.0	R 15.5	A 13.8	E 14.0	A 14.3	E 14.2
W 14.3	Q 13.7	I 16.0	L 15.3	N 14.2	I 16.3	G 13.8	N 14.3	G 14.3	N 14.5
F 15.8	W 15.0	L 16.7	Y 15.7	E 15.3	L 16.8	F 15.3	W 14.7	F 14.8	S 15.3
K 15.8	R 16.3	W 18.0	F 17.8	G 16.3	W 17.3	C 17.2	D 15.7	W 15.5	G 16.8
R 17.7	K 17.0	F 18.3	W 18.0	D 18.0	F 18.2	W 17.5	G 17.0	C 17.5	D 17.2

Conserved (1-7)

Semi-conserved (7.1-13)

Non-conserved (13.1-20)

Semi-conserved (7.1-13)

Conserved (1-7)

Table 6. Recognition of variant peptides by CTL generated after one and five stimulations with the parent peptide.

A. Gag 386 (VLAEAMSQV)

Peptide Sequence	Binding IC50 (nM)	1 Stimulation (SU)	5 Stimulations (SU)
VLAEAMSQV	49.9	31.6	222.0
VLAEAMSQA	23.8	17.0	133.5
VLAEAMSQI	70.9	21.2	246.1
VLAEAMSKV	230.5	10.8	130.9
VLAEAMSKA	69.4	NT	36.6
ILAEAMSQA	29.3	4.0	49.7
ILAEAMSKA	72.4	--	--
VLAEAMAAA	17	16.3	90.3

B. Env 134 (KLTPLCVTL)

KLTPLCVTL	77.0	278.4	683.6
KITPLCVTL	461	231.8	700.8
QLTPLCVTL	63.6	166.2	361.5
QITPLCVTL	975	105.0	166.9
ELTPLCVTL	7190	91.7	100.0
KLTPFCVTL	87.3	36.1	75.4
KLTPLCVIL	356	77.2	29.1
KLTPLCVPL	14.6	9.6	14.8

Table 7. Conservation of EP HIV-1090 epitopes across clades, calculated as identity or immunological conservation

Protein	Sequence	<u>Total</u>		<u>Clade B</u>		<u>Clade C</u>	
		Identity	Imm. Cons.	Identity	Imm. Cons.	Identity	Imm. Cons.
Pol 498	ILKEPVHGV	62%	87%	77%	86%	74%	95%
Gag 386	VLAEAMSQV	32%	93%	68%	91%	5%	94%
Pol 448	KLVGKLNWA	95%	96%	95%	95%	95%	98%
Env 134	KLTPLCVTL	80%	93%	90%	95%	89%	98%
Vpr 62	RILQQLLFI	51%	93%	68%	91%	61%	95%
Nef 221	LTFGWCFKL	49%	74%	77%	91%	47%	81%
Gag 271	MTNNPPIPV	20%	25%	91%	95%	8%	19%
Env 47	VTVYYGVPVWK	59%	87%	95%	100%	61%	92%
Pol 929	QMAVFIHNFK	84%	98%	100%	100%	94%	97%
Pol 98	VTIKIGGQLK	11%	71%	59%	91%	2%	89%
Pol 971	KIQNFRVYYR	80%	86%	91%	95%	79%	89%
Pol 347	AIFQSSMTK	53%	75%	77%	82%	44%	79%
Pol 722	KVYLAWVPAHK	14%	97%	82%	95%	3%	97%
Env 61	TTLFCASDAK	72%	89%	90%	100%	69%	92%
Nef 94	FPVRPQVPL	81%	93%	77%	95%	82%	94%
Gag 545	YPLASLRSLF	7%	29%	45%	95%	0%	0%
Rev 75	VPLQLPPL	44%	78%	68%	77%	27%	79%
Env 259	IPIHYCAPA	74%	95%	45%	95%	79%	97%
Gag 237	HPVHAGPIA	27%	54%	68%	95%	44%	94%
Pol 893	IPYNPQSQGVV	92%	96%	82%	95%	240%	97%
Env 250	CPKVSFEPI	45%	91%	77%	100%	45%	97%
Mean		54%	81%	77%	93%	59%	84%
n=		167		22		62	

Table 8. HIV Peptides predicted to be immunologically conserved

Protein	Sequence	Conserved Epitopes*	Sequence Distribution					
			All	Subtype				
				A	B	C	D	G
Pol 498	ILKEPVHGV	ILKEPVHGV	104	1	17	46	2	2
		ILREPVHGV	12			5		1
		ILKEPVHGA	10			2	1	
		ILKDPVHGV	8	5				
		KLKEPVHGV	3					
		ILKDPVHGA	2	2				
		ILKNPVHGV	2					
Gag 386	VLAEAMSQV	VLAEAMSQA	67	2	1	36	3	3
		VLAEAMSQV	54	7	15	3	1	
		VLAEAMSQT	11			9		
		VLAEAMSHA	6			4		
		ILAEAMSQV	5		3			
		ILAEAMSQA	3			2		
		VLAEAMSHV	2					
Pol 448	KLVGKLNWA	KLVGKLNWA	158	9	21	59	3	3
		KLIGKLNWA	1					
Env 134	KLTPLCVTL	KLTPLCVTL	134	8	19	55		
		QLTPLCVTL	5	2	1			
		KLTPLCVAL	3					
		RLTPLCVTL	3			3		
		KITPLCVTL	2					
Vpr 62	RILQQLLEFI	RILQQLLEFI	86	1	15	28	4	3
		RILQQLLEFV	21	2		2		
		RTLQQLLEFI	10		2	4		
		RTLQQLLEFV	10			1		
		RILQQLLEFT	6			2		
		RMLQQLLEFI	4		1	3		
		RVLQQLLEFI	3			3		
Nef 221	LTFGWCFKL	LTFGWCFKL	82	8	17	29		3
		LTFGWCYKL	31	1	2	17		
		LTLGWCFKL	4			1		
Gag 271	MTNNPPIPV	MTSNPPIPV	60	3		24	4	1

		MTNNPPIPV	33		20	5		
		MTSNPPVPV	26	1		15		1
		MTGNPPIPV	15	5		1		
		MTGNPPVPV	9			5		
		MTNNPPVPV	6			6		
		MTANPPVPV	3			2		
Env 47	VTVYYGVPVWK	VTVYYGVPVWK	99	6	21	30	3	
		VTVYYGVPVWR	40	1		18		
		VTIYYGVPVWK	2					
Pol 929	QMAVFIHNFK	QMAVFIHNFK	153	10	22	58	4	3
		QMAV FVHNFK	3			1		
		QMAV FVHNYK	2					
Pol 98	VTIKIGGQLK	VSIKVG GQIK	30			30		
		VTIKIGGQLK	18		13	1		
		VTVKIGGQLK	11	1	1		1	
		VTVRIGGQLK	6	3				
		VSIKVG GQIR	6			6		
		VSIRVG GQIK	4			4		
		VTIRIGGQLK	3		2			
		VTVKIGGQLR	3	1				
		VTVKVG GQLK	3					
Pol 971	KIQNFRVYYR	KIQNFRVYYR	133	6	20	49	4	3
Pol 347	AIFQSSMTK	AIFQSSMTK	88	5	17	27	3	2
		AIFQCSMTK	19		2	5		
		AIFQSSMTR	13		1	11		1
		AIFQASMTK	9	1		1		
		SIFQSSMTK	9	3		6		
		AIFQYSMTK	4					
		AIFQSTMTK	2			1		
Pol 722	KVYLAWVPAHK	KVYLSWVPAHK	56	8		12	1	3
		RVYLSWVPAHK	55			41		
		KVYLAWVPAHK	23	1	18		3	
		KVYLTWVPAHK	5		2			
		KIYLSWVPAHK	5			3		
		RIYLSWVPAHK	5			4		
		KIYLAWVPAHK	2		1			
		QVYLTWVPAHK	2					
Env 61	TTLFCASDAK	TTLFCASDAK	121	9	19	41	4	1
		ATLFCASDAK	7			7		

		TILFCASDAK	6					
		PTLFCASDAK	2			1		
		TTLFCASDAR	2		2			
		TTLFCASEAK	2		1			
		ATLFCASDAR	2			2		
Nef 94	FPVRPQVPL	FPVRPQVPL	135	8	17	51	4	3
		FPVKPQVPL	9	1	3	2		
		FPVRPQVPV	4			2		
Gag 545	YPLASLRS LF	EPLTSLKSLF*	22			21		
		YPLASLKSLF*	13		5		2	

*These two would not be predicted to XR. Would choose both to get maximal population coverage.

		YPLASLRS LF	11		10			
		YPLTSLKSLF	10		1		2	1
		YPPLASLKSL	10					
		YPLTSLRS LF	6		4			
		YPPLTSLKSL	6					
Rev 75	VPLQLPPL	VPLQLPPL	64	5	15	7	4	2
		VPLQLPPI	34	2	1	19		
		VPFQLPPI	26			23		
		VPFQLPPL	3		1			1
Env 259	IPIHYCAPA	IPIHYCAPA	124	8	10	49	3	2
		IPIHYCTPA	25	1	8	8		
		IPIHFCAPA	3		1	1	1	
Gag 237	HPVHAGPIA	HPVHAGPIA	39		15	21	1	
		HPVHAGPVA	34	1	3	27	2	
		HPVQAGPVA	12		3	6	1	
		HPIHAGPIA	2			2		
Pol 893	IPYNPQS QGVV	IPYNPQS QGVV	153	9	18	60	4	3
		IPYNPQS QGVI	5		3			
		IPYNPQS QGAV	2			1		
Env 250	CPKVSFEPI	CPKVSFEPI	50	5	17	3	2	0
		CPKVSFDPI	42			33		
		CPKVTFDPI	16			13		1
		CPKVTFEPI	13	3		1		1
		CPKISFDPI	9			5		
		CPKISFEPI	7		4		1	
		CPKVSWDPI	6					
		CPKVSFQPI	4		1			

* The preferred epitopes are shown in bold

Table 9. Predicted immunological conservation for a panel of HLA-A2 rest: peptides.

Source	Parent Sequence	HPV Strain	Variant Sequences	SEQ ID NO	Predicted Immunogenicity	Measured Immunogenicity (SU)*
HPV16.E7.86	TLGIVCPI	16	TLGIVCPI		+	103.7
		18	TLSFVCPW		-	
		31	SFGIVCPN		-	
		33	TVNIVCPT		-	
		45	TLSFVCPW		-	
		52	TLQVVCPG		-	
		56	ALTVTCPL		-	
		58	TCTIVCPS		-	
HPV31.E6.11	KLHELSSAL	16	KLPQLCTEL		-	
		18	KLPDLCTEL		-	
		31	KLHELSSAL		+	26.3
		33	TLHDLQAL		-	
		45	KLPDLCTEL		-	
		52	TLHELCEVL		-	
		56	SLHHLSEVL		-	
		58	TLHDLQAL		-	
HPV18/45.E6.13	KLPDLCTEL	16	KLPQLCTEL		+	15.7
		18	KLPDLCTEL		+	212.7
		31	KLHELSSAL		-	
		33	TLHDLQAL		-	
		45	KLPDLCTEL		+	205.1
		52	TLHELCEVL		-	
		56	SLHHLSEVL		-	
		58	TLHDLQAL		-	
HPV52.E6.18	VLEESVHEI	16	ELQTTIHDI		-	
		18	ELNTSLQDI		-	
		31	ALEIPYDEL		-	
		33	ALETTIHNI		-	
		45	ELNTSLQDV		-	
		52	VLEESVHEI		+	64.1
		56	VLEIPLIDL		-	
		58	ALETSVHEI		-	
HPV18.E6.47	FAFKDLFVV	16	FAFRDLCIV		-	
		18	FAFKDLFVV		+	350.6
		31	FAFTDLTIV		-	
		33	FAFADLTVV		-	31.4
		45	FAFKDLCIV		-	176.9
		52	FLFTDLRIV		-	
		56	FACTELKLV		-	
		58	FVFADLRIV		-	7.7
HPV31.E6.45	FAFTDLTIV	16	FAFRDLCIV		-	

		18	FAFKDLFVV		-	
		31	FAFTDLTIV		+	20.7
		33	FAFADLTVV		+	11.6
		45	FAFKDLCIV		-	
		52	FLFTDLRIV		-	
		56	FACTELKLV		-	
		58	FVFADLRIV		-	
HPV52.E6.45	FLFTDLRIV	16	FAFRDLCIV		-	
		18	FAFKDLFVV		-	
		31	FAFTDLTIV		-	
		33	FAFADLTVV		-	
		45	FAFKDLCIV		-	
		52	FLFTDLRIV		+	421.4
		56	FACTELKLV		-	57.5
		58	FVFADLRIV		+	94.1
HPV58.E6.45	FVFADLRIV	16	FAFRDLCIV		-	
		18	FAFKDLFVV		-	
		31	FAFTDLTIV		-	
		33	FAFADLTVV		-	
		45	FAFKDLCIV		-	
		52	FLFTDLRIV		+	13.3
		56	FACTELKLV		-	21.0
		58	FVFADLRIV		+	62.8
HPV18.E7.7	TLQDIVLHL	16	TLHEYMLDL		-	
		18	TLQDIVLHL		+	99.0
		31	TLQDYVLDL		-	
		33	TLKEYVLDL		-	
		45	TLQEIVLHL		+	
		52	TIKDYILDL		-	
		56	TLQDVVLEL		+	38.0
		58	TLREYILDL		-	
HPV16.E7.82	LLMGTLGIV	16	LLMGTLGIV		+	518.5
		18	LFLNTLSFV		-	
		31	LLMGSGFIV		+	90.1
		33	LLMGTVNIV		-	
		45	LFLSTLSFV		+	
		52	MLLGTQVV		-	
		56	LLMGALTVT		+	
		58	LLMGCTIV		-	
HPV33.E7.81	LLMGTVNIV	16	LLMGTLGIV		-	
		18	LFLNTLSFV		-	
		31	LLMGSGFIV		-	
		33	LLMGTVNIV		+	179.4
		45	LFLSTLSFV		-	
		52	MLLGTQVV		+	
		56	LLMGALTVT		-	20.8
		58	LLMGCTIV		-	
HPV52.E7.84	MLLGTQVV	16	LLMGTLGIV		-	

		18	LFLNTLSFV		-	
		31	LLMGSEFGIV		-	
		33	LLMGTVNIV		+	
		45	LFLSTLSFV		-	
		52	MLLGTLQVV		+	99.8
		56	LLMGALTVT		-	
		58	LLMGCTIV		-	
HPV56.E7.89	LLMGALTVT	16	LLMGTLGIV		-	
		18	LFLNTLSFV		-	
		31	LLMGSEFGIV		-	
		33	LLMGTVNIV		+	
		45	LFLSTLSFV		-	
		52	MLLGTLQVV		-	
		56	LLMGALTVT		+	263.5
		58	LLMGCTIV		-	43.6

* Immunogenicity was measured for all variants. Only the positive responses are shown in the table. All other responses were negative.

--	--	--	--	--	--	--

Table 10. 167 HIV-1 Variants

SEQ ID NO	Sequence Designation	Name	Accession Number	SubType	Country
	A.UG.92UG037_U51190	92UG037	U51190	A	UG
	A.BY.97BL006_AF1932	97BL006	AF193275	A	BY
	A.KE.Q23_AF004885	Q23	AF004885	A	KE
	A.SE.SE6594_AF06967	SE6594	AF069672	A	SE
	A.SE.SE7253_AF06967	SE7253	AF069670	A	SE
	A.SE.SE7535_AF06967	SE7535	AF069671	A	SE
	A.SE.SE8538_AF06966	SE8538	AF069669	A	SE
	A.SE.SE8891_AF06967	SE8891	AF069673	A	SE
	A.UG.U455_M62320	U455	M62320	A	UG
	A.SE.UGSE8131_AF107	UGSE8131	AF107771	A	SE
	A2.CY.94CY017.41_AF	94CY017.41	AF286237	A2	CY
	A2.CD.97CDKTB48_AF2	97CDKTB48	AF286238	A2	CD
	A2D.-.97KR004_AF286	97KR004	AF286239	A2D	KR
	A2G.CD.97CDKP58_AF3	97CDKP58	AF316544	A2G	CD
	AC.IN.21301_AF06715	21301	AF067156	AC	IN
	AC.RW.92RW009_U8882	92RW009	U88823	AC	RW
	AC.SE.SE9488_AF0714	SE9488	AF071474	AC	SE
	ACD.SE.SE8603_AF075	SE8603	AF075702	ACD	SE
	ACG.BE.VI1035_AJ276	VI1035	AJ276595	ACG	BE
	AD.SE.SE6954_AF0757	SE6954	AF075701	AD	SE
	AD.SE.SE7108_AF0714	SE7108	AF071473	AD	SE
	ADHK.NO.97NOGIL3_AJ	97NOGIL3	AJ237565	ADHK	NO
	ADK.CD.MAL_X04415	MAL	X04415	ADK	CD
	AG.NG.92NG003_U8882	92NG003	U88825	AG	NG
	AG.BE.VI1197_AJ2765	VI1197	AJ276596	AG	BE
	AGHU.GA.VI354_AF076	VI354	AF076474	AGHU	GA
	AGU.CD.Z321_U76035	Z321	U76035	AGU	CD
	AJ.BW.BW2117_AF1921	BW2117	AF192135	AJ	BW
	B.NL.3202A21_U34604	3202A21	U34604	B	NL
	B.US.BC_L02317	BC	L02317	B	US
	B.GB.CAM1_D10112	CAM1	D10112	B	GB
	B.DE.D31_U43096	D31	U43096	B	DE
	B.US.DH123_AF069140	DH123	AF069140	B	US
	B.GB.GB8.C1_Y13716	GB8	AJ271445	B	GB
	B.DE.HAN_U43141	HAN	U43141	B	DE
	B.FR.HXB2_K03455	HXB2	K03455	B	FR
	B.US.JRCFSF_M38429	JRCFSF	M38429	B	US
	B.GB.MANC_U23487	MANC	U23487	B	GB
	B.US.MNCG_M17449	MNCG	M17449	B	US
	B.GA.OYI_M26727	OYI	M26727	B	GA
	B.US.P896_U39362	P896	M96155	B	US
	B.US.RF_M17451	RF	M17451	B	US
	B.CN.RL42_U71182	RL42	U71182	B	CN
	B.US.SF2_K02007	SF2	K02007	B	US
	B.TW.TWCYS_AF086817	TWCYS	AF086817	B	TW
	B.AU.VH_AF146728	VH	AF146728	B	AU
	B.US.WEAU160_U21135	WEAU160	U21135	B	US
	B.KR.WK_AF224507	WK	AF224507	B	KR
	B.US.WR27_U26546	WR27	U26546	B	US

B.US.YU2_M93258	YU2	M93258	B	US
BF1.BR.93BR029.4_AF	93BR029.4	AF005495	BF1	BR
C.BR.92BR025_U52953	92BR025	U52953	C	BR
C.IN.93IN101_AB0238	93IN101	AB023804	C	IN
C.IN.93IN904_AF0671	93IN904	AF067157	C	IN
C.IN.93IN999_AF0671	93IN999	AF067154	C	IN
C.IN.94IN11246_AF06	94IN11246	AF067159	C	IN
C.IN.95IN21068_AF06	95IN21068	AF067155	C	IN
C.BW.96BW0402_AF110	96BW0402	AF110962	C	BW
C.BW.96BW1210_AF110	96BW1210	AF110972	C	BW
C.BW.96BW15B03_AF11	96BW15B03	AF110973	C	BW
C.ET.ETH2220_U46016	ETH2220	U46016	C	ET
C.BW.96BW11B01_AF11	96BW11	AF110969	C	BW
C.BW.00BW0762.1_AF44	00BW0762.1	AF443088	C	BW
C.BW.00BW0768.20_AF44	00BW0768.20	AF443089	C	BW
C.BW.00BW0874.21_AF44	00BW0874.21	AF443090	C	BW
C.BW.00BW1471.27_AF44	00BW1471.27	AF443091	C	BW
C.BW.00BW1616.2_AF44	00BW1616.2	AF443092	C	BW
C.BW.00BW1686.8_AF44	00BW1686.8	AF443093	C	BW
C.BW.00BW1759.3_AF44	00BW1759.3	AF443094	C	BW
C.BW.00BW1773.2_AF44	00BW1773.2	AF443095	C	BW
C.BW.00BW1783.5_AF44	00BW1783.5	AF443096	C	BW
C.BW.00BW1795.6_AF44	00BW1795.6	AF443097	C	BW
C.BW.00BW1811.3_AF44	00BW1811.3	AF443098	C	BW
C.BW.00BW1859.5_AF44	00BW1859.5	AF443099	C	BW
C.BW.00BW1880.2_AF44	00BW1880.2	AF443100	C	BW
C.BW.00BW1921.13_AF44	00BW1921.13	AF443101	C	BW
C.BW.00BW2036.1_AF44	00BW2036.1	AF443102	C	BW
C.BW.00BW2063.6_AF44	00BW2063.6	AF443103	C	BW
C.BW.00BW2087.2_AF44	00BW2087.2	AF443104	C	BW
C.BW.00BW2127.214_AF44	00BW2127.214	AF443105	C	BW
C.BW.00BW2128.3_AF44	00BW2128.3	AF443106	C	BW
C.BW.00BW2276.7_AF44	00BW2276.7	AF443107	C	BW
C.BW.00BW3819.3_AF44	00BW3819.3	AF443108	C	BW
C.BW.00BW3842.8_AF44	00BW3842.8	AF443109	C	BW
C.BW.00BW3871.3_AF44	00BW3871.3	AF443110	C	BW
C.BW.00BW3876.9_AF44	00BW3876.9	AF443111	C	BW
C.BW.00BW3886.8_AF44	00BW3886.8	AF443112	C	BW
C.BW.00BW3891.6_AF44	00BW3891.6	AF443113	C	BW
C.BW.00BW3970.2_AF44	00BW3970.2	AF443114	C	BW
C.BW.00BW5031.1_AF44	00BW5031.1	AF443115	C	BW
C.BW.96BW01B21_AF11	96BW01B21	AF110960	C	BW
C.BW.96BW0407_AF11	96BW0407	AF110963	C	BW
C.BW.96BW0502_AF11	96BW0502	AF110967	C	BW
C.BW.96BW06.J4_AF29	96BW06.J4	AF290028	C	BW
C.BW.96BW11.06_AF11	96BW11.06	AF110970	C	BW
C.BW.96BW1210_AF11	96BW1210	AF110972	C	BW
C.BW.96BW15B03_AF11	96BW15B03	AF110973	C	BW
C.BW.96BW16.26_AF11	96BW16.26	AF110978	C	BW
C.BW.96BW17A09_AF11	96BW17A09	AF110979	C	BW
C.BW.96BWMO1.5_AF44	96BWMO1.5	AF443074	C	BW
C.BW.96BWMO3.2_AF44	96BWMO3.2	AF443075	C	BW
C.BW.98BWMC12.2_AF44	98BWMC12.2	AF443076	C	BW
C.BW.98BWMC13.4_AF44	98BWMC13.4	AF443077	C	BW
C.BW.98BWMC14.a3_AF44	98BWMC14.a3	AF443078	C	BW

C.BW.98BWMO14.10_AF44	98BWMO14.10	AF443079	C	BW
C.BW.98BWMO18.d5_AF44	98BWMO18.d5	AF443080	C	BW
C.BW.98BWMO36.a5_AF44	98BWMO36.a5	AF443081	C	BW
C.BW.98BWMO37.d5_AF44	98BWMO37.d5	AF443082	C	BW
C.BW.99BW3932.12_AF44	99BW3932.12	AF443083	C	BW
C.BW.99BW4642.4_AF44	99BW4642.4	AF443084	C	BW
C.BW.99BW4745.8_AF44	99BW4745.8	AF443085	C	BW
C.BW.99BW4754.7_AF44	99BW4754.7	AF443086	C	BW
C.BW.99BWMC16.8_AF44	99BWMC16.8	AF443087	C	BW
CRF01_AE.CF.90CF11697_	90CF11697	AF197340	CRF01_AE	CF
CRF01_AE.CF.90CF402_U5	90CF402	U51188	CRF01_AE	CF
CRF01_AE.CF.90CF4071_A	90CF4071	AF197341	CRF01_AE	CF
CRF01_AE.TH.93TH057_AF	93TH057	AF197338	CRF01_AE	TH
CRF01_AE.TH.93TH065_AF	93TH065	AF197339	CRF01_AE	TH
CRF01_AE.TH.93TH253_U5	93TH253	U51189	CRF01_AE	TH
CRF01_AE.TH.95TNIH047_	95TNIH047	AB032741	CRF01_AE	TH
CRF01_AE.TH.CM240_U547	CM240	U54771	CRF01_AE	TH
CRF01_AE.TH.TH022_AB03	TH022	AB032740	CRF01_AE	TH
CRF02_AG.SN.98SEMP1211	98SEMP1211	AJ251056	CRF02_AG	SN
CRF02_AG.FR.DJ263_AF06	DJ263	AF063223	CRF02_AG	FR
CRF02_AG.FR.DJ264_AF06	DJ264	AF063224	CRF02_AG	FR
CRF02_AG.GH.G829_AF184	G829	AF184155	CRF02_AG	GH
CRF02_AG.NG.IBNG_L3910	IBNG	L39106	CRF02_AG	NG
CRF02_AG.SE.SE7812_AF1	SE7812	AF107770	CRF02_AG	SE
CRF03_AB.RU.KAL153-2_A	KAL153-2	AF193276	CRF03_AB	RU
CRF03_AB.RU.RU98001_AF	RU98001	AF193277	CRF03_AB	RU
CRF04_cpx.CY.94CY032-3	94CY032-3	AF049337	CRF04_cpx	CY
CRF04_cpx.GR.97PVCH_AF	97PVCH	AF119820	CRF04_cpx	GR
CRF04_cpx.GR.97PVMY_AF	97PVMY	AF119819	CRF04_cpx	GR
CRF05_DF.BE.VI1310_AF1	VI1310	AF193253	CRF05_DF	BE
CRF05_DF.BE.VI961_AF07	VI961	AF076998	CRF05_DF	BE
CRF06_cpx.ML.95ML127_A	95ML127	AJ288982	CRF06_cpx	ML
CRF06_cpx.ML.95ML84_AJ	95ML84	AJ245481	CRF06_cpx	ML
CRF06_cpx.SN.97SE1078_	97SE1078	AJ288981	CRF06_cpx	SN
CRF06_cpx.AU.BFP90_AF0	BFP90	AF064699	CRF06_cpx	AU
CRF11_cpx.CM.97CM-MP81	97CM-MP818	AJ291718	CRF11_cpx	CM
CRF11_cpx.GR.GR17_AF17	GR17	AF179368	CRF11_cpx	GR
D.CD.84ZR085_U88822	84ZR085	U88822	D	CD
D.UG.94UG1141_U8882	94UG1141	U88824	D	UG
D.CD.ELI_K03454	ELI	K03454	D	CD
D.CD.NDK_M27323	NDK	M27323	D	CD
F1.BR.93BR020.1_AF0	93BR020.1	AF005494	F1	BR
F1.FI.FIN9363_AF075	FIN9363	AF075703	F1	FI
F1.FR.MP411_AJ24923	MP411	AJ249238	F1	FR
F1.BE.VI850_AF07733	VI850	AF077336	F1	BE
F2.CM.MP257_AJ24923	MP257	AJ249237	F2	CM
F2KU.BE.VI1126_AF07	VI1126	AF076475	F2KU	BE
G.NG.92NG083_U88826	92NG083	U88826	G	NG
G.BE.DRCBL_AF084936	DRCBL	AF084936	G	BE
G.SE.SE6165_AF06164	SE6165	AF061642	G	SE
H.CF.90CF056_AF0054	90CF056	AF005496	H	CF
H.BE.VI991_AF190127	VI991	AF190127	H	BE
H.BE.VI997_AF190128	VI997	AF190128	H	BE
J.SE.SE7022_AF08239	SE7022	AF082395	J	SE
J.SE.SE7887_AF08239	SE7887	AF082394	J	SE

WO 2005/012502

PCT/US2004/009510

K.CD.EQTB11C_AJ2492	EQTB11C	AJ249235	K	CD
K.CM.MP535_AJ249239	MP535	AJ249239	K	CM
N.CM.YBF30_AJ006022	YBF30	AJ006022	N	CM
O.SN.99SE-MP1299_ZX	SEMP1299	AJ302646	O	SN
O.SN.99SE-MP1300_ZX	SEMP1300	AJ302647	O	SN
O.CM.ANT70_L20587	ANT70	L20587	O	CM
O.CM.MVP5180_L20571	MVP5180	L20571	O	CM
U.CD.-.83CD0031	83CD0031	AF286236	U	CD

Table 11. HIV Gag Sequence Alignment

GCG Multiple Sequence File.

Written by Omiga 1.1

Name: 00BW0762_1	Len: 556	Check: 2513	Weight: 1.00
Name: 00BW0768_2	Len: 556	Check: 8965	Weight: 1.00
Name: 00BW0874_2	Len: 556	Check: 9574	Weight: 1.00
Name: 00BW1471_2	Len: 556	Check: 5395	Weight: 1.00
Name: 00BW1616_2	Len: 556	Check: 4692	Weight: 1.00
Name: 00BW1686_8	Len: 556	Check: 7822	Weight: 1.00
Name: 00BW1759_3	Len: 556	Check: 7777	Weight: 1.00
Name: 00BW1773_2	Len: 556	Check: 9727	Weight: 1.00
Name: 00BW1783_5	Len: 556	Check: 9681	Weight: 1.00
Name: 00BW1795_6	Len: 556	Check: 9667	Weight: 1.00
Name: 00BW1811_3	Len: 556	Check: 4422	Weight: 1.00
Name: 00BW1859_5	Len: 556	Check: 7320	Weight: 1.00
Name: 00BW1880_2	Len: 556	Check: 1603	Weight: 1.00
Name: 00BW1921_1	Len: 556	Check: 883	Weight: 1.00
Name: 00BW2036_1	Len: 556	Check: 2591	Weight: 1.00
Name: 00BW2063_6	Len: 556	Check: 5152	Weight: 1.00
Name: 00BW2087_2	Len: 556	Check: 5183	Weight: 1.00
Name: 00BW2127_2	Len: 556	Check: 5469	Weight: 1.00
Name: 00BW2128_3	Len: 556	Check: 9621	Weight: 1.00
Name: 00BW2276_7	Len: 556	Check: 4153	Weight: 1.00
Name: 00BW3819_3	Len: 556	Check: 4227	Weight: 1.00
Name: 00BW3842_8	Len: 556	Check: 9312	Weight: 1.00
Name: 00BW3871_3	Len: 556	Check: 501	Weight: 1.00
Name: 00BW3876_9	Len: 556	Check: 773	Weight: 1.00
Name: 00BW3886_8	Len: 556	Check: 2351	Weight: 1.00
Name: 00BW3891_6	Len: 556	Check: 129	Weight: 1.00
Name: 00BW3970_2	Len: 556	Check: 8768	Weight: 1.00
Name: 00BW5031_1	Len: 556	Check: 3966	Weight: 1.00
Name: 96BW01B21	Len: 556	Check: 602	Weight: 1.00
Name: 96BW0407	Len: 556	Check: 9836	Weight: 1.00
Name: 96BW0502	Len: 556	Check: 6402	Weight: 1.00
Name: 96BW06_J4	Len: 556	Check: 254	Weight: 1.00
Name: 96BW11_06	Len: 556	Check: 6801	Weight: 1.00
Name: 96BW1210	Len: 556	Check: 6016	Weight: 1.00
Name: 96BW15B03	Len: 556	Check: 6072	Weight: 1.00
Name: 96BW16_26	Len: 556	Check: 9409	Weight: 1.00
Name: 96BW17A09	Len: 556	Check: 2723	Weight: 1.00
Name: 96BWM01_5	Len: 556	Check: 5051	Weight: 1.00
Name: 96BWM03_2	Len: 556	Check: 496	Weight: 1.00
Name: 98BWMC12_2	Len: 556	Check: 1164	Weight: 1.00
Name: 98BWMC13_4	Len: 556	Check: 4961	Weight: 1.00
Name: 98BWMC14_a	Len: 556	Check: 7351	Weight: 1.00
Name: 98BWM014_1	Len: 556	Check: 288	Weight: 1.00
Name: 98BWM018_d	Len: 556	Check: 6836	Weight: 1.00
Name: 98BWM036_a	Len: 556	Check: 4386	Weight: 1.00
Name: 98BWM037_d	Len: 556	Check: 6900	Weight: 1.00
Name: 99BW3932_1	Len: 556	Check: 292	Weight: 1.00
Name: 99BW4642_4	Len: 556	Check: 1347	Weight: 1.00
Name: 99BW4745_8	Len: 556	Check: 7980	Weight: 1.00
Name: 99BW4754_7	Len: 556	Check: 9892	Weight: 1.00
Name: 99BWMC16_8	Len: 556	Check: 4279	Weight: 1.00
Name: A2_CD_97CD	Len: 556	Check: 9920	Weight: 1.00
Name: A2_CY_94CY	Len: 556	Check: 3070	Weight: 1.00
Name: A2D_97KR	Len: 556	Check: 6350	Weight: 1.00
Name: A2G_CD_97C	Len: 556	Check: 856	Weight: 1.00
Name: A_BY_97BL0	Len: 556	Check: 9315	Weight: 1.00
Name: A_KE_Q23_A	Len: 556	Check: 2442	Weight: 1.00

WO 2005/012502

PCT/US2004/009510

Name: A_SE_SE659	Len: 556	Check: 8612	Weight: 1.00
Name: A_SE_SE725	Len: 556	Check: 8315	Weight: 1.00
Name: A_SE_SE753	Len: 556	Check: 2915	Weight: 1.00
Name: A_SE_SE853	Len: 556	Check: 9112	Weight: 1.00
Name: A_SE_SE889	Len: 556	Check: 8732	Weight: 1.00
Name: A_SE_UGSE8	Len: 556	Check: 8696	Weight: 1.00
Name: A_UG_92UG0	Len: 556	Check: 6290	Weight: 1.00
Name: A_UG_U455_	Len: 556	Check: 164	Weight: 1.00
Name: AC_IN_2130	Len: 556	Check: 8482	Weight: 1.00
Name: AC_RW_92RW	Len: 556	Check: 977	Weight: 1.00
Name: AC_SE_SE94	Len: 556	Check: 8752	Weight: 1.00
Name: ACD_SE_SE8	Len: 556	Check: 9655	Weight: 1.00
Name: ACG_BE_VI1	Len: 556	Check: 3777	Weight: 1.00
Name: AD_SE_SE69	Len: 556	Check: 732	Weight: 1.00
Name: AD_SE_SE71	Len: 556	Check: 8506	Weight: 1.00
Name: ADHK_NO_97	Len: 556	Check: 5257	Weight: 1.00
Name: ADK_CD_MAL	Len: 556	Check: 5301	Weight: 1.00
Name: AG_BE_VI11	Len: 556	Check: 1610	Weight: 1.00
Name: AG_NG_92NG	Len: 556	Check: 4188	Weight: 1.00
Name: AGHU_GA_VI	Len: 556	Check: 8242	Weight: 1.00
Name: AGU_CD_Z32	Len: 556	Check: 2601	Weight: 1.00
Name: AJ_BW_BW21	Len: 556	Check: 8389	Weight: 1.00
Name: B_AU_VH_AF	Len: 556	Check: 9166	Weight: 1.00
Name: B_CN_RL42_	Len: 556	Check: 3865	Weight: 1.00
Name: B_DE_D31_U	Len: 556	Check: 9464	Weight: 1.00
Name: B_DE_HAN_U	Len: 556	Check: 6999	Weight: 1.00
Name: B_FR_HXB2_	Len: 556	Check: 141	Weight: 1.00
Name: B_GA_OYI_	Len: 556	Check: 7578	Weight: 1.00
Name: B_GB_CAM1_	Len: 556	Check: 8023	Weight: 1.00
Name: B_GB_GB8_A	Len: 556	Check: 2675	Weight: 1.00
Name: B_GB_MANC_	Len: 556	Check: 8961	Weight: 1.00
Name: B_KR_WK_AF	Len: 556	Check: 7372	Weight: 1.00
Name: B_NL_3202A	Len: 556	Check: 7117	Weight: 1.00
Name: B_TW_TWCYS	Len: 556	Check: 8767	Weight: 1.00
Name: B_US_BC_L0	Len: 556	Check: 7091	Weight: 1.00
Name: B_US_DH123	Len: 556	Check: 5049	Weight: 1.00
Name: B_US_JRCSE	Len: 556	Check: 975	Weight: 1.00
Name: B_US_MNCG_	Len: 556	Check: 688	Weight: 1.00
Name: B_US_P896_	Len: 556	Check: 8809	Weight: 1.00
Name: B_US_RF_M1	Len: 556	Check: 9306	Weight: 1.00
Name: B_US_SF2_K	Len: 556	Check: 9799	Weight: 1.00
Name: B_US_WEAU1	Len: 556	Check: 9636	Weight: 1.00
Name: B_US_WR27_	Len: 556	Check: 3349	Weight: 1.00
Name: B_US_YU2_M	Len: 556	Check: 8828	Weight: 1.00
Name: BF1_BR_93B	Len: 556	Check: 7935	Weight: 1.00
Name: C_BR_92BR0	Len: 556	Check: 4429	Weight: 1.00
Name: C_BW_96BW0	Len: 556	Check: 1550	Weight: 1.00
Name: C_BW_96BW1	Len: 556	Check: 7158	Weight: 1.00
Name: C_BW_96BW1	Len: 556	Check: 6016	Weight: 1.00
Name: C_BW_96BW1	Len: 556	Check: 6072	Weight: 1.00
Name: C_ET_ETH22	Len: 556	Check: 4314	Weight: 1.00
Name: C_IN_93IN1	Len: 556	Check: 6959	Weight: 1.00
Name: C_IN_93IN9	Len: 556	Check: 9362	Weight: 1.00
Name: C_IN_93IN9	Len: 556	Check: 3298	Weight: 1.00
Name: C_IN_94IN1	Len: 556	Check: 6744	Weight: 1.00
Name: C_IN_95IN2	Len: 556	Check: 8559	Weight: 1.00
Name: CRF01_AE_C	Len: 556	Check: 4763	Weight: 1.00
Name: CRF01_AE_C	Len: 556	Check: 4315	Weight: 1.00
Name: CRF01_AE_C	Len: 556	Check: 3920	Weight: 1.00
Name: CRF01_AE_T	Len: 556	Check: 5074	Weight: 1.00
Name: CRF01_AE_T	Len: 556	Check: 4026	Weight: 1.00

WO 2005/012502

PCT/US2004/009510

Name: CRF01_AE_T	Len: 556	Check: 2981	Weight: 1.00
Name: CRF01_AE_T	Len: 556	Check: 6648	Weight: 1.00
Name: CRF01_AE_T	Len: 556	Check: 6948	Weight: 1.00
Name: CRF01_AE_T	Len: 556	Check: 6129	Weight: 1.00
Name: CRF02_AG_F	Len: 556	Check: 8134	Weight: 1.00
Name: CRF02_AG_F	Len: 556	Check: 8925	Weight: 1.00
Name: CRF02_AG_G	Len: 556	Check: 9991	Weight: 1.00
Name: CRF02_AG_N	Len: 556	Check: 9388	Weight: 1.00
Name: CRF02_AG_S	Len: 556	Check: 9371	Weight: 1.00
Name: CRF02_AG_S	Len: 556	Check: 3289	Weight: 1.00
Name: CRF03_AB_R	Len: 556	Check: 5554	Weight: 1.00
Name: CRF03_AB_R	Len: 556	Check: 5307	Weight: 1.00
Name: CRF04_cpx_	Len: 556	Check: 5488	Weight: 1.00
Name: CRF04_cpx_	Len: 556	Check: 6046	Weight: 1.00
Name: CRF04_cpx_	Len: 556	Check: 3807	Weight: 1.00
Name: CRF05_DF_B	Len: 556	Check: 6702	Weight: 1.00
Name: CRF05_DF_B	Len: 556	Check: 5563	Weight: 1.00
Name: CRF06_cpx_	Len: 556	Check: 8164	Weight: 1.00
Name: CRF06_cpx_	Len: 556	Check: 1975	Weight: 1.00
Name: CRF06_cpx_	Len: 556	Check: 1486	Weight: 1.00
Name: CRF06_cpx_	Len: 556	Check: 4050	Weight: 1.00
Name: CRF11_cpx_	Len: 556	Check: 1515	Weight: 1.00
Name: CRF11_cpx_	Len: 556	Check: 9474	Weight: 1.00
Name: D_CD_84ZR0	Len: 556	Check: 4739	Weight: 1.00
Name: D_CD_ELI_K	Len: 556	Check: 8533	Weight: 1.00
Name: D_CD_NDK_M	Len: 556	Check: 6100	Weight: 1.00
Name: D_UG_94UG1	Len: 556	Check: 7681	Weight: 1.00
Name: F1_BE_VI85	Len: 556	Check: 8540	Weight: 1.00
Name: F1_BR_93BR	Len: 556	Check: 241	Weight: 1.00
Name: F1_FI_FIN9	Len: 556	Check: 7510	Weight: 1.00
Name: F1_FR_MP41	Len: 556	Check: 9846	Weight: 1.00
Name: F2_CM_MP25	Len: 556	Check: 1247	Weight: 1.00
Name: F2KU_BE_VI	Len: 556	Check: 2186	Weight: 1.00
Name: G_BE_DRCBL	Len: 556	Check: 8049	Weight: 1.00
Name: G_NG_92NG0	Len: 556	Check: 254	Weight: 1.00
Name: G_SE_SE616	Len: 556	Check: 6407	Weight: 1.00
Name: H_BE_VI991	Len: 556	Check: 808	Weight: 1.00
Name: H_BE_VI997	Len: 556	Check: 4749	Weight: 1.00
Name: H_CF_90CF0	Len: 556	Check: 4718	Weight: 1.00
Name: J_SE_SE702	Len: 556	Check: 2564	Weight: 1.00
Name: J_SE_SE788	Len: 556	Check: 2692	Weight: 1.00
Name: K_CD_EQTB1	Len: 556	Check: 6586	Weight: 1.00
Name: K_CM_MP535	Len: 556	Check: 3803	Weight: 1.00
Name: N_CM_YBF30	Len: 556	Check: 8603	Weight: 1.00
Name: O_CM_ANT70	Len: 556	Check: 365	Weight: 1.00
Name: O_CM_MVP51	Len: 556	Check: 9160	Weight: 1.00
Name: O_SN_99SE_	Len: 556	Check: 9380	Weight: 1.00
Name: O_SN_99SE_	Len: 556	Check: 8796	Weight: 1.00
Name: U_CD_83C	Len: 556	Check: 4633	Weight: 1.00

//

	1				50
00BW0762_1	MGARASILRG	.EKLDKWEKI	RLRPGGRKHY	MIKHIVWASR	ELERFALNPG
00BW0768_2	MGARASVLRG	.EKLDKWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
00BW0874_2	MGARASILRG	.GKLDTWEKI	RLRPGGKKQY	MIKHLVWASR	ELERFALNPG
00BW1471_2	MGARASILRG	.GKLDTWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
00BW1616_2	MGARASILRG	.GKLDEWEKI	RLRPGGKKRY	MMKHLVWASR	ELERFALNPG
00BW1686_8	MGARASILRG	.GKLDTWEKI	RLRPGGKKHY	MIKHLVWASR	ELERFALNPG
00BW1759_3	MGARASILRG	.GKLDKWERI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
00BW1773_2	MGASASILRG	.GKLDKWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNSG

00BW1783_5	MGARASILRG	.GKLDTWEKI	RLRPGGKKHY	MMKHLVWASR	ELERFALNPG
00BW1795_6	MGARASILRG	.GKLDKWEKI	RLRPGGKKHY	MMKHLVWASR	ELERFALNPG
00BW1811_3	MGARASILRG	.GKLDKWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
00BW1859_5	MGARASVLKG	.EKLDWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
00BW1880_2	MGARASILRG	.GKLDKWEKI	RLRPGGKKQY	MIKHLVWASR	ELERFALNPG
00BW1921_1	MGARASILRG	.GKLDTWEKI	RLRPGGKKRY	MLKHLIWTSR	ELERFALNPD
00BW2036_1	MGARASILRG	.EKLDTWEKI	KLPRPGGKKHY	MLKHLVWASR	ELERFALNPG
00BW2063_6	MGARASILRG	.GKLDKWEKI	RLRPGGKKQY	MIKHLVWASR	ELERFALNPG
00BW2087_2	MGARASILRG	.GKLDTWEKI	KLPRPGGKKSY	KLKHLVWASR	ELERFALNPS
00BW2127_2	MGARASILRG	XEKLDWEKI	RLRPGGKKRY	RLKHLVWASR	ELENFALNPG
00BW2128_3	MGARASILRG	.GQLDKWEKI	RLRPGGKKHY	MLKHLVWASG	ELEKFALNPG
00BW2276_7	MGARASVLKG	.DKLDWEKI	KLPRPGGKKHY	MLKHLVWASR	ELERFALNPG
00BW3819_3	MGARASILRG	.GKLDAWERT	RLRPGGKKHY	RLKHLVWASR	ELERFALNPG
00BW3842_8	MGARASVLRG	.EKLDTWEKI	KLPRPGGKKHY	MLKHIVWASR	ELERFALNPG
00BW3871_3	MGARASILRG	.GKLDTWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
00BW3876_9	MGARASILKG	.GKLDTWEKI	RLRPGGKKHY	MVKHLVWASR	ELERFALNPG
00BW3886_8	MGARASILRG	.GKLDKWEKI	RLRPGGKKCY	MIKHIIWASR	ELERFALNPG
00BW3891_6	MGARASILRG	.GKLDKWEKI	RLRPGGKKKY	MLKHLVWASR	ELERFALNSG
00BW3970_2	MGARASILRG	.GKLDAWERI	KLPRPGGKKHY	MLKHLVWASR	ELERFALNPS
00BW5031_1	MGARASILRG	.GKLDRWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPS
96BW01B21	MGARASILRG	.GKLDQWEKI	RLRPGGKKCY	MLKHLVWASR	ELERFALNPG
96BW0407	MGARASILRG	.GKLDAWERI	RLRPGGKKCY	MMKHLVWASR	ELEKFALNPG
96BW0502	MGARASILRG	.EKLDKWEKI	RLRPGGKKHY	MLKHLVWASR	ELEGFALNPG
96BW06_J4	MGARASILRG	.GNLDTWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
96BW11_06	MGARASILRG	.GKLDKWEKI	RLRPGGKKRY	MIKHLVWASR	ELERFALNPG
96BW1210	MGARASILRG	.EKLDTWEKI	RLRPGGKKRY	MMKHLVWASR	ELENFALNPA
96BW15B03	MGARASISRG	.EKLDTWEKI	RSRPGGKKCY	MLKHIVWASR	ELERFALNPG
96BW16_26	MGARASILRG	.GKLDKWEKI	RLRPGGKKRY	MLKHLVWASR	ELERFALNPG
96BW17A09	MGARASILRG	.GKLDTWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
96BWM01_5	MGARASILRR	.GKLDKWEKI	RLRPGGKKRY	MIKHLVWASR	ELERYALNPG
96BWM03_2	MGARASILRG	.EKLDKWEKI	RLRPGGKKHY	MLKHIVWASR	ELEKFALNPG
98BWMC12_2	MGARASILRG	.EKLDTWEKI	RLRPGGKKQY	RIKHLVWASR	ELDRFALNSG
98BWMC13_4	MGARASILRG	.GKLDKWEKI	RLRPGGKKHY	MMKHLVWASR	ELGRFALNPG
98BWMC14_a	MGARASILRG	.GKLDKWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
98BWM014_1	MGARASILRG	.GKLDKWEKI	RLRPGGKKQY	RLKHLVWASR	ELERFALNPG
98BWM018_d	MGARASILRG	.GKLDTWEKI	KLPRPGGKKHY	MMKHLVWASR	ELERFALNPS
98BWM036_a	MGARASILRG	.GKLDTWEKI	RLRPGGKKPY	MLKHLVWASR	ELERFALNPG
98BWM037_d	MGARASILRG	.GKLDWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
99BW3932_1	MGARASILRG	.GKLDWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPE
99BW4642_4	MGARASVLKG	.EKLDTWEKI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
99BW4745_8	MGARASILRG	.GKLGNERI	KLPRPGGKTY	MLKHLVWASR	ELERFALNPS
99BW4754_7	MGARASILRG	.EKLDRWEKI	WLRPGGKNHY	MLKHLVWASR	ELERFALNPA
99BWMC16_8	MGARASILRG	.GKLDTWEKI	RLRPGGKKHY	MIKHLVWASR	ELERFALNPG
A2_CD_97CD	MGARASVLSG	.GKLEAWEKI	RLRPGGKKKY	RLKHLVWASR	ELEKFSINPS
A2_CY_94CY	MGARASILSG	.GKLDWEKI	RLRPGGKKKY	RLKHLVWASR	ELEKFSINPG
A2D_97KR	MGARASVLSG	.GKLDWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
A2G_CD_97C	..ARASILSG	.GKLEAWEKI	RLRPGGKKKY	RLKHLVWASR	ELEKFSINPG
A_BY_97BL0	XGARASVLSG	.GKLEAWEKI	RLRPGGKKKY	RIKHLVWASR	ELERFALNPG
A_KE_Q23_A	MGARASVLSG	.GKLEAWEKI	RLRPGGKKKY	RMKHLIWASR	ELDRFALNPS
A_SE_SE659	...RASVLSG	.GKLDWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPS
A_SE_SE725	...RASVLSG	.GKLDWEKI	RLRPGGKKKY	RMKHLVWASR	ELERFALNPS
A_SE_SE753	MGARASVLSG	.GKLDWEKI	RLRPGGKKQY	RLKHLVWASR	ELERFALNPS
A_SE_SE853	...RASVLSG	.GRLDAWEKI	RLRPGGKKKY	RMKHLVWASR	ELDRFALNPS
A_SE_SE889	...RASKLSG	.EKLEAWEKI	RLRPGGKKKY	KLKHMVWARR	ELEKSALNPS
A_SE_UGSE8	MGARASVLSG	.GKLDWEKI	RLRPGGKNKY	RLKHLVWASR	ELEKFALNPG
A_UG_92UG0	MGARASVLSG	.GKLDWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPS
A_UG_U455_	MGARASVLSG	.KKLDSWEKI	RLRPGGKNKY	RLKHLVWASR	ELEKFTLNPG
AC_IN_2130	MGARASILRG	.GKLDKWEKI	RLRPGGKKHY	MIKHLVWASR	ELERFALNPG
AC_RW_92RW	MGARASILRG	.GKLDWEKI	KLKPGGKKTY	MMKHLVWASR	ELERFALNPD
AC_SE_SE94	...RASVLSG	.GKLDWEKI	RLRPGGKKKY	RMKHLVWASR	ELDRFALNPS
ACD_SE_SE8	MGARASILSG	.GKLDWEKI	RLRPGGKKKY	RLKHLVWASR	ELDRFALNPS

ACG_BE_VI1	MGARASVLTG	.GKLDTWERI	RLRPAGKKKY	RMKLLVWASR	ELERFAINPG
AD_SE_SE69	MGARASVLSG	.ERLDEWEKI	QLRPGGKKRY	RLKHIVWASR	ELERFALNPG
AD_SE_SE71	...RASVLSG	.GKLDaweKI	RLRPGGRKKY	KLKHIVWASR	ELERFALNPS
ADHK_NO_97	MGARASILSG	.GKLDKWEKI	RLRPGGKKQY	RLKHLVWASR	ELDRFALNPS
ADK_CD_MAL	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
AG_BE_VI11	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	RMKHLVWASR	ELERFAINPG
AG_NG_92NG	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	RMKHLVWASR	ELERFALNPD
AGHU_GA_VI	MGARASILSG	.GKLDaweKI	RLRPGGKKKY	QIKHIVWASR	ELERFALNPG
AGU_CD_Z32	MGARASVLSG	.GKLDKWEKI	RLRPGGKKQY	KLKHIVWASR	ELERFALNPG
AJ_BW_BW21	MGARASVLSG	.GKLDaweKI	RLRPGGKKQY	RMKHLVWASR	ELERFALNPG
B_AU_VH_AF	MGARASVLSG	.GELDRWEKI	RLRPGGKKTY	KLKHIVWASR	ELERFAVNPG
B_CN_RL42_	MGARASVLSG	.GQDRWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFAVNPG
B_DE_D31_U	MGARASVLSG	.GELDRWEKI	RLRPGGKKKY	RLKHIVWASR	ELERFAVNPG
B_DE_HAN_U	MGARASVLSG	.GELDKWEKI	RLRPGGKKKY	QLKHIVWASR	ELERFAVNPG
B_FR_HXB2_	MGARASVLSG	.GELDRWEKI	RLRPGGKKKY	KLKHIVWASR	ELERFAVNPG
B_GA_OYI_	MGARASVLSG	.GELDKWEKI	RLRPGGKKKY	QLKHIVWASR	ELERFAINPG
B_GB_CAM1_	MGARASVLSG	.GELDKWEKI	RLRPGGKKKY	KLKHIVWASR	ELERFAVNPG
B_GB_GB8_A	MGARASVLSG	.GELDRWEKI	RLRPGGKKKY	RLKHVVWASR	ELERFAVNPG
B_GB_MANC_	MGARASVLSG	.GKLDREKI	RLRPGGKKKY	KLKHIVWASR	ELERFTVNLG
B_KR_WK_AF	MGARASILSG	.GELDQWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFAVNPG
B_NL_3202A	MGARASVLSG	.GELDKWEKI	RLRPGGKKRY	KLKHIVWASR	ELERFAVNPG
B_TW_TWCYS	MGARASILSA	.GELDKWERV	RLRPGGKKKY	RLKHLVWASR	ELERFAVNPG
B_US_BC_L0	MGARASVLSG	.GKLDKWEKI	RLRPGGKKKY	KLKHLVWASR	ELERFAVNPG
B_US_DH123	MGARASVLSG	.GKLDsWEKI	RLRPGGKKKY	KLKHIVWASR	ELERFAVNPG
B_US_JRCsf	MGARASVLSG	.GELDRWEKI	RLRPGGKKKY	RLKHIVWASR	ELERFAVNPG
B_US_MNCG_	MGARASVLSG	.GELDRWENI	RLRPGGKKKY	KLKHVVWASR	ELERFAVNPG
B_US_P896_	MGARASVLSG	.GELDRWEKI	RLRPGGKKKY	KLKHIVWASR	ELERFAVNPS
B_US_RF_M1	MGARASVLSG	.GKLDKWEKI	RLRPRGKKRY	KLKHIVWASR	ELERFAVNPS
B_US_SF2_K	MGARASVLSG	.GELDKWEKI	RLRPGGKKKY	KLKHIVWASR	ELERFAVNPG
B_US_WEAU1	MGARASVLSG	.GELDRWEKI	RLRPGGNKKY	KLKHIVWASR	ELERFAVNPG
B_US_WR27_	MGARASMLSG	.GELDKWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFAVNPG
B_US_YU2_M	MGARASVLSA	.GELDKWEKI	RLRPGGKKQY	RLKHIVWASR	ELERFAVDPG
BF1_BR_93B	MGARASVISG	.GELDKWEKI	RLRPGGHKKY	RLKHIVWASR	ELERFAVNPG
C_BR_92BR0	MGARASILRG	.GKLDawERI	KLKPGGKKHY	MMKHLVWASR	ELERFALDPG
C_BW_96BW0	MGARASILRG	.GKLDaweKI	RLRPGGKKQY	RIKHLVWASR	ELERFALNPG
C_BW_96BW1	MGARASILRG	.GKLDTWEKI	RLRPGGKKRY	MIKHLVWASR	ELERFALNPG
C_BW_96BW1	MGARASILRG	.EKLDTWEKI	RLRPGGKKRY	MMKHLVWASR	ELENFALNPA
C_BW_96BW1	MGARASISRG	.EKLDTWEKI	RSRPGGKKCY	MLKHIVWASR	ELERFALNPG
C_ET_ETH22	MGARASILRG	.EKLDaweKI	KLKPGGKKHY	MLKHLVWANR	ELEKFALNPD
C_IN_93IN1	MGARASILRG	.GKLDKWERI	RLRPGGKKHY	MLKHLVWASR	ELERFALNPG
C_IN_93IN9	MGARASILRG	.EKLDKWEKI	RLRPGGKKHY	MLKHLVWASR	ELDRFALNPG
C_IN_93IN9	MGARASILRG	.EKLDKWERI	RLRPGGKKHY	MLKHLVWASR	ELDRFALNPG
C_IN_94IN1	MGARASILRG	.GKLDKWEKI	RLRPGGKKHY	MLKHLVWASR	ELEKFALNPG
C_IN_95IN2	MGARASILRG	.GKLDKWEKI	RLRPGGKKRY	MLKHLVWASR	ELDRFAVNPG
CRF01_AE_C	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	RMKHLVWASR	ELERFALNPG
CRF01_AE_C	MGARASILSG	.GKLDaweKI	RLRPGGKKQY	RMKHLVWASR	ELERFALNPG
CRF01_AE_C	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF01_AE_T	MSARASVLSG	.EKLNAWEKI	RLRPGGRKKY	KLKHLVWASR	ELEKFALNPG
CRF01_AE_T	MGARASVLSG	..KLDaleKI	RLRPGGKKKY	KMKHLVWASR	ELERFALNPG
CRF01_AE_T	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	KMKHLVWASR	ELERFALNPG
CRF01_AE_T	MGAKASVLSG	.GKLDaweKI	RLRPGGRKKY	HLKHIVWASR	ELERFALNPG
CRF01_AE_T	MGARASVLSG	.GKLDaweKI	RLRPGGRKKY	RLKHLVWASR	ELERFALNPS
CRF01_AE_T	MGARASILSG	.GKLDaweKI	RLRPGGRKKY	RMKHLVWASR	ELERFALNPG
CRF02_AG_F	MGARASVLSG	.GKLDsWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF02_AG_F	MGARASVLSG	.GKLDsWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF02_AG_G	MGARVSVLRG	.GQLDTWEKI	RLRPGGKKKY	KMKLLVWASR	ELERFAVNPG
CRF02_AG_N	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF02_AG_S	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF02_AG_S	.GARASVLSG	.GKLDawERI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF03_AB_R	MGARASVLSG	.GKLDaweKI	RLRPGGKEKY	RIKHLVWASR	ELERFALNPS
CRF03_AB_R	MGARASVLSG	.GKLDaweKI	RLRPGGKKKY	RIKHLVWASR	ELERFAINPS

CRF04_cpx_	MGARASVLSG	.GKLDAWERI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF04_cpx_	MGARASVLSG	.GKLDAWERI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF04_cpx_	MGARASVLSG	.GRLDAWEKI	RLRPGGKKRY	RIKHLIWASR	ELERFALNPG
CRF05_DF_B	MGARASVLSG	.GKLDAWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
CRF05_DF_B	MGARASVLSG	.GKLDAWEKI	RLRPGGKKKY	RLKHIVWASR	ELERFAINPG
CRF06_cpx_	MGARASVLSG	.GKLDEWEKI	RLRPGGKKKY	KMKHLVWASR	ELERFAINPG
CRF06_cpx_	MGARASVLSG	.GKLDEWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPS
CRF06_cpx_	MGARASVLSG	.GKLDEWEKI	RLRPGGKKKY	RMKHLVWASR	ELERFALNPG
CRF06_cpx_	MGARASVLSG	.GKLDEWEKI	RLRPGGKKKY	RLKHLVWASR	ELDRFALNPG
CRF11_cpx_	MGARASVLSG	.GKLDSWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPS
CRF11_cpx_	MGARASVLSG	.GKLDAWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPS
D_CD_84ZR0	MGARASVLSG	.GKLDAWEKI	RLRPGGKKKY	KLKHIVWASR	ELERFALNPG
D_CD_ELI_K	MGARASVLSG	.GKLDKWEKI	RLRPGGKKKY	RLKHIVWASR	ELERYALNPG
D_CD_NDK_M	MGARASVLSG	.GKLDTWERI	RLRPGGKKKY	ALKHLIWASR	ELERFTLNPG
D_UG_94UG1	MGARASVLSG	.GKLDEWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
F1_BE_VI85	MGARASILSG	.GKLDEWEKI	QLRPGGKKRY	KMKHLIWASR	ELERFALDPG
F1_BR_93BR	MGARASVLSG	.GKLDAWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALDPG
F1_FI_FIN9	MGARASVLSG	.GKLDAWEKI	RLRPGGKKQY	RIKHLVWASR	ELERFAIDPG
F1_FR_MP41	MGARASVLSG	.GKLDAWERI	RLRPGGKKKY	RMKHLVWASR	ELERFAVDPG
F2_CM_MP25	MGARASVLSG	.GKLDAWEKI	RLRPGGKKKY	RLKHIVWASR	ELKRFALNPG
F2KU_BE_VI	MGARASVLSG	.GKLDSWEKI	RLRPGGRKKY	RLKHLVWASR	ELERFALNPG
G_BE_DRCBL	MGARASVLSG	.GKLDAWEKI	RLRPGGKKRY	RMKHLVWASR	ELDRFALNPG
G_NG_92NG0	VGARASVLSG	.GKLDSWEKI	RLRPGGRKKY	KLKHIVWASR	ELGRFALNRD
G_SE_SE616	MGARASVLTG	.GKLDAWEKI	RLRPGGRKSY	KIKHLVWASR	ELERFALNPD
H_BE_VI991	MGARASVLSG	.GKLDAWEKI	RLRPGGRKKY	RLKHLVWASR	ELERFALNPD
H_BE_VI997	MGARASVLSG	.GRLDTLEKI	RLRPGGKKKY	RLKHIVWASR	ELERFALNPG
H_CF_90CF0	MGARASVLSG	.GKLDAWEKI	RLRPGGKKKY	RLKHLVWASR	ELERFALNPG
J_SE_SE702	MGARASILSG	.GKLDDWEKI	RLRPGGKKQY	RIKHLVWASR	ELDRFALNPG
J_SE_SE788	MGARASILSG	.GKLDDWEKI	RLRPGGKKKY	RIKHLVWASR	ELDRFALNPG
K_CD_EQTB1	MGARASVLSG	.GKLDKWEKI	QLRPGGKKKY	RLKHLVWASR	ELERFALNPN
K_CM_MP535	MGARASVLSG	.GKLDAWEKI	RLRPGGKKKY	KLKHLVWASR	ELERFALNPG
N_CM_YBF30	MGARASVLTG	.GKLDQWEST	YLRPGGKKKY	RMKHLVWASR	ELERFACNPG
O_CM_ANT70	MGASASVLTG	.SKLDAWEQI	RLKPGSKKKY	RLKHLVWASR	ELERFACNPE
O_CM_MVP51	MGARASVLTG	.SKLDAWERI	RLRPGSKKAY	RLKHLVWASR	ELERYACNPG
O_SN_99SE_	MGARASVLTG	.SKLDAWEQI	RLKPGCKKKY	RLKHLVWASR	ELDRSACNPE
O_SN_99SE_	MGARASVLSG	.SKLDTWEQI	RLKPGCKKKY	RLKHLVWASR	ELERFACNPE
U_CD_83C	MGARASVLSG	.GKLDAWEKI	RLRPGGRKKY	RLKHLVWASR	ELEKFALNPG

51

100

00BW0762_1	LLETSEGCKQ	IIKQLQPALQ	TGTEELRSLY	NTVATLYCVH	KKIDVRDTKE
00BW0768_2	LLETSEGCKQ	IIKQLQPALQ	TGTEELRSLF	NTVATPYCVH	EKIEVRDTKE
00BW0874_2	LLETAEGCRQ	IIKQLHPALQ	TGTEELRSLY	NTVATLYCVH	RGIEIRDTE
00BW1471_2	LLETADGCKQ	IIKQLQPALQ	TGTEELRSLF	NTVATLYCVH	KGIKVQDTKE
00BW1616_2	LLETSDGCKQ	IMKQLQPALQ	TGTEELKSLF	NTVATLYCVH	ANIDVRDTKE
00BW1686_8	LLETSDGCKQ	IIQQQLQPALK	TGTEELRSLF	NTVATLYCAH	KGISIQDTKE
00BW1759_3	LLETAEGCKR	IIKQLQPALQ	TGTEELKSLH	NTVATLYCVH	KEIDVRDTKE
00BW1773_2	LLETAEGCKQ	IIKQLQPALQ	TGTEELKSLY	NTVATLYCVH	AGIEVRDTKE
00BW1783_5	LLETSEGCKQ	IIQQQLQPALK	TGTEELRSLY	NTVATLYCVH	AKIEVRDTKE
00BW1795_6	LLETSEGCKQ	IMKQLQPAIQ	TGTEELRSLF	NTVATLYCVH	EGIDVQDTKE
00BW1811_3	LLETAAGCKQ	IIRQLHPALQ	TGTEELRSLY	NTVATLYCVH	AEIEVRDTKE
00BW1859_5	LLETSEGCRQ	IMRQLQPALQ	TGTEELRSLF	NTVATLYCVH	EKIPVRDTKE
00BW1880_2	LLETAEGCKQ	IIRQLHPALQ	TGTQELRSLY	NTVATLYCVH	KDIEVRDTKE
00BW1921_1	LLETSEGCKQ	IIQQQLQPALQ	TGTEELRSLY	NTVATLYCVH	KGIEVQDTKE
00BW2036_1	LLETSEGCKQ	IMKQLQPALQ	TGTEELRSLF	NTVATLYCVH	KKIQVQDTKE
00BW2063_6	LLETADGCKQ	IMKQLQPALQ	TGTEELRSLY	NTVATLYCVH	EGIDVRDTKE
00BW2087_2	LLETSEGCKQ	IIKQLQPALQ	TGTEELRSLF	NTVATLYCVH	KGIEVRDTKE
00BW2127_2	LLETSEGCRQ	IIKQLQPALQ	TGTEELRSLY	NTIATLYCAH	QRIKVQDTKE
00BW2128_3	LLETSEGCKQ	IIKQLHPALQ	TGTEELKSLY	NTVATLYCVH	AGIAVRDTKE
00BW2276_7	LLETSEGCKQ	IMQQQLQPALK	TGTEELISLY	NTVATLYCVH	EKIQVQDTKE
00BW3819_3	LLETSEGCKQ	IIKQLQPALQ	TGTEELRSLY	NTVATLYCVH	QKIEIRDTE
00BW3842_8	LMETSEGCKQ	IIQQQLQPALQ	TGTEELRSLF	NTIATLYCVH	KGIKVQDTKE

00BW3871_3	LLETSDGCKQ	IIKQLQPALQ	TGTEELRSLF	NTVATLYCVH	KGIKVQDTKE
00BW3876_9	LLETSEGCEQ	IMKQLQPALQ	TGTEELRSLY	NTVATLYCVH	ARIEVKDTKE
00BW3886_8	LLETAEGCTQ	IMKQLQPALQ	TGTEELRSLF	NTVATLYCVH	AGIDVRDTKE
00BW3891_6	LLETSDGCKQ	IIQQLQPALK	TGTEELRSLY	NTVATLYCVH	NNIEIRDTE
00BW3970_2	LLETSEGCKQ	IIKQLQPALQ	TGTEELKSLF	NTVATLYCVH	EKIEIRDTE
00BW5031_1	LLETADGCKQ	IIKQLQPALK	TGTEELRSLY	NTVATLYCVH	RNIEVQDTKE
96BW01B21	LLETSEGCKQ	IMEQLQPALQ	TGTEELKSLY	NTVATLYCVH	ENIEVRDTKE
96BW0407	LLETSEGCKQ	IIRQLHPALQ	TGSEELKSLF	NTVATLYCVH	EKIEVRDTKE
96BW0502	LLETAEGCKQ	IMKQLQPALQ	TGTEELRSLY	NTVATLYCVH	AGIEVRDTKE
96BW06_J4	LLETSEGCKQ	IMQQLQPALQ	TGTEELRSLF	NTVATLYCVH	KGIKVQDTKE
96BW11_06	LLETSEGCKQ	IMKQLQPALQ	TGTEELRSLY	NTVATLYCVH	EGIEVQDTKE
96BW1210	LLETAEGCKQ	IMKQLQPALK	TGTEELKSLY	NTVATLYCVH	EKIEVRDTKE
96BW15B03	LLETSEGCKQ	IMKQLQPALQ	TGTEELRSLF	NTVATLYCVH	EKIKVKDTKE
96BW16_26	LLETSEGCRQ	IMIQLQPALK	TGTEELKSLY	NTVAVLFCVH	AGIEVQDTKE
96BW17A09	LLETAEGCKQ	IIKQLQPALQ	TGTEELKSLY	NTVATLYCVH	QGIEVRDTKE
96BWM01_5	LLETSEGCKQ	IIKQLQPTIQ	TGTEELKSLY	NTVATLYCVH	KGIDVRDTKE
96BWM03_2	LLETSEGCKQ	IIKQLQPALQ	TGTEELRSLY	NTVATLYCVH	EKIEVRDTKE
98BWMC12_2	LLETAEGCTQ	IMRQLQPALQ	TGTEELRSLY	NTVATLYCVL	KGIKVRDTKE
98BWMC13_4	LLETSEGCKQ	IMRQLQPAIQ	TGTEELRSLF	NTVATLYCVH	EGIDVKDTKE
98BWMC14_a	LLETSEGCKQ	IMKQLQPALQ	TGTEELRSLF	NTVATLYCVH	MGIEVRDTKE
98BWM014_1	LLETAEGCKQ	IIQQLQPALK	TGTEELRSLF	NTVATLYWVH	KNVEVRDTKE
98BWM018_d	LLETSEGCRK	IMKQLQPSIQ	TGTEELRSLY	NTIATLYCVH	EKIEVRDTKE
98BWM036_a	LLETAEGCKQ	ILQQLQPALK	TGTEELKSLY	NTVATLYCVH	QGIEVRDTKE
98BWM037_d	LLENSEGCKQ	IMKQLQPALQ	TGTEELKSLF	NTVATLYCVH	EKIEVRDTKE
99BW3932_1	LLETADGCKQ	IIKQLQPAFQ	TGTEELRSLY	NSIAVLYCVH	RRLTIQDTKE
99BW4642_4	LLETSEGCKQ	IMKQLQPALQ	TGTEELRSLY	NTVATLYCVH	ERIQVRDTKE
99BW4745_8	LLETTEGCRQ	IIRQLQPALQ	TGTEELKSLY	NTVATLYCVH	NNIEIRDTE
99BW4754_7	LLETAEGCKQ	IIKQLHPALQ	TGTEELRSLY	NTVATLYCVH	NKIEVQDTKE
99BWMC16_8	LLETSEGCKQ	IMKQLQPALQ	TGTEELRSLY	NTIATLYCVH	ESIEVRDTKE
A2_CD_97CD	LLETETGCRR	IFGQLQPALE	TGTEELRSLY	NTIAVLYFVH	QKIEVKDTKE
A2_CY_94CY	LLETPEGCRQ	IIRQLQPALQ	TGTEELKSLY	NTVVVLYWVH	QRVDVKDTKE
A2D_97KR	LLETSEGCKQ	IIGQLEPSIK	TGSEELKSLF	NAVVTLYCVH	QRIEVRDTKE
A2G_CD_97C	LLETAAGCRQ	ILGQLQPALQ	TGTEELRSLY	NTVAVLYCVH	QKIEVKDTKE
A_BY_97BL0	LLETSEGCKQ	ILEQLQPTLK	TGSEELKSLY	NTVATLYCVH	QRIEIKDTKE
A_KE_Q23_A	LLETADGCQQ	IMEKLQPALK	TGTEELKSLF	NTVATLYCVH	QRIDVKDTKE
A_SE_SE659	LLETTEGCQQ	IIEQLQPALN	TGTEELKSLF	NTVATLYCVH	QRIDVKDTKE
A_SE_SE725	LLETTEGCQQ	IMEQLQSALK	TGTEELRSLF	NTVAVLYCVH	QRIEIKDTKE
A_SE_SE753	LLETAEGCQQ	IMKQLQPALK	TGTEELRSLY	NTVATLYCVH	QRINVTDTKE
A_SE_SE853	LLETTEGCQQ	IIGQLQPAFK	TGTEELKSLY	NTVATLWCVH	QRIDVKDTKE
A_SE_SE889	LLETTEGCQQ	IMDKLQPALR	TGTEELRSLY	NTVATLYCVH	QKIAVKDTKE
A_SE_UGSE8	LLETAEGCQQ	IIEQLQSAPK	TGTEELRSLY	NTVATLYCVH	QRIDVKDTKE
A_UG_92UG0	LLETTEGCQQ	IMEQLQSALR	TGTEELRSLY	NTVATLYCVH	QRIEIKDTKE
A_UG_U455_	LLETAEGCQQ	ILGQLQPALQ	TGTEELRSLY	NTVAVLYCVH	QRIDVKDTKE
AC_IN_2130	LLETSEGCKQ	IIKQLQPALQ	TGTEELRSLH	NTVATLYCVH	AGIEIRDTE
AC_RW_92RW	LLETPEGCKQ	IMRQLQPALQ	TGTEELRSLY	NTVATLYCVH	QKIDVKDTKE
AC_SE_SE94	LLETSEGCKQ	ILEQLQPALK	TGTEELKSLF	NTVATLYCVH	QRIEIKDTKE
ACD_SE_SE8	LLETTEGCQQ	IMDQLQPALK	TGTEELRSLY	NTVATLWCVH	KRIDVKDTKE
ACG_BE_VI1	LLETAEGCQQ	IIEQLQSTLK	TGSEELKSLF	NTVATLWCVH	LRIEIKDTKE
AD_SE_SE69	LLETPEGCKQ	IMGQLQPAIQ	TGSEELKSLF	NTVATLYCVH	AKIKVTDTKE
AD_SE_SE71	LLETTKGCQQ	IMEQLQPAFK	TGTEELKSLY	NTVATLYCVH	QQIEVKDTKE
ADHK_NO_97	LLETSEGCKQ	VMEQLQPALK	TGSEELKSLF	NLVAVLWCVH	QRIDVRDTKE
ADK_CD_MAL	LLETGEGCQQ	IMEQLQSTLK	TGSEELKSLY	NTVATLYCVH	QRIDVKDTKE
AG_BE_VI11	LLETAEGCQQ	IIEQLQSTLR	TGSEELKSLY	NTLATLWCVH	KKIEVRDTKE
AG_NG_92NG	LLETTEGCQQ	IMRQLQPSLQ	TGTEELKSLF	NTVATLYCVH	QRIEIKDTKE
AGHU_GA_VI	LLETAEGCQQ	IMEQLQSALK	TGSEELKSLF	NTVATLYCVH	QKIDVKDTKE
AGU_CD_Z32	LLETTEGCQQ	ILSOFQPAK	TGTEELISLY	NTVATLYCVH	QNIEVKDTKE
AJ_BW_BW21	LLETAEGCQQ	IMEQLQSALK	TGSEELRSLY	NTVATLYCVH	QRIEIKDTKE
B_AU_VH_AF	LLETSEGCRQ	ILVQLQPSLP	TGSEELKSLF	NTVATLYCVH	QKIEIRDTE
B_CN_RL42_	LLETSEGCRQ	ILEQLQPALQ	TGSEELRSLF	NTVATLYCVH	LRIEIKDTKE
B_DE_D31_U	LLETSEGCRQ	ILGQLQPSLQ	TGSEELRSLF	NTVATLYCVH	QRIEIKDTKE
B_DE_HAN_U	LLETSEGCRQ	IMGQLQPSLQ	TGSEELRSLY	NTVATLYCVH	QKIEVKDTKE

B_FR_HXB2_	LLETSEGCRO	ILGQLQPSLQ	TGSEELRSly	NTVATLYCVH	QRIEIKDTKE
B_GA_OYI_	LLETSEGCRO	ILGQLQPSLK	TGSEEIRSLY	NTVATLYCVH	QKIEVKDTKE
B_GB_CAM1_	LLETSEGCRO	ILGQLQPSLQ	TGSEELRSly	NTVATLYCVH	QKIDVKDTKE
B_GB_GB8_A	LLETSEGCRO	ILEQLQPSLQ	TGSEERRSLF	NTVATLYCVH	QRIDVKDTKE
B_GB_MANC_	LLETSEGCRO	ILVQLQPSLQ	TGSEELKSLY	NTVAVLYCVH	QGIEVKDTKE
B_KR_WK_AF	LLETSEGCRO	ILGQLQPSLQ	TGSEELKSLF	NAVAVLYCVH	QRIEIKDTKE
B_NL_3202A	LLETAEGCRO	ILGQLQPALQ	TGSEELRSly	NTVAVLYCVH	QRIEVKDTKE
B_TW_TWCYS	LLETSEGCRO	ILGQLQSSLQ	TGSEELRSly	NTVATLYCIH	QKIEIKDTKD
B_US_BC_L0	LLETAEGCRO	ILGQLQPSLQ	TGSEELKSLY	NTIAVLYCVH	QMIDVKDTKE
B_US_DH123	LLETSEGCRO	ILGQLQPSLQ	TGSEELRSly	NTVATLYCVH	ERIEVKDTKE
B_US_JRCSF	LLESSEGCRO	ILGQLQPSLK	TGSEELTSly	NTVATLYCVH	QRIEIKDTKE
B_US_MNCG_	LLETSEGCRO	ILGQLQPSLQ	TGSEELKSLY	NTVATLYCVH	QKIEIKDTKE
B_US_P896_	LLETSEGCRO	ILGQLQSSLQ	TGSEELKSLY	NTVATLYCVH	QRIEVKDTKE
B_US_RF_M1	LLETAEGCRO	ILGQLQPALQ	TGSEELKSLY	NAVATLYCVH	QNIEVRDTKE
B_US_SF2_K	LLETSEGCRO	ILGQLQPSLQ	TGSEELRSly	NTVATLYCVH	QRIDVKDTKE
B_US_WEAU1	LLETSEGCRO	ILGQLQPSLQ	TGSEELRSly	NTVAVLYCVH	QKIEVKDTKE
B_US_WR27_	LLETAEGCRO	ILRQLQPSLQ	TGSEELRSly	NTVATLYCVH	QRIGVKDTKE
B_US_YU2_M	LLETSEGCRO	ILGQLQPSLQ	TGSEELRSly	NTVATLYCVH	QKIEVKDTKE
BF1_BR_93B	LLETSEGCRO	ILEQLQPALK	TGSEELRSly	NTVATLYCVH	QKIDVKDTKE
C_BR_92BR0	LLETSEGCKQ	IMKQLQPALQ	TGTKEELSLH	NTVATLYCVH	EKIDVRDTKE
C_BW_96BW0	LLETSEGCKQ	IIQQLQPALQ	TGTEELRSly	NTVATLYCVH	KKIEVRDTKE
C_BW_96BW1	LLGTSEGCKQ	IMKQLQPALQ	TGTEEIRSLY	NTVATLYCVH	EGIEVQDTKE
C_BW_96BW1	LLETAEGCKQ	IMKQLQPALK	TGTEELKSLY	NTVATLYCVH	EKIEVRDTKE
C_BW_96BW1	LLETSEGCKQ	IMKQLQPALQ	TGTEELRSly	NTVATLYCVH	EKIKVKDTKE
C_ET_ETH22	LLDTSAGCKQ	IIKQLQPALQ	TGTEELKSLF	NTVATLYCVH	QKIEIKDTKE
C_IN_93IN1	LLETAEGCKQ	IIKQLQPALQ	TGTEELKSLH	NTVATLYCVH	AGIEVRDTKE
C_IN_93IN9	LLETSEGCKQ	IIKQLQPALQ	TGTEELRSLH	NTVVTLYCVH	AGIEVRDTKE
C_IN_93IN9	LLEAAEGCKQ	IIKQLQPALQ	TGTEELRSLH	NTVATLYCVH	EGIEVRDTKE
C_IN_94IN1	LLETSEGCKQ	IMKQLQPALQ	TGTEELRSly	NTVATLYCVH	AEIEVRDTKE
C_IN_95IN2	LLETAEGCKQ	IIKQLQPALQ	TGTEELRSly	NTVATLYCVH	AGIEVRDTKE
CRF01_AE_C	LLETAEGCQQ	LIEQLQSNLK	TGSEEIKSLF	NTIATLWCVH	QRIDVKDTKE
CRF01_AE_C	LLETAEGCQQ	LIEQLQSTIK	TGSEELKSLF	NTIATLWCVH	QRIDVKDTKE
CRF01_AE_C	LLETAEGCQQ	LLEQLQSTLK	TGSEELKSLF	NTIATLWCVH	QRINVTDTKE
CRF01_AE_T	LLETAEGCQQ	LIEQLQSTLK	TGSEELKSLF	NTIATLWCVH	QKIEVKDTKK
CRF01_AE_T	LLETAEGCQQ	IIQLQOSALK	TGSEELKSLY	NTVVTLWCVH	QRIDVKDTKE
CRF01_AE_T	LLETAEGCQQ	LIEQLQSTLK	TGSEELKSLY	NTIATLWCVH	QRIEVKDTKE
CRF01_AE_T	LLETAEGCQQ	LIEQLQSTLK	TGSEELRSly	NTIVTLWCVH	QRIEVKDTKE
CRF01_AE_T	FLETAEGCQQ	IIQLQOSALK	TGLEELKSLF	NTVATLWCVH	QRIEVKDTKE
CRF01_AE_T	LLETAEGCQQ	LIGQLQSNLK	TGSEELRSly	NTVATLWCVH	QRIEVKDTKE
CRF02_AG_F	LLETAEGCQQ	LIEQLQTAGL	TGSEELKSLY	NTIATLWCVH	KRIDIKDTKE
CRF02_AG_F	LLETAEGCQQ	LMKQLQSALG	TGSEELKSLY	NTIATLWCVH	QRIDIKDTKE
CRF02_AG_G	LLETTEGCQQ	ILEQLQPALK	AGSEELKSLY	NTVATLYCVH	QKIDVRDTKE
CRF02_AG_N	LLETAEGCQQ	LMEQLQSALR	TGSEELKSLF	NTIATLWCVH	QRIDIKDTKE
CRF02_AG_S	LLETAEGCQQ	LMEQLQSTLR	TGSEELKSLF	NTIATLWCVH	QRIDIKDTKE
CRF02_AG_S	LLETAEGCQQ	IIQLQOSALK	TGSEELKSLF	NTVATLYCVH	QRIEVKDTKE
CRF03_AB_R	LLETSEGCQQ	ILEQLQPTLK	TGSEELKSLY	NTVATLYCVH	QRIEIKDTKE
CRF03_AB_R	LLETSEGCQQ	ILEQLQPTLK	TGSEELKSLY	NTVATLYCVH	RRIEIKDTKE
CRF04_cpx_	LLETAEGCQQ	LMEQLQSTLK	TGSEELRSly	NTITTLWCVH	QRIDVQDTKE
CRF04_cpx_	LLETAEGCQQ	LMEQLQPTLR	TGSEEVKSLF	NTLATLWCVH	QRIDVKDTKE
CRF04_cpx_	LLETAEGCQQ	LMEQLQSTLK	TGSEELKSLF	NLIATLWCVH	QRIEVKDTKE
CRF05_DF_B	LLETSEGCKQ	IISQLQPAIQ	TGTEELKSLY	NTVATLYCVH	ERIEVKDTKE
CRF05_DF_B	LLETSEGCKQ	IIAQQLQPAIQ	TGSEELRSly	NTVATLYCVH	EKIEVKDTKE
CRF06_cpx_	LLETAEGCQQ	IIQLQOSALK	TGSEELKSLF	NTVATLYCVH	QRIKVTDTKE
CRF06_cpx_	LLETAEGCQQ	IMEQLQSALR	TGSEELKSLY	NTVATLYCVH	QRIKVTDTKE
CRF06_cpx_	LLETPEGCQQ	IMEQLQSTLR	TGSEELKSLY	NTVATLYCVH	QRIKVTDTKE
CRF06_cpx_	LLETAEGCQQ	IIQLQOSALK	TGSEELRSly	NTVATLYCVH	QRIKVQDTKE
CRF11_cpx_	LLETGEGCQQ	LMGQLQPALG	TGTEELRSly	NTLATLYCVH	HRIEIKDTKE
CRF11_cpx_	LLETGEGCKQ	IMGQLQPALG	TGTEELRSly	NTVAALYCVH	LRIEVKDTKE
D_CD_84ZR0	LLETSEGCKQ	IMEQLQPALK	TGSEELRSly	NTVATLYCVH	KRIEIKDTKE
D_CD_ELI_K	LLETSEGCKQ	IIGQLQPAIQ	TGTEELRSly	NTVATLYCVH	KGIDVKDTKE
D_CD_NDK_M	LLETSEGCKQ	IIGQLQPSIQ	TGSEEIRSLY	NTVATLYCVH	ERIEVKDTKE

D_UG_94UG1	LLETSEGCRO	IIRQLQPSIQ	TGSEELKSLY	NTVVTLYCVH	ERIKVASTKE
F1_BE_VI85	LLETSEGCQK	IIRQLQPSLQ	TGSEELKSLF	NTVAVLYYVH	QRAGVTDTK
F1_BR_93BR	LLETSEGCRC	IIGQLQPSLQ	TGSEELKSLY	NTIAYLYYVH	QKVEVKDTKE
F1_FI_FIN9	LLETSEGCQK	IIAQIQPSIQ	TGSEELRSLY	NTIAYLYFVH	QKIEVKDTKE
F1_FR_MP41	LLETPEGCKQ	IIRQLQPSLQ	TGSEELRSLF	NTVAVLYCVH	QKIEIKDTKE
F2_CM_MP25	LLETTEGCKK	IIGQLQPSLQ	TGSEELKSLF	NTIVVLYYVH	QKIEVRDTKE
F2KU_BE_VI	LLESANGCRO	IIEQIQPALQ	TGSEELRSLF	NTVVTLYWVH	QRIEVKDTKE
G_BE_DRCBL	LLETAEGCQK	IMAQLQPALQ	TGTEELKSLF	NTVATLYCVH	QKIEVRDTKE
G_NG_92NG0	LLETAEGCVQ	IMKQLQP.AL	TGTEELRSLF	NTVATLYCVH	QKIEVKDTKE
G_SE_SE616	LLETAEGCQQ	IMRQLQPSLQ	TGTEELKSLY	NAVATLYCVH	QRIEVKDTKE
H_BE_VI991	LLETADGCQQ	ILGQLQPALK	TGTEDLQSLY	NTIAYLYCVH	QRIDVKDTKE
H_BE_VI997	LLESAEGCLO	IIEQLRPSIK	TGTEELXSLF	NTVATLYCVL	QRIEVKDTKE
H_CF_90CF0	LLETPEGCLO	IIEQIQPAIK	TGTEELKSLF	NLVAVLYCVH	RKIDVKDTKE
J_SE_SE702	LLESAKGCQQ	ILVQLQPALQ	TGTEELKSLY	NTVATLYCVH	QRIEIKDTKE
J_SE_SE788	LLESAKGCQQ	ILVQLQPALQ	TGTQEIKSLY	NTVATLYCVH	QRIEIKDTME
K_CD_EQTB1	LLETVEGCRO	IIRQLQPSLQ	TGSEELRSLF	NTVATLYWVH	QSIQVRDTKE
K_CM_MP535	LLETTEGCRO	IITQIQPSIQ	TGSEELKSLY	NTIAYLYFVH	QKIEVKDTKE
N_CM_YBF30	LMDTADGCAK	LLNQLEPALK	TGSEELRSLY	NALAVLYCVH	SRIQIHNTQE
O_CM_ANT70	LLETAEGNEK	LLQQLEPALK	TGSDSLQSLW	NAIVVLWCVH	NRYKIGDTQQ
O_CM_MVP51	LLETAEGTEQ	LLQQLEPALK	TGSEDLKSLW	NAIAYVLWCVH	NRFDIRDTQQ
O_SN_99SE_	LLETAEGNEK	LLQQLEPALK	TGSESLQSLW	NTIAYVLWCVH	NRIKVEDTQQ
O_SN_99SE_	LLETAEGNEE	LLQQLEPALK	TGSESLQSLW	NTIAYVLWCVH	KRFKVEDTQQ
U_CD_83C	LLETAEGCQQ	IIGQLQPALQ	TGSEELKSIY	NLVATLYCVH	QRIEVKDTKE

101

150

00BW0762_1	ALDQIEKEQN	QSQQK.....	.TQQAETADK	K.....	VSQNYPIVQN
00BW0768_2	ALDKMEEEQN	KSQQK.....	TQQAEEA.AG	K.....	VSQNYPIVQN
00BW0874_2	ALDKIEKEQN	ESQQ	KTQQAKAADG	K.....	VSQNYPIVQN
00BW1471_2	ALDKIEEEQN	K.....	CQQQAKAAEG	K.....	VSQNYPIVQN
00BW1616_2	ALDKLEEEQN	KSQQE.....	TQQQTKAADG	KIS.....	..HNYPIVQN
00BW1686_8	ALDKIEEEQN	KSQQ.....	KTQQVAAAAG	Q.....	VSQNYPIVQN
00BW1759_3	ALDKIEEEQK	KSQQK.....	.LQQEEAADR	KV.....S	..QNYPIVQN
00BW1773_2	ALDKIEEEQN	KCQQK.....	.TQQAKEADG	K.....	VSQNYPIVQN
00BW1783_5	ALDKIEEEQN	KCQQK.....	.TQQAECTQQ	AGAA...GGK	VSQNYPIVQN
00BW1795_6	ALDKVEEEQN	KSQQ.....	KMQQAEEAADK	KVS.....	..QNYPIVQN
00BW1811_3	ALDKIEEEQN	KSQQK.....	.TQQAQEATA	TG.....K	VSQNYPIVQN
00BW1859_5	ALDKIEEEQN	KCQQ.....	KTQQTEAAAG	K.....	VSQNYPIVQN
00BW1880_2	ALDKIEEEQN	KSQQK.....	.AQQAEEAADK	Q.....	VSQNYPIVQN
00BW1921_1	ALDKIEEEQN	KSQQ.....	KTQQAEEAAAG	K.....	VSQNYPIVQN
00BW2036_1	ALDKIEEEQN	TCQQ.....	KTQQAEEAAAG	K.....	ISQNYPIVQN
00BW2063_6	ALDKIEEEQQ	KSQQ.....	KTQQAEEAADK	KVS.....	..QNYPIVQN
00BW2087_2	ALDKIEEEQN	KCQQK.....	.TQQAKAADK	G.....E	VSQNYPIVQN
00BW2127_2	ALDEIEEEQN	KCQQK.....	.AQQEKATDG	K	VSQNYPIVQN
00BW2128_3	ALDKIEEEQN	KIQQK.....	.TQQAKEADK	K	VSQNYPIVQN
00BW2276_7	ALDKIEEEQN	KSQQKSQQ..	KTQQAQAADG	K.....	VSQNYPIVQN
00BW3819_3	ALDKIEEEQN	KCQQK.....	.TQQGKAADK	K	ISQNYPIVQN
00BW3842_8	ALDKIEEEQN	KSQQKTQQQQ	KTQQTEAAAG	K.....	VSQNYPIVQN
00BW3871_3	ALDKIEEEQN	KSQQ.....	KTQQAEEAAAE	K.....	VSQNYPIVRN
00BW3876_9	ALDKIEEEQN	KRQQK.....	.AQQEKAADG	K	VSQNYPIVQN
00BW3886_8	ALDKIEEEQN	KSQQK.....	.TQQAQAADK	K	VSQNYPIVQN
00BW3891_6	ALDRIEEEQK	KCQQK.TQQQ	KTQQVEAADG	K.....	VSQNYPIVQN
00BW3970_2	ALDKIEEEQK	KSQQK.....	.TQQAKAADG	KI.....S	..QNYPIVQN
00BW5031_1	ALDKVEEEQN	NCQQK.....	.TQQAEEAAGG	K	ISQNYPIVQN
96BW01B21	ALDKIEEEQN	KSQQK.....	.TQQAKTDD.	.G.....K	ISPNYPIVQN
96BW0407	ALDKIEEEQN	KCQQK.....	.IQQAEEAADK	G.....K	VSQNYPIVQN
96BW0502	ALDKIEEEQN	KSQQK.....	.TQQAKEADG	K.....	VSQNYPIVQN
96BW06_J4	ALDKIEEEQN	KSQQ.....	KTQQAEEAAAG	K.....	VSQNYPIVQN
96BW11_06	AVDKIEEEQN	KSQQ.....	KTQQAEEAADK	KVS.....	..QNYPIVQN
96BW1210	ALDKIEEEQN	KCQQK.....	.EQQAEEAAAK	G.....K	VSQNYPIVQN
96BW15B03	ALDKIEEEQN	KSQE.....	KTQQAEEAAAG	K.....	ISQNYPIVQN
96BW16_26	ALDKIEEEQN	KSQQK.....	.TQQAKEATG	K.....	VSQNYPIVQN

96BW17A09	ALDKIEEEQN	K.....	FQQQAKAADG	K.....	VSQNYPIVQN
96BWM01_5	ALDKIEEEQN	KIQQ.....	KMQQAEEADK	KISQ...DKK	ISQNYPIVQN
96BWM03_2	ALDRIEEEQN	KCQQK.....	AQQQEEA.AA	K.....	VSQNYPIVQN
98BWMC12_2	ALDKIKEEQN	KSKQ	KTQQAKAADE	K.....	VSQNYPIVQN
98BWMC13_4	ALDKVEEEQN	KIQQ.....	KMQQAESADK	KVS.....	..QNYPIVQN
98BWMC14_a	ALDKIEEEQN	KVQQ.....	KTQQAEAAAAG	K.....	VSQNYPIVQN
98BWM014_1	ALDKIEEEQN	KSQQK.....	TQQAEAADKG	K.....	VSQNYPIVQN
98BWM018_d	ALDKVEEEQK	QCQQR.....	.TQQAEAADK	G.....K	VSQNYPIVQN
98BWM036_a	ALDKIEEEQN	KIQQQ.....	.TQKATAADK	G.....K	VSQNYPIVQN
98BWM037_d	ALDKIEEEQN	KSQQ.....	..K.TQQADG	KV.....S	..QNYPIVQN
99BW3932_1	ALDKIEEEQN	KCQQK.....	.TQQEKMAEG	K	VSQNYPIVQN
99BW4642_4	ALDKIEEEQN	KCQS.....	KTQQAEAAAAG	K.....	VSQNYPIVQN
99BW4745_8	ALDRIEEEQN	KSQQ	KAQQAQAAEG	K.....	VSQNYPVVQN
99BW4754_7	ALDKIEEEQN	KSQQK.....	.TTQAEAADK	K.....	VSQNYPIVQN
99BWMC16_8	ALDKIEEEQN	KCQQK.....	.TQQAEAADK	.G.....K	VSQNYPIVQN
A2_CD_97CD	ALDKIEEEQN	KCKQK.....	.TQQAAADTG	SSSSQNYRGS	SSQNYPIVQN
A2_CY_94CY	ALDKIEEEQN	..KQK.....	.TQHAAADTG	NS.....S	S.QNYPIVQN
A2D___97KR	ALDKLEEEQN	KHKQK.....	.TQPAAADTG	SS.....	GSQNYPIVQN
A2G_CD_97C	ALDKIEEEQN	TCKQR.....	.TQHAAADTG	SSRSQDYRGS	SSQNYPIVQN
A_BY_97BL0	ALDKIEEIQN	XSKQ.....	KTQQAAATGTG	S.....SS.K	VSQNYPIXQN
A_KE_Q23_A	ALDKIEEIKN	KSKQ.....	KTQQAAADTG	N.....SS.N	VSQNYPIVQN
A_SE_SE659	ALDKIEEMQK	KSKQ.....	KTSQATADTG	S.....SS.K	VSQNYPIVQN
A_SE_SE725	ALDKIEEIQK	KSKQ.....	KAQQAAADTG	N.....SS.K	VSQNYPIVQN
A_SE_SE753	ALDKIEEIQN	KSEQ.....	KTQQAAADTG	N.....S..K	VSQNYPIVQN
A_SE_SE853	ALDKLEEIQK	KSKQ.....	KTQQAVADTG	S.....SS.K	VSHNYPVVQN
A_SE_SE889	ALDKIEEIQN	KNKQ.....	KTQQAAADTG	N.....SS.N	VSRNYPIVQN
A_SE_UGSE8	ALDKIEEIQN	KSKQ.....	KKEQAAADTG	N.....SS.K	VSQNYPIVQN
A_UG_92UG0	ALDKIEEIQK	KSKQ.....	KTQQAAADTG	S.....SS.K	VSQNYPIVQN
A_UG_U455_	ALNKIEEMQN	KNKQR.....	.TQQAAANTG	SS.....	..QNYPIVQN
AC_IN_2130	ALDKIEEEQE	KSQQK.....	.TQQAKEADG	K.....	VSQNYPIVQN
AC_RW_92RW	ALDKIEEEQN	KSQQK.....	.TQQAEAADK	G.....K	VSQNYPIVQN
AC_SE_SE94	ALDKIEEIKN	KSKQ.....	KAQQAAADTG	N.....SG.K	VSQNYPIVQN
ACD_SE_SE8	ALDKIEEIQK	KSKQ.....	KAQQAAADTG	N.....SS.N	VSQNYPIVQN
ACG_BE_VI1	ALDKLEEVQN	KSKQR.....	.TQQAAAATG	S	GSQNFPIVQN
AD_SE_SE69	ALDKIEEEQT	KSK.....K	KAQQATADTK	N.....SS.Q	VSQNYPIVQN
AD_SE_SE71	ALDKIEEMQN	RSKQ.....	KTQQAAADTG	N.....NS.K	VSQNYPIVQN
ADHK_NO_97	ALDKIEEIQN	KSKQ.....	KTQQAAAATG	N.....GS.N	ISQNYPIVQN
ADK_CD_MAL	ALDKIEEIQN	KSRQKTQQAA	AAQQAAAATK	N.....SS.S	VSQNYPIVQN
AG_BE_VI11	AVDKIEEMQN	KSKQ.....	KTQQAAAADT	G.....S	SSQNYPIVQN
AG_NG_92NG	ALEEVEKIQK	NSQQE.....	.TQKAAMGKG	N.....SS.Q	VSQNYPIVQN
AGHU_GA_VI	ALDKIEEIQN	KSKQ.....	KTQQAAAATG	N.....SS.K	VSQNYPIVQN
AGU_CD_Z32	ALDEVEKAQK	SSQQK.....	.TQKATMAEE	N.....SSNQ	VSKNFPPIVQN
AJ_BW_BW21	ALDKIEEIQN	KSKQ.....	KAQQAAAATG	N.....SS.S	VSQNYPIVQN
B_AU_VH_AF	ALEKIEEEQN	KSKKKAQQ.A	AADAAAADAG	N.....SS.K	VSQNYPIVQN
B_CN_RL42_	ALEKIEEEQN	KSKK.....	KAQQTAAAGTG	NNS.....Q	ASQNYPIVQN
B_DE_D31_U	ALDKIEEEQN	KSKK.....	KAQPAARDAG	N.....NS.Q	VSQNYPIVQN
B_DE_HAN_U	ALDKVEEEQN	NSKK.....	KAQQEAADAG	N.....RN.Q	VSQNYPIVQN
B_FR_HXB2_	ALDKIEEEQN	KSKK.....	KAQQAAADTG	H.....SN.Q	VSQNYPIVQN
B_GA_OYI_	ALDKIEEEQN	KSKK.....	KAQQTAAADTG	NSS.....Q	VSQNYPIVQN
B_GB_CAM1_	ALEKIEEEQN	KSKK.....	KAQQAAAGTG	N.....SS.Q	VSQNYPIVQN
B_GB_GB8_A	ALDKIEEEQS	KSKK.....	KAQQAAADKG	NSNS...N.Q	VSQNYPIVQN
B_GB_MANC_	ALDKIEEEQN	KSKK.....	KAQPAAADTG	N.....SS.Q	VSHNYPPIVQN
B_KR_WK_AF	ALEKIEEEQS	KSKK.....	KAQQATADTG	S.....SS.Q	VSQNYPIVQN
B_NL_3202A	ALDKIEEEQN	KSKK.....	KAQQAAADTG	N.....SS.Q	VSQNYPIVQN
B_TW_TWCYS	ALEKIEEEQN	KSKK.....	KAQQAAATGTG	NNS.....P	VSQNYPIVQN
B_US_BC_L0	ALEKIEEEQN	NSKK.....	KAQQAAADTG	N.....SS.Q	VSRNYPIVQN
B_US_DH123	ALDKVEEEQN	KSKK.....	KAQQAAADTG	N.....SS.Q	VSQNYPIVQN
B_US_JRCSE	ALEKIEEEQT	KSMK.....	KAQQAAADTG	N.....SS.Q	VSQNYPIVQN
B_US_MNCG_	ALEKIEEEQN	KSKK.....	KAQQAAADTG	NRG.N.SS.Q	VSQNYPIVQN
B_US_P896_	ALDKIEEEQN	KSKK.....	KAQQAAADTG	N.....SS.Q	VSQNYPIVQN
B_US_RF_M1	ALDKIEEEQN	KSKK.....	KAQQAAADTG	N.....GS.Q	VSQNYPIVQN

B_US_SF2_K	ALEKIEEEQN	KSKK.....	KAQQAAAAAG	TG..N.SS.Q	VSQNYPIVQN
B_US_WEAU1	ALDKIEEEQN	KSKK.....	KAQQAAADTE	N.....RN.Q	VSQNYPIVQN
B_US_WR27_	ALEKIEEEQN	KSKK.....	KAQQAEADTG	N.....NS.K	VSQNYPIVQN
B_US_YU2_M	ALEKIEEEQN	KSKK.....	KAQQAAADTG	N.....SS.Q	VSQNYPIVQN
BF1_BR_93B	ALEKIEEEQN	.SKK.....	KAQQAAANTG	NNS.....Q	VSQNYPIVQN
C_BR_92BR0	ALDKIKEEQN	KSQOK.....	.TQQAEAADK	G.....K	VSQNYPIVQN
C_BW_96BW0	ALDKIEEEQN	KCQOK.....	.IQQAEEAADK	G.....K	VSQNYPIVQN
C_BW_96BW1	AVDKIEEEQN	KSQQ.....	KTQQAEAADK	RVS.....	..QNYPIVQN
C_BW_96BW1	ALDKIEEEQN	KCQOK.....	.EQQAEEAAK	G.....K	VSQNYPIVQN
C_BW_96BW1	ALDKIEEEQN	KSQE.....	KTQQAEAAAG	K.....	ISQNYPIVQN
C_ET_ETH22	ALDKIEEEQN	ESQOK.....	.TQQAGAADR	G.....K	DSQNYPIVQN
C_IN_93IN1	ALDKIEEEQN	KIQOK.....	.TQQAKEADG	K.....	VSQNYPIVQN
C_IN_93IN9	ALDKIEEEQN	KNQOK.....	.TQQAKEADE	K.....	VSQNYPIVQN
C_IN_93IN9	ALDKIEEEQN	KSQOK.....	.TQQAKEADG	K.....	VSQNYPIVQN
C_IN_94IN1	ALDKIEEEQN	KSQOK.....	.TQQAKEADS		..QNYPIVQN
C_IN_95IN2	ALDKIEEEQN	KIKOK.....	.TQQAKEDDG	K.....	VSQNYPIVQN
CRF01_AE_C	ALDKIEEVQN	KNKQ.....	KTQQAAAAGTG	S.....NS.K	VSQNYPIVQN
CRF01_AE_C	ALDKIEEVQN	KSKQ.....	KAQQAAAAGTG	S.....SS.K	VSQNYPIVQN
CRF01_AE_C	ALDKIEEIQN	KNKQ.....	KTQQAAADTG	S.....SS.K	VSQNYPIVQN
CRF01_AE_T	ALDKIEEVQK	KSQQ.....	KTQQAAAAGTG	S.....SS.K	VSQNYPIVQN
CRF01_AE_T	ALDKIEEVQN	KSQQ.....	KTQQAAAAGTG	S.....SS.K	VSQNYPIVQN
CRF01_AE_T	ALDKIEEVQK	KSQQ.....	KKQQAADTG	S.....SS.K	VSQNYPIVQN
CRF01_AE_T	ALDKIEEVQN	KSQQ.....	KTQQAAAAGTG	S.....SS.T	VSHNYPIVQN
CRF01_AE_T	ALDKIEEVQN	KSQR.....	KTQQAAAAGTG	S.....SS.K	VSQNYPIVQN
CRF01_AE_T	ALDKIEEAQK	KSQQ.....	KTQQAAAAGTG	S.....SS.K	VSQNYPIVQN
CRF02_AG_F	ALEKIEEVQN	KSKQ.....	KTQQAAAATG	S	SSQNYPIVQN
CRF02_AG_F	ALDKIEEVQN	KSKQ.....	KTQQAEAAATG	S	SSQNYPIVQN
CRF02_AG_G	ALDKLEEIQN	KSKQK.....	.TQQAAANS.	Q	VSQNYPIVQN
CRF02_AG_N	ALDKMEEIQN	KSKQ.....	KVQQTAAATG	S	SSQNYPIVQN
CRF02_AG_S	ALDKVEEVQK	KSKQ.....	KTQQAAAAAG	S	SSQNYPIVQN
CRF02_AG_S	ALDKLEEIQE	KSKQ.....	KTQQAAAATG	S	SSQNFPIVQN
CRF03_AB_R	ALDKIEEIQN	KSKQ.....	KTQQAATGTG	S.....SS.K	VSQNYPIVQN
CRF03_AB_R	ALDKIEEIQN	KSKQ.....	KTQQAATGTG	S.....SS.K	VSQNYPIVQN
CRF04_cpx_	ALDKIEEIQS	KSKQ.....	KTQQAAAAAG	G.....SS.N	VSQNYPIVQN
CRF04_cpx_	ALDKVEEVQK	KSKQ.....	KTQQAAADTG	G.....SS.K	VSQNYPIVQN
CRF04_cpx_	ALDKVEEQN	KSKQ.....	KTQQAAAGN.	SS.N	VSQNYPIAQN
CRF05_DF_B	ALEKIEEEQN	KSKR.....K	KAQQAEAGAG	N.....SS.Q	ASQNYPIVQN
CRF05_DF_B	ALEKIEEEQN	KSKS.....K	RAQQAEAGTK	N.....SG.P	VSQNFPIVQN
CRF06_cpx_	AIDKIEEIQK	KSKQK.....	.AHQAAATG.	N.....SS.N	LSQNYPIVQN
CRF06_cpx_	ALDKIEEIQN	KSKQK.....	.AQQAAAATG	N.....SS.N	LSQNYPIVQN
CRF06_cpx_	ALHKIEEIQN	KSKQK.....	.VQQAAAATG	N.....SS.Q	LSQNYPIVQN
CRF06_cpx_	AVDKVEEIQN	KSKQK.....	.TQQTTAATG	N.....SS.N	LSQNYPIVQN
CRF11_cpx_	ALDKIEEIQN	KSKQ.....	KKQQAADTG	N.....SS.N	VSQNYPIVQN
CRF11_cpx_	ALDKVEEIQN	KSK.....	.KQQAADSG	N.....SN.K	VSQNYPIVQN
D_CD_84ZR0	ALDKIEEEQN	NSKK.....	RAQQAAAGTG	N.....SS.Q	VSQNYPIVQN
D_CD_ELI_K	ALEKMEEEQN	KSK.....K	KAQQAAADTG	N.....NS.Q	VSQNYPIVQN
D_CD_NDK_M	AVEKMEEEQN	KSK.....K	KTQQAAADS.	S.Q	VSQNYPIVQN
D_UG_94UG1	ALDKIEEEQA	KSK.....K	KAQQATADTR	N.....SS.Q	VSQNYPIVQN
F1_BE_VI85	ALDKLEEEQN	KSQQ.....	.KTQQAAADK	G	VSQNYPIVQN
F1_BR_93BR	ALEKLEEEQN	KGRQ.....	.KTQQATAEK	G	VSQNYPIVQN
F1_FI_FIN9	ALDKLEEEQN	KSQOKT....	.QAAAAADK	G	VSQNYPIVQN
F1_FR_MP41	ALEKLEEEQN	KGQQ.....	.KTQQAAADK	G	VSQNYPIVQN
F2_CM_MP25	ALDKLQEEQD	KHQQ.....	.KTQQATADK	GVSK.....G	VSQNYPIVQN
F2KU_BE_VI	ALDKLEEEQK	TQKQQ.....	.KTQQEAAGK	G	VSQNYPIVQN
G_BE_DRCBL	ALEEVEKIQK	KSQQ.....	KEN	S.....SS.Q	VSQNYPIVQN
G_NG_92NG0	APEEVEKIQK	NSQQE.....	.IQQAANKNEG	N.....SN.P	VSQNYPIVQN
G_SE_SE616	ALEEVEKIQK	KSQEK.....	.IQQAAMDKG	N.....SN.Q	VSQNYPIVQN
H_BE_VI991	ALGKIEEIQN	KNKQR.....	.TQQAPAAAD	K.EK...DSK	ISQNYPIVQN
H_BE_VI997	ALGKIEEIQN	KRQOK.....	.TQQATANK.	..ER...DNK	VSQNYPIVQN
H_CF_90CF0	ALDKIEEIQN	KSQOK.....	.TQQAAADK.	..EK...DNK	VSQNYPIVQN
J_SE_SE702	ALDKIEEIQN	KNKQQ.....	.TQKAETDKK	DNS.....Q	VSQNYPIVQN

J_SE_SE788	ALEKIEEIQN	KNKQQ.....	.AQKAETDKK	DNS.....Q	VSQNYPIVQN
K_CD_EQTB1	ALDKLEEEQ.	NRTQQ.....	.KTQQGKADK	G	VSQNYPIVQN
K_CM_MP535	ALDKLEEEQ.	NKSQR.....	.KTQQEAADK	G	VSQNYPIVQN
N_CM_YBF30	ALDKIKEKQE	QHKPE.....P	KNPEAGAAAA	TDS.....N	ISRNYPLVQT
O_CM_ANT70	AIQKLKEVMG	SRKS.....	.ADAAKEDTS	A.....R..Q	AGQNYPIVSN
O_CM_MVP51	AIQKLKEVMA	SRKS.....	.AEAAKEETS	P.....R..Q	TSQNYPIVTN
O_SN_99SE	AIQKLKEVMG	SRKS.....	.AGTAKEDTS	A.....R..Q	TGQNYPIVTN
O_SN_99SE	AIQKLKEVMG	SRKS.....	.AGAAKEDTS	A.....R..Q	TGQNYPVVAN
U_CD___83C	ALDKIEEVQN	KGKQK.....	.AQQAADK.	..GS...NSQ	VSQNYPIVQN

151

200

00BW0762_1	IQGQMVHQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW0768_2	LQGQMVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW0874_2	LQGQMVHNPI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW1471_2	AQGQMVHQSI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPSDLNTM
00BW1616_2	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW1686_8	LQGQMVHQAI	SPRTLNAWVK	VIEEKGFNPE	VIPMFTALSE	GATPQDLHTM
00BW1759_3	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPSDLNTM
00BW1773_2	LQGQMVHQPI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW1783_5	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPGDLNTM
00BW1795_6	LQGQMVHQAI	SPRTLNAWVK	VIEDKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW1811_3	AQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALPE	GATPQDLNTM
00BW1859_5	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW1880_2	LQGQMVHQPL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW1921_1	LQGQMVHQSI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW2036_1	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	IIPMFTALSE	GATPQDLNTM
00BW2063_6	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW2087_2	AQGQMVHQTI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW2127_2	LQGQMVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW2128_3	LQGQMVHQPL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDSNTM
00BW2276_7	IQGQMVHQPL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW3819_3	LQGQMVHQNI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW3842_8	LQGQMVHQAI	SPRTLNAWVK	VIEEKGFNPE	VIPMFTALSE	GATPQDLNTM
00BW3871_3	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GAAPQDLNTM
00BW3876_9	LQGQMVHQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW3886_8	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
00BW3891_6	LQGQMVHQSL	SPRTLNAWVK	VIEEKAFSPE	IIPMFTALSE	GATPQDLNTM
00BW3970_2	LQGQMVHQPI	APRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNSM
00BW5031_1	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFNPE	IIPMFTALSE	GATPRDLNTM
96BW01B21	LQGQMVHQPL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
96BW0407	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
96BW0502	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
96BW06_J4	LQGQMVHQPI	SPRTLNAWVK	VIEEKGFNPE	VIPMFTALSE	GATPQDLNTM
96BW11_06	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
96BW1210	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	IIPMFTALSE	GATPQDLNTM
96BW15B03	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
96BW16_26	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
96BW17A09	AQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
96BWM01_5	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQILNTM
96BWM03_2	LQGQMVHQAI	SPRTLNAWVK	VIEEKGFNPE	VIPMFTALSE	GATPQDLNTM
98BWMC12_2	VQGQMVHQPI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLTTM
98BWMC13_4	LQGQMVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
98BWMC14_a	LQGQMVHQAI	SPRTLNAWVK	VIEEKGFNPE	VIPMFTALSE	GATPQDLNTM
98BWM014_1	IQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
98BWM018_d	LQGQVQHAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
98BWM036_a	LQGQMVHQAP	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
98BWM037_d	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
99BW3932_1	LQGQMVHQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
99BW4642_4	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
99BW4745_8	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
99BW4754_7	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM

99BWMC16_8	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFTALSE	GATPQDLNTM
A2_CD_97CD	AQGQMVHQAV	SPRTLNAWVK	VVEEKAFSPE	VIPMFTALSE	GATPQDLNTM
A2_CY_94CY	AQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFTALSE	GATPQDLNTM
A2D___97KR	AQGQMTYQNL	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
A2G_CD_97C	AQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
A_BY_97BL0	AQGQMT HQSM	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
A_KE_Q23_A	AQGQMIHQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
A_SE_SE659	AQGQMIHQSL	SPRTLNAWVK	VIEEKGFNPE	VIPMFSALSE	GATPQDLNMM
A_SE_SE725	AQGQMVHQSL	SPRTLNAWVK	VIEEKAFSPE	VIPVFSALSE	GATPQDLNMM
A_SE_SE753	AQGQMVHQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
A_SE_SE853	AQGQMIHQNL	SPRTLNAWVK	VIEEKGFNPE	VIPMFSALSE	GATPQDLNTM
A_SE_SE889	AQGQMVHQSL	SPRTLNAWVK	VIEEKGFSPPE	VIPMFSALSE	GATPGDLNMM
A_SE_UGSE8	AQGQMVHQVM	SPRTLNAWVK	VIEERAFSPE	VIPMFSALSE	GATPHDLNMM
A_UG_92UG0	AQGQMIHQSL	SPRTLNAWVK	VIEEKALSPE	VIPMFSALSE	GATPQDLNMM
A_UG_U455_	AQGQPVHQAL	SPRTLNAWVK	VVEDKAFSPE	VIPMFSALSE	GATPQDLNMM
AC_IN_2130	LQGQMVHQAL	SPRTLNAWVK	VIEEKAFSPE	IIPMFTALSE	GATPQDLNTM
AC_RW_92RW	AQGQMVHQAI	SPRTLNAWVK	VIEEKAFSQE	VIPMFTALSE	GATPQDLNTM
AC_SE_SE94	PQGQMVHQPI	SPRTLNAWVK	VIDEKAFSPE	VIPMFTALSE	GATPQDLNTM
ACD_SE_SE8	AQGQMVHQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
ACG_BE_VI1	AQE QMVHQSM	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
AD_SE_SE69	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GAAPQDLNTM
AD_SE_SE71	AQGQMIHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
ADHK_NO_97	AQGQMVHQAI	SPRTLNAWVK	AIEEKAFSPE	VIPMFSALSD	GATPQDLNMM
ADK_CD_MAL	AQGQMIHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
AG_BE_VI11	AQGQMVHQSM	SPRTLNAWVK	VIEEKGFSPPE	VIPMFSALSE	GATPQDLNMM
AG_NG_92NG	AQGQVVHQPI	SPRTLNAWVK	VIEEKNFSPPE	VIPMFTALSE	GATPQDLNTM
AGHU_GA_VI	AQGQMVHQAM	SPRTLNAWVK	VIEEKGFSPPE	VIPMFSALSE	GATPQDLNMM
AGU_CD_Z32	AQGQMVHQAL	TPRTLNAWVK	VVEEKNFSPE	VIPMFSALSE	GATPQDLNTM
AJ_BW_BW21	AQGQMVHQAV	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
B_AU_VH_AF	LQGQMVHQPI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_CN_RL42_	LQGQMVHQPI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_DE_D31_U	LQGQMVHQPI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_DE_HAN_U	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_FR_HXB2_	IQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_GA_OYI__	LQGQMVHQPI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALAE	GATPQDLNTM
B_GB_CAM1_	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_GB_GB8_A	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALAE	GATPQDLNTM
B_GB_MANC_	LQGQMVHQSI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GAIPQDLNTM
B_KR_WK_AF	LQGQMVHQPI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_NL_3202A	LQGQMVHQAL	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_TW_TWCYS	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_BC_L0	IQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_DH123	IQGQMVHQAL	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_JRCSF	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_MNCG_	IEGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_P896_	IQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_RF_M1	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_SF2_K	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_WEAU1	LQGQMVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
B_US_WR27_	LQGQMVHQAL	SPRTLFAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLYTM
B_US_YU2_M	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
BF1_BR_93B	LQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
C_BR_92BR0	LQGQMVHQPI	SARTLNAWVK	VVEEKAFSPE	VIPMFTALSE	GATPQDLNTM
C_BW_96BW0	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
C_BW_96BW1	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
C_BW_96BW1	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	IIPMFTALSE	GATPQDLNTM
C_BW_96BW1	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
C_ET_ETH22	MQGQMVHQPI	SARTLNAWVK	VVEEKAFSPE	VIPMFTALSE	GATPQDLNTM
C_IN_93IN1	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
C_IN_93IN9	LQGQMVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
C_IN_93IN9	LQGQMVHQPI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM

C_IN_94IN1	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
C_IN_95IN2	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
CRF01_AE_C	AQGQMVHQPL	SPRTLNAWVK	VVEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF01_AE_C	AQGQMVHQAL	SPRTLNAWVK	VVEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF01_AE_C	AQGQMIHQAL	SPRTLNAWVK	VVEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF01_AE_T	AQGQMVHQPV	SPRTLNAWVK	VVEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF01_AE_T	AQGQMVHQPL	SPRTLNAWVK	VVEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF01_AE_T	AQGQMVHQPL	SPRTLNAWVK	VIEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF01_AE_T	AQGQMVHQPI	SPRTLNAWVK	VIEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF01_AE_T	AQGQMAHQPL	SPRTLNAWVK	VVEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF01_AE_T	AQGQMVHQPV	SPRTLNAWVK	VVEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF02_AG_F	AQGQMT HQPM	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNMM
CRF02_AG_F	AQGQMT HQPM	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNMM
CRF02_AG_G	AQGQMVHQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFTALSE	GATPQDLNMM
CRF02_AG_N	AKGQMT HQSM	SPRTLNAWVK	VIEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF02_AG_S	AQGQMT HQSI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF02_AG_S	AQGQMT HQSI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF03_AB_R	AQGQMT HQSM	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF03_AB_R	AQGQMT HQSM	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF04_cpx_	AQGQMT HQSI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF04_cpx_	AQGQMT HQAM	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF04_cpx_	AQGQMT HQSI	SPRTLNAWVK	VVEEKGFNPE	VIPMFSALSE	GATPQDLNMM
CRF05_DF_B	IQGQMVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
CRF05_DF_B	LQGQMVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
CRF06_cpx_	AQGQMT HQAM	SPRTLNAWVK	VIEDKAFSPE	VIPMFTALSE	GATPQDLNMM
CRF06_cpx_	AQGQMT HQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF06_cpx_	AQGQMT HQAI	SPRTLNAWVK	AIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF06_cpx_	AQGQMT HQAI	SPRTLNAWVK	AIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF11_cpx_	AQGQMT HQPV	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNMM
CRF11_cpx_	AQGQMT HQAM	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNMM
D_CD_84ZR0	LQGQMVHQAL	SPRTLNAWVK	VIEEKAFNPE	VIPMFSALSE	GATPQDLNTM
D_CD_ELI_K	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
D_CD_NDK_M	LQGQMVHQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
D_UG_94UG1	LQGQMVHHPL	SPRTLNAWVK	VIEEKAFNPE	VIPMFSALSE	GATPQDLNTM
F1_BE_VI85	LQGQMT HQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPTDLNTM
F1_BR_93BR	LQGQMT HQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
F1_FI_FIN9	LQGQMT HQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
F1_FR_MP41	LQGQMT HQPI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
F2_CM_MP25	LQGQMT HQSL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
F2KU_BE_VI	LQGQMT HQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
G_BE_DRCBL	AQGQMT HQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFTALSE	GATPQDLNTM
G_NG_92NG0	AQGQMT HQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNTM
G_SE_SE616	AQGQMT HQAI	TPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNLM
H_BE_VI991	AQGQMT HQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNAM
H_BE_VI997	AQGQMT HQPI	SXRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNAM
H_CF_90CF0	AQGQMT HQAI	SPRTLNAWVK	VVEEKAFSPE	VIPMFSALSE	GATPQDLNAM
J_SE_SE702	LQGQPVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
J_SE_SE788	LQGQPVHQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
K_CD_EQTB1	LQGQMT HQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM
K_CM_MP535	LQGQMT HQAL	SPRTLNAWVK	VIEEKAFSPE	VIPMFTALSE	GATPQDLNTM
N_CM_YBF30	AQGQMT HQPL	TPRTLNAWVK	VIEEKAFSPE	VIPMFMALSE	GATPSDLNTM
O_CM_ANT70	AQGQMT HQAI	SPRTLNAWVK	AVEEKAFNPE	IIPMFMALSE	GAISYDINTM
O_CM_MVP51	AQGQMT HQAI	SPRTLNAWVK	AVEEKAFNPE	IIPMFMALSE	GAVPYDINTM
O_SN_99SE_	AQGQMT HQSL	SPRTLNAWVK	AVEEKAFNPE	IIPMFMALSE	GAIPYDTNTM
O_SN_99SE_	AQGQMT HQSL	SPRTLNAWVK	AVEEKAFNPE	IIPMFMALSE	GAIPYDTNTM
U_CD_83C	MQGQMT HQAI	SPRTLNAWVK	VIEEKAFSPE	VIPMFSALSE	GATPQDLNTM

201

250

00BW0762_1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
00BW0768_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PAHAGPVAPG	QMREPRGSDI
00BW0874_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVHAGPVAPG	QMREPRGSDI

00BW1471_2	LNTVGGHQ.A	AMQMLKDTIN	EEVAEWDR LH	PVQAGPIAPG	QMRDPRGSDI
00BW1616_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR MH	PVQAGPVAPG	QMRDPRGSDI
00BW1686_8	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
00BW1759_3	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
00BW1773_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
00BW1783_5	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVQAGPVAPG	QIREPRGSDI
00BW1795_6	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
00BW1811_3	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
00BW1859_5	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVHAGPIAPG	QMREPRGSDI
00BW1880_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMRDPRGSDI
00BW1921_1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QLREPRGSDI
00BW2036_1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPNPAG	QMREPRGSDI
00BW2063_6	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPAAPG	QMREPRGSDI
00BW2087_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
00BW2127_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
00BW2128_3	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVHAGPIAPG	QMREPRGSDI
00BW2276_7	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
00BW3819_3	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
00BW3842_8	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVQAGPVAPG	QIREPRGSDI
00BW3871_3	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
00BW3876_9	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR IH	PVHAGPVAPG	QMRDPRGSDI
00BW3886_8	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
00BW3891_6	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
00BW3970_2	LNAVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QLREPRGSDI
00BW5031_1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
96BW01B21	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR TH	PVHAGPVAPG	QLREPRGSDI
96BW0407	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVHAGPIAPG	QMREPRGSDI
96BW0502	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVQAGPVAPG	QMRDPRGSDI
96BW06_J4	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVQAGPVAPG	QMRDPRGSDI
96BW11_06	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
96BW1210	LNTVGGHQ.A	AMQMLKDTIN	EEAAGWDR LH	PVHAGPVAPG	QMREPRGSDI
96BW15B03	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
96BW16_26	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVHAGPIAPG	QMREPRGSDI
96BW17A09	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
96BWM01_5	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
96BWM03_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR MH	PVHAGPVAPG	QMREPRGSDI
98BWMC12_2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
98BWMC13_4	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVQAGPVAPG	QIREPRGSDI
98BWMC14_a	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMRDPRGSDI
98BWM014_1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR TH	PVHAGPVAPG	QMREPRGSDI
98BWM018_d	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
98BWM036_a	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
98BWM037_d	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVQAGPIAPG	QIREPRGSDI
99BW3932_1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR TH	PVHAGPVAPG	QMREPRGSDI
99BW4642_4	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
99BW4745_8	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR IH	PVHAGPVAPG	QMREPRGSDI
99BW4754_7	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVHAGPIAPG	QIREPRGSDI
99BWMC16_8	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR TH	PVHAGPIAPG	QMREPRGSDI
A2_CD_97CD	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVQAGPIPPG	QMREPRGSDI
A2_CY_94CY	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVHAGPIPPG	QMREPRGSDI
A2D_97KR	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGSDI
A2G_CD_97C	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR VH	PVHAGPIPPG	QMREPRGS..
A_BY_97BL0	LNIVGGHQ.A	AMQMLKDTIN	EEAAXXDR LH	PAQAGPFPPG	QMREPRGSDI
A_KE_Q23_A	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGSDI
A_SE_SE659	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGS..
A_SE_SE725	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PAHAGPVAPG	QMREPRGS..
A_SE_SE753	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGSDI
A_SE_SE853	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGS..
A_SE_SE889	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGS..
A_SE_UGSE8	LNIVGGHQ.A	AMEMLKDTIN	EEAAEWDR TH	PIHAGPVAPG	QMREPRGSDI
A_UG_92UG0	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI

A_UG_U455_	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPIPPG	QMREPRGSDI
AC_IN_2130	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PAQAGPIAPG	QMREPRGSDI
AC_RW_92RW	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVQAGPVAPG	QIREPRGSDI
AC_SE_SE94	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PAQAGPVAPG	QMREPRGS..
ACD_SE_SE8	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGLIAPG	QMREPRGSDI
ACG_BE_VI1	LNIVGGHQ.A	AMQMLKDTIN	EEAADWDRTH	PVHAGPNPPG	QMREPRGSDI
AD_SE_SE69	LSTVGGHQ.A	AMQILKETIN	EEAADWDRHLH	PVHAGPNAPG	QMREPRGSDI
AD_SE_SE71	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPIPPG	QMREPRGS..
ADHK_NO_97	LNIVGGHQ.A	AMQMLKDTIN	EEAADWDRHLH	PVHAGPIPPG	QMREPRGSDI
ADK_CD_MAL	LNIVGGHQ.A	AMQMLKDTIN	EEAADWDRVH	PVHAGPIPPG	QMREPRGSDI
AG_BE_VI11	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
AG_NG_92NG	LNTVGGHQ.A	AMQMLKDSIN	EEAAEWDRHLH	PQQAGPIPPG	QIREPRGSDI
AGHU_GA_VI	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVQAGPIPPG	QIREPRGSDI
AGU_CD_Z32	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PPQAGPIPPG	QIREPRGSDI
AJ_BW_BW21	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPIPPG	QMREPRGSDI
B_AU_VH_AF	LNTVGGHQAA	AMQMLKETIN	EEAADWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_CN_RL42_	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPVAPG	QMREPRGSDI
B_DE_D31_U	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_DE_HAN_U	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_FR_HXB2_	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPIAPG	QMREPRGSDI
B_GA_OYI_	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_GB_CAM1_	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_GB_GB8_A	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPVAPG	QMREPRGSDI
B_GB_MANC_	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPVAPG	QMREPRGSDI
B_KR_WK_AF	LNTVGGHQ.A	AMQMLKETIN	EEAADWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_NL_3202A	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_TW_TWCYS	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_US_BC_L0	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVQAGPVAPG	QMREPRGSDI
B_US_DH123	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_US_JRCSF	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_US_MNCG_	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPITPG	QMREPRGSDI
B_US_P896_	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVQAGPVAPG	QMREPRGSDI
B_US_RF_M1	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_US_SF2_K	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPIAPG	QMREPRGSDI
B_US_WEAU1	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
B_US_WR27_	LNTVGGHQ.A	AMQMLKETIN	DEAAEWDRHLH	PVQAGPVAPG	QMREPRGSDI
B_US_YU2_M	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
BF1_BR_93B	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
C_BR_92BR0	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPVAPG	QMREPRGSDI
C_BW_96BW0	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
C_BW_96BW1	LTTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPVAPG	QMREPRGSDI
C_BW_96BW1	LNTVGGHQ.A	AMQMLKDTIN	EEAAGWDRHLH	PVHAGPVAPG	QMREPRGSDI
C_BW_96BW1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
C_ET_ETH22	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPVAPG	QMRDPRGSDI
C_IN_93IN1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PIHAGPIAPG	QMREPRGSDI
C_IN_93IN9	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PIHAGPIAPG	QMREPRGSDI
C_IN_93IN9	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPIAPG	QIREPRGSDI
C_IN_94IN1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPIAPG	QMREPRGSDI
C_IN_95IN2	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVPAGPIAPG	QLREPRGSDI
CRF01_AE_C	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRHLH	PVHAGPIPPG	QMREPRGSDI
CRF01_AE_C	LNIVGGHQ.A	AMQILKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGADI
CRF01_AE_C	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF01_AE_T	LNIVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF01_AE_T	LNIVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF01_AE_T	LNIVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF01_AE_T	LNIVGGHQ.A	AMQMLKETIN	EEPAEWDRVH	PVHAGPIPPG	QIREPRGSDI
CRF01_AE_T	LNIVGGHQ.A	AMQMLKETIN	EEPAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF01_AE_T	LNIVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF01_AE_T	LNIVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF02_AG_F	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF02_AG_F	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF02_AG_G	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRTH	PIHAGPNPPG	QMREPRGSDI

CRF02_AG_N	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF02_AG_S	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF02_AG_S	LNIVGGHQ.A	AMQMLKDTIN	EEAADWDRTH	PVHAGPIPPG	QMREPRGSDI
CRF03_AB_R	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PAQAGPFPPG	QMREPRGSDI
CRF03_AB_R	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PAQAGPFPPG	QMREPRGSDI
CRF04_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEAADWDRTH	PVHAGPIPPG	QMREPRGSDI
CRF04_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEASEWDRAH	PVHAGPIPPG	QMREPRGSDI
CRF04_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PAHAGPNPAG	QMREPRGSDI
CRF05_DF_B	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDR LH	PVQAGPVAPG	QMRDPRGSDI
CRF05_DF_B	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDRVH	PAQAGPIAPG	QIREPRGSDI
CRF06_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QIREPRGSDI
CRF06_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
CRF06_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRMH	PVQAGPMPPG	QMREPRGSDI
CRF06_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PAQAGPIPPG	QIRDPRGSDI
CRF11_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVQAGPIAPG	QMRDPRGSDI
CRF11_cpx_	LNIVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPAPPG	QMREPRGSDI
D_CD_84ZR0	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDR LH	PVQAGPVAPG	QMREPRGSDI
D_CD_ELI_K	LNTVGGHQ.A	AMQMLKETIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
D_CD_NDK_M	LNTVGGHQ.A	AMQMLKETIN	DEAAEWDR LH	PVHAGPVAPG	QMREPRGSDI
D_UG_94UG1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPVAPG	QLREPRGSDI
F1_BE_VI85	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPAPPG	QMREPRGSDI
F1_BR_93BR	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PTQAGPIPPG	QIREPRGSDI
F1_FI_FIN9	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGSDI
F1_FR_MP41	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PAHAGPILPG	QMREPRGSDI
F2_CM_MP25	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGSDI
F2KU_BE_VI	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIAPG	QMREPRGSDI
G_BE_DRCBL	LNTVGGHQ.A	AMQMLKETIN	DEAAEWDR LH	PQQAGPIAPG	QIRDPTGSDI
G_NG_92NG0	LNTVGGHQ.A	AMQMLKDTIN	DEAAEWDR LH	PQQAGPIPPG	QIREPSGSDI
G_SE_SE616	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR MH	PQQAGPFPPG	QIREPRGSDI
H_BE_VI991	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGSDI
H_BE_VI997	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR LH	PVHAGPIPPG	QMREPRGSDI
H_CF_90CF0	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIPPG	QMREPRGSDI
J_SE_SE702	LNTIGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPVAPG	QVREPRGSDI
J_SE_SE788	LNTIGGHQ.A	AMQMLKDTIN	EEAAEWDRVH	PVHAGPIAPG	QVREPRGSDI
K_CD_EQTB1	LNTVGGHQ.A	AMQMLKDTIN	EEAAEWDR MH	PVQAGPIPPG	QIREPRGSDI
K_CM_MP535	LNTVGGHQ.A	AMQMLKDTIN	DEAAEWDR LH	PVHAGPIPPG	QMREPRGSDI
N_CM_YBF30	LNTVGGHQ.A	AMQMLKEVIN	EEAADWDRTH	PVPVGPLPPG	QLRDPRGSDI
O_CM_ANT70	LNAIGGHQ.G	ALQVLKEVIN	EEAVEWDRTH	PPVVGPLPPG	QIREPTGSDI
O_CM_MVP51	LNAIGGHQ.G	ALQVLKEVIN	EEAAEWDRTH	PPAMGPLPPG	QIREPTGSDI
O_SN_99SE_	LNAIGGHQ.G	ALQVLKEVIN	EEAAEWDRTH	PPAAGPLPVG	QIREPTGSDI
O_SN_99SE_	LNAIGGHQ.G	ALQVLKEVIN	EEAAEWDRTH	PQAAGPLPPG	QIREPTGSDI
U_CD_83C	LNTVGGHQ.A	AMQMLKDTIN	EEAADWDR LH	PVHAGPIPPG	QMREPRGSDI

	251				300
00BW0762_1	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW0768_2	AGTTSNLQEQ	IAWMTA.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW0874_2	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1471_2	AGTTSTLQEQ	VAWMTS.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1616_2	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1686_8	AGTTSNLQEQ	VAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1759_3	AGTTSTLQEQ	IAWMTS.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1773_2	AGTTSTLQEQ	ITWMTS.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1783_5	AGTTSTLQEQ	ITWMTS.NPP	IPVGDIIYKRW	IVLGLNKIVR	MYSPPVSILDI
00BW1795_6	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1811_3	AGTTSTLQEQ	IAWMTN.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1859_5	AGTTSTLQEQ	IAWMTG.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW1880_2	AGTTSTLQEQ	ITWMTS.NPP	IPVGDIIYKRW	IVLGLNKIVR	MYSPPVSILDI
00BW1921_1	AGTTSTLQEQ	IAWMTN.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW2036_1	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW2063_6	AGTTSTLAEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW2087_2	AGTTSTLQEQ	IAWMTN.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW2127_2	AGTTTTLQEQ	INWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW2128_3	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW2276_7	AGTTSTLQEQ	IAWMTS.NPA	IPVGDIIYKRW	IILGLDKIVR	MYSPPVSILDI
00BW3819_3	AGTTSTLQEQ	VAWMTS.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW3842_8	AGTTSNLQEQ	ITWMTS.NPP	IPVGEIYKRW	IVLGLNKIVR	MYSPPVSILDI
00BW3871_3	AGTTSTLQEQ	IDWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW3876_9	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW3886_8	AGTTSTLQEQ	IAWMTN.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW3891_6	AGTTSNLQEQ	INWMTA.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
00BW3970_2	AGTTSTLQEQ	IAWMTN.NPP	VPVGDIIYKRW	IVLGLNKIVR	MYSPPVSILDI
00BW5031_1	AGTTSTLQEQ	IAWMTS.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW01B21	AGTTSNLQEQ	IAWMTA.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW0407	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW0502	AGATSTLQEQ	IAWMTS.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW06_J4	AGTTSTLQEQ	IGWMTN.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW11_06	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW1210	AGTTSNLQEQ	INWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW15B03	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW16_26	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
96BW17A09	AGTTSTLQEQ	IAWMTN.NPP	IPVGDIIYKRW	ITMGLNKIVR	MYSPPVSILDI
96BWM01_5	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
96BWM03_2	AGSTSTLQEQ	IAWMTS.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
98BWMC12_2	AGSTSNLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
98BWMC13_4	AGTTSTLQEQ	IAWMTR.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
98BWMC14_a	AGTTSTLQEQ	VGWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
98BWM014_1	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
98BWM018_d	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
98BWM036_a	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIIYKRW	IIMGLNKIVR	MYSPPVSILDI
98BWM037_d	AGTNSTLQEQ	IAWMTN.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
99BW3932_1	AGTTSTLQEQ	IQWMTS.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
99BW4642_4	AGTTSTLQEQ	VTWMTS.NPP	VPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
99BW4745_8	AGTTSTLQEQ	INWMTG.NPP	IPVGDIIYKRW	IIMGLNKIVR	MYSPPVSILDI
99BW4754_7	AGTTSTLQEQ	ITWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
99BWMC16_8	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
A2_CD_97CD	AGATSNLQEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
A2_CY_94CY	AGTTSTLQEQ	IGWMTS.DPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
A2D_97KR	AGTTSTLQEQ	IGWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
A2G_CD_97C					
A_BY_97BL0	AGTTSTLQEQ	IGXMTS.NPP	IPXGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
A_KE_Q23_A	AGTTSTPQEQ	IGWMTG.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI
A_SE_SE659					
A_SE_SE725					
A_SE_SE753	AGTTSTPQEQ	IGWMTG.NPP	IPVGDIIYKRW	IILGLNKIVR	MYSPPVSILDI

A_SE_SE853					
A_SE_SE889					
A_SE_UGSE8	AGTTSTLQEQ	IAWMTS.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
A_UG_92UG0	AGTTSTPQEQ	IAWMTG.NPP	IPVGDIYKRW	MILGLNKIVR	MYSPPVSILDI
A_UG_U455	AGTTSTVQEQ	IGWMTG.NPP	IPVGDIYRRW	IILGLNKIVR	MYSPPVSILDI
AC_IN_2130	AGTTSTLQEQ	IAWMTG.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
AC_RW_92RW	AGTTSTLQEQ	IAWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
AC_SE_SE94					
ACD_SE_SE8	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
ACG_BE_VI1	AGTTSTLQEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
AD_SE_SE69	AGTTSTLQEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
AD_SE_SE71					
ADHK_NO_97	AGTTSTLQEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
ADK_CD_MAL	AGTTSTLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
AG_BE_VI11	AGTTSTLQEQ	VGWMTS.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
AG_NG_92NG	AGTTSTLQEQ	ITWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
AGHU_GA_VI	AGTTSTTQEQ	IGWMTG.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
AGU_CD_Z32	AGTTSTLQEQ	IRWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
AJ_BW_BW21	AGTTSTLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
B_AU_VH_AF	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_CN_RL42	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTCILDI
B_DE_D31_U	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_DE_HAN_U	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_FR_HXB2	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_GA_OYI	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_GB_CAM1	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_GB_GB8_A	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
B_GB_MANC	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_KR_WK_AF	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPASILDI
B_NL_3202A	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_TW_TWCYS	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_US_BC_L0	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPSSILDI
B_US_DH123	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IIMGLNKIVR	MYSPTSILDI
B_US_JRCSF	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
B_US_MNCG	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPSSILDI
B_US_P896	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPSSILDI
B_US_RF_M1	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPISILDI
B_US_SF2_K	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_US_WEAU1	AGTTSTLQEQ	IAWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
B_US_WR27	AGXTSTLXXX	IGWMTN.XPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
B_US_YU2_M	AGTTSTLQEQ	IGWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPTSILDI
BF1_BR_93B	AGTTSTLQEQ	IQWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPTSILGI
C_BR_92BR0	AGTTSTLQEQ	ITWMTN.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_BW_96BW0	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_BW_96BW1	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_BW_96BW1	AGTTSNLQEQ	INWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_BW_96BW1	AGTTSTLQEQ	IAWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_ET_ETH22	AGTTSTLQEQ	IAWMTG.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_IN_93IN1	AGTTSSLQEQ	IAWMTG.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_IN_93IN9	AGTTSSLQEQ	IAWMTG.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_IN_93IN9	AGTTSTLQEQ	IAWMTG.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
C_IN_94IN1	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDT
C_IN_95IN2	AGTTSTLQEQ	IAWMTN.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF01_AE_C	XGTTSNLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF01_AE_C	AGTTSTLHEQ	IGWMTS.NPP	IPVGEIYKKW	IILGLNKIVR	MYSPPVSILDI
CRF01_AE_C	AGTTSTLQEQ	IGWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF01_AE_T	AGTTSTLQEQ	IGWMTN.NPP	IPVGDIYKRW	IILGLNKIVR	MYRPPVSILDI
CRF01_AE_T	AGTTSTLQEQ	IGWMTN.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF01_AE_T	AGTTSTLQEQ	IGWMTN.NPP	IPVGDIYKRW	IILGLNKIVR	MYQPVSILDI
CRF01_AE_T	AGTTTTLQEQ	IGWMTN.NPP	IPVGSIIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF01_AE_T	AGTTSTLQEQ	IGWMTN.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI

CRF01_AE_T	AGTTSTLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF02_AG_F	AGTTSTPQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF02_AG_F	AGTTSTLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF02_AG_G	AGTTSNLQEQ	IAWMTG.NPP	IPVGEIYKRW	IVLGLNKIVR	MYSPPVGILDI
CRF02_AG_N	AGTTSTLQEQ	IGWMTS.NPP	IPVGEIYKRW	IVLGLNKIVR	MYSPPVSILDI
CRF02_AG_S	AGTTSTLQEQ	IGWMTS.NPP	IPVGEIYKRW	IVLGLNKIVR	MYSPPVSILDI
CRF02_AG_S	AGTTSTLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF03_AB_R	AGSSSTLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF03_AB_R	AGTTSTLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF04_cpx_	AGTTSTLQEQ	IGWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	TYSPPISILDI
CRF04_cpx_	AGTTSTLQEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF04_cpx_	AGTTSTLQEQ	VGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF05_DF_B	AGTTSTLQEQ	IAWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF05_DF_B	AGTTSTLQEQ	ITWMTN.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF06_cpx_	AGTTSTLQEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVGILDI
CRF06_cpx_	AGTTSTLQEQ	IGWMTG.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF06_cpx_	AGTTSTLLEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF06_cpx_	AGTTSTLQEQ	IRWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
CRF11_cpx_	AGTTSTLQEQ	IGWMTG.NPP	IPVGDIYRRW	IILGLNKIVR	MYSPPVSILDI
CRF11_cpx_	AGTTGNLQEQ	IGWMTG.NPP	IPVGEIYRRW	IILGLNKIVR	MYSPPVSILDI
D_CD_84ZR0	AGTTSTLQEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
D_CD_ELI_K	AGTTSTLQEQ	IAWMTS.NPP	IPVGEIYKRW	IIVGLNKIVR	MYSPPVSILDI
D_CD_NDK_M	AGTTSTLQEQ	IAWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
D_UG_94UG1	AGTTSNLQEQ	IGWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
F1_BE_VI85	AGTTSTLQEQ	IQWMTG.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
F1_BR_93BR	AGTTSTLQEQ	IQWMTG.NPP	VPVGEMYKRW	IILGLNKIVR	MYSPPVGILDI
F1_FI_FIN9	AGTTSTLQEQ	IQWMTS.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
F1_FR_MP41	AGTTSTLQEQ	IQWMTS.NPP	VPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
F2_CM_MP25	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
F2KU_BE_VI	AGATSNLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
G_BE_DRCBL	AGATSTLQEQ	IRWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
G_NG_92NG0	AGTTSTLQEQ	IRWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
G_SE_SE616	AGTTSSLQEQ	ITWMTG.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
H_BE_VI991	AGTTSTLQEQ	VAWMTG.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
H_BE_VI997	AGTTSTLQEQ	IAWMTG.NPS	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
H_CF_90CF0	AGTTSTLQEQ	IAWMTG.NPA	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI
J_SE_SE702	AGTTSNLQEQ	IGWMTG.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
J_SE_SE788	AGTTSTLQEQ	IGWMTG.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
K_CD_EQTB1	AGTTSTLQEQ	ITWMTS.NPP	IPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
K_CM_MP535	AGTTSTLQEQ	IAWMTS.NPP	VPVGEIYKRW	IILGLNKIVR	MYSPPVSILDI
N_CM_YBF30	AGTTSTLAEQ	VAWMTA.NPP	VPVGDIYRRW	IVLGLNRIVR	MYSPPVSILEI
O_CM_ANT70	AGTTSTQQEQ	IHWTTTRPNQP	IPVGDIYRKW	IVLGLNKMVK	MYSPPVSILDI
O_CM_MVP51	AGTTSTQQEQ	IIWTTTRGANS	IPVGDIYRKW	IVLGLNKMVK	MYSPPVSILDI
O_SN_99SE_	AGTTSTQQEQ	VHWITRPNQP	IPVGDIYRKW	IVLGLNKVVK	MYSPPVSILDI
O_SN_99SE_	AGTTSTQQEQ	VHWITRPNQP	IPVGDIYRKW	IVLGLNKVVK	MYSPPVSILDI
U_CD___83C	AGTTSTLQEQ	IGWMTS.NPP	IPVGDIYKRW	IILGLNKIVR	MYSPPVSILDI

301

350

00BW0762_1	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR
00BW0768_2	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW0874_2	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR
00BW1471_2	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW1616_2	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW1686_8	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKAILR
00BW1759_3	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLIQN	ANPDCKTILK
00BW1773_2	KQGPKEPFRD	YVDRFFKVLK	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW1783_5	KQGPKEPFRD	YVDRFFKTLR	AEQSTQEVKN	WMTDTLLIQN	ANPDCKTILR
00BW1795_6	KQGPKEPFRD	YVDRFFKTLR	AEQSTQEVKN	WMTDTLLVQN	ANPDCKTILK
00BW1811_3	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW1859_5	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW1880_2	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR

00BW1921_1	KQGPKEPFRD	YVDRFFKTLR	AEQSSQEVKN	WMTDTLLVQN	ANPDCKTILR
00BW2036_1	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW2063_6	KQGPKEPFRD	YVDRFFKTLR	AEQSTQEVKN	WMTDTLLVQN	ANPDCKTILR
00BW2087_2	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILK
00BW2127_2	KQGPKEPFRD	YVDRIFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
00BW2128_3	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTKTLLVQN	ANPDCKTILR
00BW2276_7	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW3819_3	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW3842_8	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
00BW3871_3	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTGTLLVQN	ANPDCKTILR
00BW3876_9	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILK
00BW3886_8	KQGPKEPFRD	YVDRFFRTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW3891_6	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW3970_2	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
00BW5031_1	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	GNPDCKNILR
96BW01B21	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
96BW0407	KQGPKEPFRD	YVDRFFKTLR	AEQSTQEVKN	WMTDTLLVQN	ANPDCKTILR
96BW0502	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
96BW06_J4	KQGPKEPFRD	YVDRFFKTLR	AEQCTQDVKN	WMTDTLLVQN	ANPDCKTILR
96BW11_06	KQGPKEPFRD	YVDRFFKTLR	AEQSSQEVKN	WMTDTLLVQN	ANPDCKTILK
96BW1210	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
96BW15B03	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKPILR
96BW16_26	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
96BW17A09	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
96BWM01_5	KQGPKEPFRD	YVDRFFKTLR	AEQSSQEVKN	WMTDTLLVQN	ANPDCKTILR
96BWM03_2	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
98BWMC12_2	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLIQN	ANPDCKTILR
98BWMC13_4	KQGPKEPFRD	YVDRFFKTLR	AEQATREVKV	WMTDTLLVQN	ANPDCKTILR
98BWMC14_a	KQGPKEPFRD	YVDRFFKCLR	AEQATQEVKD	WMTETLLVQN	ANPDCKTILR
98BWM014_1	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
98BWM018_d	KQGPKEPFRD	YVDRFFKTLR	AEQSTQEVKN	WMTDTLLVQN	ANPDCKTILR
98BWM036_a	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR
98BWM037_d	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR
99BW3932_1	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
99BW4642_4	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR
99BW4745_8	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMADTLLVQN	ANPDCKTILR
99BW4754_7	KQGPKEPFRD	YVDRFFKTLR	AEQSTQDVKN	WMTDTLLVQN	ANPDCKTILR
99BWMC16_8	KQGPKEPFRD	YVDRFFKTLR	AEQASQDVKN	WMTDTLLVQN	ANPDCKTILR
A2_CD_97CD	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKSILR
A2_CY_94CY	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKRSILR
A2D___97KR	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILR
A2G_CD_97C					
A_BY_97BL0	RQGPKEPFRD	YVDRFFKTLX	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
A_KE_Q23_A	KQGPKEPFRD	YVDRFFKTFR	AEQATQDVKN	WMTDTLLVQN	ANPDCKAILR
A_SE_SE659					
A_SE_SE725					
A_SE_SE753	KQGPKEPFRD	YVDRFFKTLR	AEQASQDVKN	WMTETLLIQN	ANPDCKSILR
A_SE_SE853					
A_SE_SE889					
A_SE_UGSE8	KQGPKEPFRD	YVDRFFKALR	AEQATQEVKG	WMTDTLLVQN	ANPDCKSILR
A_UG_92UG0	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKG	WMTETLLIQN	ANPDCKSILR
A_UG_U455_	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKSILR
AC_IN_2130	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
AC_RW_92RW	KQGPKEPFRD	YVDRFFKTLR	AEQASQDVKN	WMTDTLLVQN	ANPDCKTILR
AC_SE_SE94					
ACD_SE_SE8	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILK
ACG_BE_VI1	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKRTILR
AD_SE_SE69	RQGPKEPFRD	YVDRFYKTLR	AEQASRDVKN	WMTETLLVQN	ANPDCKTILK
AD_SE_SE71					
ADHK_NO_97	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILK
ADK_CD_MAL	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILK

AG_BE_VI11	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
AG_NG_92NG	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR
AGHU_GA_VI	RQGPKEPFRE	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILR
AGU_CD_Z32	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKG	WMTDTLLVQN	ANPDCKTILR
AJ_BW_BW21	RQGPKEPFRD	YVDRFFKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK
B_AU_VH_AF	KQGPKEPFRD	YVDRFYRTL	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_CN_RL42	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK
B_DE_D31_U	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILK
B_DE_HAN_U	KQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILK
B_FR_HXB2	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_GA_OYI	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK
B_GB_CAM1	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_GB_GB8_A	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK
B_GB_MANC	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_KR_WK_AF	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_NL_3202A	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_TW_TWCYS	RQGPKEPFRD	YVDRFYKTLR	TEQASQDVKN	WMTETLLVQN	ANPDCKTILK
B_US_BC_L0	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_US_DH123	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	SNPDCKTILK
B_US_JRCSF	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILK
B_US_MNCG	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	RTTETLLVQN	ANPDCKTILK
B_US_P896	KQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_US_RF_M1	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETFLVQN	ANPDCKTILK
B_US_SF2_K	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK
B_US_WEAU1	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILK
B_US_WR27	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
B_US_YU2_M	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
BF1_BR_93B	RQGPKEPFRD	YVDRFYKTLR	AEQTSQDVKN	WMTETLLVQN	ANPDCKTILK
C_BR_92BR0	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
C_BW_96BW0	KQGPKEPFRD	YVDRFFKTLR	AEQSTQEVKN	WMTDTLLVQN	ANPDCKTILR
C_BW_96BW1	KQGPKEPFRD	YVDRFFKTLR	AEQSSQEVKN	WMTDTLLVQN	ANPDCKTILK
C_BW_96BW1	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
C_BW_96BW1	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKPILR
C_ET_ETH22	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
C_IN_93IN1	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
C_IN_93IN9	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
C_IN_93IN9	KQGPKEPFRD	YVDRFFRTL	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
C_IN_94IN1	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILR
C_IN_95IN2	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
CRF01_AE_C	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_C	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_C	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_T	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_T	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_T	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_T	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_T	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_T	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF01_AE_T	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF02_AG_F	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	RMTETLLVQN	ANPDCKSILR
CRF02_AG_F	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILR
CRF02_AG_G	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILR
CRF02_AG_N	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILR
CRF02_AG_S	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKSILR
CRF02_AG_S	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTETLLVQN	SNPDCKTILK
CRF03_AB_R	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
CRF03_AB_R	RQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
CRF04_cpx	RQGPKEPFRD	YVDRFFKCLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF04_cpx	RQGPKEPFRD	YVDRFFKCLR	AEQATQEVKN	WMTETLLVQN	ANPDCKSILK
CRF04_cpx	RQGTKEPFRD	YVDRFFKCLR	AEQATQDVKN	WMTETLLVQN	ANPDCKSILK
CRF05_DF_B	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKTILK
CRF05_DF_B	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK

CRF06_cpx_	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR
CRF06_cpx_	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILK
CRF06_cpx_	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILK
CRF06_cpx_	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILK
CRF11_cpx_	KQGPKEPFRD	YVDRFFKALR	AEQATQEVKG	WMTDTLLIQN	ANPDCKSILR
CRF11_cpx_	RQGPKEPFRD	YVDRFFKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKSILR
D_CD_84ZR0	RQGPKEPFRD	YVDRFYKTLR	AEQASQEVKN	WMTETLLVQN	ANPDCKTILK
D_CD_ELI_K	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK
D_CD_NDK_M	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK
D_UG_94UG1	RQGPKEPFRD	YVDRFYKTLR	AEQASQDVKN	WMTETLLVQN	ANPDCKTILK
F1_BE_VI85	KQGPKEPFRD	YVDRFFKVLR	AEQASQDVKG	WMTDTLLVQN	ANPDCKTILK
F1_BR_93BR	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKG	WMTDTLLVQN	ANPDCKTILK
F1_FI_FIN9	RQGPKEPFRD	YVDRFFKALR	AEQATQEVKG	WMTDTLLVQN	ANPDCKIILK
F1_FR_MP41	RQGPKEPFRD	YVDRFFKTLR	AEQASQEVKN	WMTESLLVQN	SNPDCKTILK
F2_CM_MP25	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKG	WMTETLLVQN	ANPDCKTILK
F2KU_BE_VI	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLIQN	ANPDCKTILK
G_BE_DRCBL	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKS	WMTDTLLIQN	ANPDCKIILK
G_NG_92NG0	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKG	WMTDTLLVQN	ANPDCKTILR
G_SE_SE616	RQGPKEPFRD	YVDRFFKCLR	AEQASQDVKG	WMTDTLLVQN	ANPDCKTILR
H_BE_VI991	KQGPKEPFRD	YVDRFFRVLR	AEQATQDVKN	WMTDTLLVQN	ANPDCRTILK
H_BE_VI997	KQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILR
H_CF_90CF0	KQGPKEPFRD	YVDRFFKTLR	AEQATQDVKN	WMTETLLVQN	ANPDCKTILR
J_SE_SE702	RQGPKEPFRD	YVDRFFKALR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILK
J_SE_SE788	RQGPKEPFRD	YVDRFFKALR	AEQATQDVKN	WMTDTLLVQN	ANPDCKTILK
K_CD_EQTB1	RQGPKEPFRD	YVDRFFRVLR	AEQATQEVKN	WMTETLLVQN	ANPDCRTILK
K_CM_MP535	RQGPKEPFRD	YVDRFFKTLR	AEQATQEVKN	WMTDTLLVQN	ANPDCKTILK
N_CM_YBF30	KQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKQLLK
O_CM_ANT70	KQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKQILK
O_CM_MVP51	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	SNPDCKQILK
O_SN_99SE_	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKQILK
O_SN_99SE_	RQGPKEPFRD	YVDRFYKTLR	AEQATQEVKN	WMTETLLVQN	ANPDCKQILK
U_CD_83C	KQGPKEPFRD	YVDRFFKVLR	AEQSSQEVKN	WMTDTLLIQN	ANPDCKTILK

351

400

00BW0762_1	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQANN...	...TNIMLQK
00BW0768_2	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQHN....	...TSIMMQK
00BW0874_2	GLGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQSN....	...S.IMMQR
00BW1471_2	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQANN...	...TNVMMQK
00BW1616_2	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAGN...	...ATIMMQK
00BW1686_8	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQANS...	...PNILMQR
00BW1759_3	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSKANN...	...ANILMQR
00BW1773_2	ALGPAASLEE	MMTACQGVGG	PSHKARVLAE	AMSQANT...	...TNIMMQK
00BW1783_5	ALGQGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQVGN...	...PQVMMQR
00BW1795_6	ALGSGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQANN...	...VQIMMQK
00BW1811_3	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQTTS...	...MQIMMQR
00BW1859_5	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQANN...	...ANIMMQR
00BW1880_2	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQATN...	...TSIMMQK
00BW1921_1	ALGPGATLEE	MMTACQGVGG	PSHKARALAE	AMSQANN...	...TNIMMQK
00BW2036_1	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVNG...	...SNVLMQR
00BW2063_6	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQANN...	...INVMMQK
00BW2087_2	ALGSGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQANS...	...TNIMIQR
00BW2127_2	ALGSGVTLEE	MMSACRGVGG	PSHKARVLAE	AMSQANN...	...TNIMMQR
00BW2128_3	ALGPAATLVE	MMTACQGVGG	PSHKARVLAE	AMSQTAS...	...TNILMQR
00BW2276_7	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQAN....	...SNIMMQR
00BW3819_3	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQANN...	...INVMMQR
00BW3842_8	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQANS...	...TNIMMQR
00BW3871_3	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQANN...	...SNIMMQK
00BW3876_9	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQAGS...	...VNILMQR
00BW3886_8	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQANN...	...INVMMQR
00BW3891_6	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQTN....	...SNIMMQN
00BW3970_2	AIGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQTH....	...SNIMMQR

00BW5031_1	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQANN...	...INIMMQR
96BW01B21	ALGPGASIEE	MMTACQGVGG	PSHKARVLAE	AMSQANS...	...MNIMMQR
96BW0407	ALGPGASLEE	MMTACQGVGG	PSHKARVLGE	AMSQA.N...	...TNVMMQK
96BW0502	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQANS...	...VNIMMQR
96BW06_J4	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQANS...	...TSILMQR
96BW11_06	TLGPGASLEE	MMTACQGVGG	PSHKARILAE	AMSQANN...	...PNIMMQR
96BW1210	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSHAGN...	...AGIMMQR
96BW15B03	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQATS...	...ANILMQR
96BW16_26	ASGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQANN...	...TNIMIQR
96BW17A09	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSHANN...	...TSIMMQR
96BWM01_5	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQANN...	...IQVMMQK
96BWM03_2	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQAN...	...TNIMMQR
98BWMC12_2	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSHTN....	...S.IMMQR
98BWMC13_4	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQANN...	...INIMMQR
98BWMC14_a	SLGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAN....	...TSIMMQR
98BWM014_1	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQANS...	...TNILIQR
98BWM018_d	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQAGN...	...ANIMMQR
98BWM036_a	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQANN...	...ANIMMQR
98BWM037_d	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQAN....	...SNIMMQR
99BW3932_1	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQANN...	...ASVMMQR
99BW4642_4	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMGQVNS...	...TNIMMQR
99BW4745_8	GLGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQTN....	...SNIMMQR
99BW4754_7	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQANN...	...PNIMMQR
99BWMC16_8	ALGPGASIEE	MMTACQGVGG	PSHKARVSAE	AMSQANQ...	...ANIMMQR
A2_CD_97CD	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQNT..	...NIMIQR
A2_CY_94CY	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSHVQSTN.	..T.NIMMQR
A2D___97KR	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQVQNTN.	..SNIMMMQR
A2G_CD_97C					
A_BY_97BL0	ALGSEATLEE	MMTACQGVGG	PGHKAXVXAE	AMSQVQN...	...ANIMMQR
A_KE_Q23_A	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQH...	...TNIMMQR
A_SE_SE659					
A_SE_SE725					
A_SE_SE753	ALGAGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQH...	...PNIMMQR
A_SE_SE853					
A_SE_SE889					
A_SE_UGSE8	GLGAGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAQQ...	...TNIMMQR
A_UG_92UG0	ALGAGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQH...	...TNIMMQR
A_UG_U455_	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQQT..	SIMMQR
AC_IN_2130	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQTN....	...NSILMQR
AC_RW_92RW	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQQ...	...PNIMMQR
AC_SE_SE94					
ACD_SE_SE8	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQAN....	..ANTAIMMQR
ACG_BE_VI1	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQANN...	TVMMQR
AD_SE_SE69	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQATNN..	..INAAIMMQR
AD_SE_SE71					
ADHK_NO_97	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQN...	...VSVMMQR
ADK_CD_MAL	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQATN...	..STAAIMMQR
AG_BE_VI11	ALGTGATLEE	MVTACQGVGG	PGHKARVLAE	AMSQVH....	..QTNIMMQR
AG_NG_92NG	ALGAGATLEE	MLTACQGVGG	PSHKARVLAE	AMSRAT.G..	..TSAAIMMQR
AGHU_GA_VI	GLGAGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVHN...	...TSIMMQR
AGU_CD_Z32	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAS.S..	..TAAAIMMQR
AJ_BW_BW21	ALGAGATLEE	MMTACQGVG.	PGHKAKILAE	AMSQVQN...	...TNIMMQR
B_AU_VH_AF	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	..NSATIMMQR
B_CN_RL42_	ALGPAATLEE	MMTACQGVGG	PSHKARILAE	AMSQVTNSAI	TNSATIMMQR
B_DE_D31_U	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	..NSATVMMQR
B_DE_HAN_U	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	..GSAAIMMQR
B_FR_HXB2_	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	..NSATIMMQR
B_GA_OYI_	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVN....	..SVTVMMQR
B_GB_CAM1_	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	..NSATIMMQR
B_GB_GB8_A	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVN....	..SATIMMQR
B_GB_MANC_	ALGPAATLEE	MMTACQGVGG	PSHKARILAE	AMSQVT....	..NSATIMMQR

B_KR_WK_AF	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAT....	.NSATIMMQR
B_NL_3202A	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NSATIMMQR
B_TW_TWCYS	ALGPAATLEE	MMTACQGVGG	PSHKARVLAE	AMSRVP....	.NSTNIMMQR
B_US_BC_L0	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NSATIMMQR
B_US_DH123	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQITN...	.TSATIMMQG
B_US_JRCSF	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NPATIMMQR
B_US_MNCG_	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NSATIMMQR
B_US_P896_	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NSATIMMQR
B_US_RF_M1	ALGPAATLEE	MMTACQGVGG	PSHKARILAE	AMSQVT....	.NSATIMLQK
B_US_SF2_K	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NPANIMMQR
B_US_WEAU1	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NTATIMMQR
B_US_WR27_	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.GSNAIMMQK
B_US_YU2_M	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NSATIMMQR
BF1_BR_93B	ALGPAATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVT....	.NSGTIMMQR
C_BR_92BR0	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSKVNN...	...TNIMMQR
C_BW_96BW0	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQT.N...	...TNVMMQR
C_BW_96BW1	ALGPGASLEE	MMTACQGVGG	PSHKARILAE	AMSQANN...	...SNIMMQK
C_BW_96BW1	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSHAGN...	...AGIMMQR
C_BW_96BW1	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQATS...	...ANILMQR
C_ET_ETH22	ALGPGASLEE	MMTACQGVGG	PAHKARVLAE	AMSQVNN...	...TTIMMQK
C_IN_93IN1	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQAN....	...STILMQR
C_IN_93IN9	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQTN....	...SAILMQK
C_IN_93IN9	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQAN....	...SNILMQR
C_IN_94IN1	ALGPGASLEE	MMTACQGVGG	PSHKAKVLAE	AMSQTN....	...SAILMQR
C_IN_95IN2	ALGPGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQTN....	...SAILMQR
CRF01_AE_C	ALGTGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQ....	..HANIMMQR
CRF01_AE_C	ALGTGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVH....	..HTNIMMQK
CRF01_AE_C	ALGTGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQ....	..HANIMMQR
CRF01_AE_T	SLGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSHAQ....	..HANIMMQR
CRF01_AE_T	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAQ....	..HANIMMQR
CRF01_AE_T	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAQ....	..HANIMMQR
CRF01_AE_T	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSHAQ....	..QANIMMQR
CRF01_AE_T	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSHAQ....	..HATIMMQR
CRF01_AE_T	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSHVQ....	..QANIMMQR
CRF02_AG_F	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQ....	..QSNIMMQR
CRF02_AG_F	ALGPGASLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQ....	..QSNIMMQR
CRF02_AG_G	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQSP..	...NILMQR
CRF02_AG_N	ALGTGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQ....	..QANVMMQR
CRF02_AG_S	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQVQ....	..QPNIMMQR
CRF02_AG_S	ALGPGATLEE	MMSACQGVGG	PGHKARVLAE	AMSQAQ....	..QSNIMMQR
CRF03_AB_R	ALGSGATLEE	MMTVCQGVGG	PGHKARVLAE	AMSQVQN...	...ANIMMQK
CRF03_AB_R	ALGSGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVQN...	...ANIMMQK
CRF04_cpx_	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQASN...	.AAAAIMMQK
CRF04_cpx_	ALGTGATLEE	MMTACQGVGG	PAHKARVLAE	AMSQASN...	.AAAAIMMQK
CRF04_cpx_	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQASS...	.AAAAIMMQK
CRF05_DF_B	ALGQQATLEE	MMTACQGVGG	PSHKARVLAE	AMSQATN...	.SAATAMMQR
CRF05_DF_B	ALGPQATLEE	MMTACQGVGG	PSHKARVLAE	AMSQATG...	.SPAVMMQ.R
CRF06_cpx_	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQAS....	.VGAIMMQ.K
CRF06_cpx_	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQASGT..	.LTAIMMQ.K
CRF06_cpx_	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQASGT..	.VAAAIMMQK
CRF06_cpx_	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVSGA..	.TAAIMMQ.K
CRF11_cpx_	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQVQQ...	...TNVMMQR
CRF11_cpx_	ALGVGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAQH...	...TNIMMQR
D_CD_84ZR0	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQATS...	.ASAAIMMQK
D_CD_ELI_K	ALGPQATLEE	MMTACQGVGG	PSHKARVLAE	AMSQATN...	.SVTTAMMQR
D_CD_NDK_M	ALGPQATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVTG...	.SATAVMMQR
D_UG_94UG1	ALGPAATLEE	MMTACQGVGG	PSHKARVLAE	AMSQATN...	.ANTAIMMQR
F1_BE_VI85	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAN....	...SAIMMQK
F1_BR_93BR	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQATN...	...TAIMMQK
F1_FI_FIN9	GLGIGATLEE	MMTACRGVGG	PGHKARILAE	AMSQAN....	...TTIMMQK
F1_FR_MP41	ALGPGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQATN...	...AAIMMQK

F2_CM_MP25	ALGPGATLEE	MMTACQGVGG	PSHKARILAE	AMSKATG...	...AAIMMQK
F2KU_BE_VI	ALGPGASLEE	MMTACQGVGG	PAHKARVLAE	AMSQATN...	...TAIMMQK
G_BE_DRCBL	GLGQGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQAS.G..	.AAAAIMMQK
G_NG_92NG0	ALGPGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQASGA..	.AAAAIMMQK
G_SE_SE616	ALGQGASLEE	MMTACQGVGG	PSHKARVLAE	AMSQAS.G..	.AAAAIMMQK
H_BE_VI991	ALGRGASIEE	MMTACQGVGG	PSHKARVLAE	AMSQVTNAS.	...AAIMMQK
H_BE_VI997	ALGQGASIEE	MMTACQGVGG	PSHKARVLAE	AMSQVTNAN.	...AAIMMQK
H_CF_90CF0	ALGQGASIEE	MMTACQGVGG	PSHKARVLAE	AMSQVTNTN.	...TAIMMQK
J_SE_SE702	ALGSGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVTN...	...TNIMMQK
J_SE_SE788	ALGSGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVTN...	...TNIMMQK
K_CD_EQTB1	ALGSGATLEE	MMTACQGVGG	PGHKARVLAE	AMSQVTN...	...SAVMMQK
K_CM_MP535	ALGPGASLEE	MMTACQGVGG	PSHKARILAE	AMSQVTN...	...PVVMMQK
N_CM_YBF30	ALGPGATLEE	MMTACQGVGG	PAHKARVLAE	AMSQVQPP..	..TTSVFQK
O_CM_ANT70	SLGPGATLEE	MMVACQGVGG	PTHKARVLAE	AMATAQQDLK	GGYTAVFMQK
O_CM_MVP51	ALGPEATLEE	MMVACQGVGG	PTHKAKILAE	AMASAQQDLK	GGYTAVFMQK
O_SN_99SE_	SLGPGATLEE	MMIACQGVGG	PTHKARVLAE	AMA.AAQDLK	GGYTAVFMQK
O_SN_99SE_	SLGPGATLEE	MMIACQGVGG	PTHKARVLAE	AMS.AAQDLK	GGYSVAVFMQK
U_CD____83C	ALGTGATLEE	MMTACQGVGG	PSHKARVLAE	AMSQTN....	...TAIMMQK

401

450

00BW0762_1	NNFKGPRRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW0768_2	SNFKGPKRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW0874_2	SNFKGHRKIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1471_2	GNFKGPRRVI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKECT..E
00BW1616_2	SNFKGPRRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1686_8	SNFKGSK.RV	KCFNCGKEGH	IARNCRAPRK	RGCWKCCKEG	HQMKDCT..E
00BW1759_3	SNFKGPKRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1773_2	SNFKGPRRTV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1783_5	NNLKGTRRTV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1795_6	NNFKGPRRII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1811_3	SNFKGSKRSV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1859_5	SNFKGPRKII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1880_2	SNFKGPRRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW1921_1	SNFKGPRRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW2036_1	GNFKGPKRII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW2063_6	NNFKGPRRLV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW2087_2	SNFKGSKRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW2127_2	GNFRGAKRSV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW2128_3	SNFKGSKRSV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW2276_7	SNFKGPRKIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCTN.D
00BW3819_3	SNFKGSKRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDYT..E
00BW3842_8	GNFKGPKRTI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW3871_3	SNFKGPRRII	KCFNCGKEGH	LARNCRAPRK	RGCWKCCKEG	HQMKDCTT.E
00BW3876_9	GNFKGSKRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW3886_8	GNFKGAKRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW3891_6	SNFKGSRRIV	KCFNCGKVGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW3970_2	SNFKGPKRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
00BW5031_1	NNFKGPKRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
96BW01B21	SNFKNPKRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
96BW0407	SNFKGPRRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
96BW0502	SNFKGPRRNV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
96BW06_J4	GNFKGPKRII	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
96BW11_06	NNFKGPRRIV	KCFNCGKEGH	IARNCKAPRK	KGCWKCCKEG	HQMKDCT..E
96BW1210	GNFKGPRKIP	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCS..E
96BW15B03	SNFKGPKRII	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
96BW16_26	SNFKGPRRSV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
96BW17A09	GNFKGPRRTI	KCFNCGKEGH	LAKNCRAPRK	XGCWKCCKEG	HQMKECT..E
96BWM01_5	NNFKGPRRTI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
96BWM03_2	GNFKGPKRII	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCNN.E
98BWMC12_2	SNFKGPKRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCN..E
98BWMC13_4	SNFKGPKRII	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E

98BWMC14_a	GNFKGPRRII	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
98BWM014_1	SHFKGSKRTV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
98BWM018_d	GNFKGPRRII	KCSNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
98BWM036_a	SNFKGPKRTV	KCFNCGKEGH	VARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
98BWM037_d	GNFKGSKRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
99BW3932_1	SNFKGPKRII	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
99BW4642_4	SNFKGPKRII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
99BW4745_8	SNFKGPRRTV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
99BW4754_7	XNFKGPRRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
99BWMC16_8	SNFKGPKRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
A2_CD_97CD	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
A2_CY_94CY	GNFRGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
A2D_97KR	GNFRGQK..I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
A2G_CD_97C					
A_BY_97BL0	SNFRGPKR.I	KCXNCGKEGH	LARNCRAPRK	XGCWKCCKEG	HQMKDCTE..
A_KE_Q23_A	GNFKGQKR.I	KCFNCGKEGH	LARNCRALRK	KGCWKCCKEG	HQMKDCTE..
A_SE_SE659					
A_SE_SE725					
A_SE_SE753	GNFRGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
A_SE_SE853					
A_SE_SE889					
A_SE_UGSE8	GNFRGQKK.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
A_UG_92UG0	GNFKGQKR.I	KCFNCGKEGH	LAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
A_UG_U455_	GNFRGPRR.I	KCFNCGKEGH	LAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
AC_IN_2130	SNFKGFKRTV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
AC_RW_92RW	GNFKGQRRII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
AC_SE_SE94					
ACD_SE_SE8	GNFKGPKKII	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
ACG_BE_VI1	GNFKGPKRII	KCFNCGKEGH	VARNCRAPRK	KGCWKCCKEG	HQMKDCTT.E
AD_SE_SE69	GNFKGQRKII	KCFNCGKLGH	IARNCKAPRK	KGCWKCCKEG	HQMKDCTE..
AD_SE_SE71					
ADHK_NO_97	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
ADK_CD_MAL	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
AG_BE_VI11	GNFRGQKT.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
AG_NG_92NG	NNFKGPRRGI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
AGHU_GA_VI	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
AGU_CD_Z32	SNFKGPRKTI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDC.E..
AJ_BW_BW21	SNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_AU_VH_AF	GNFRNQRKTIV	KCFNCGKEGH	IARNCRAPRK	RGCWKCCKEG	HQMKDCTE..
B_CN_RL42_	GNFRNQRKIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HLMKDCTE..
B_DE_D31_U	GNFRNQRKTIV	KCFNCGKEGH	IAKNCRAPRR	KGCWKCCKEG	HQMKDCTE..
B_DE_HAN_U	GNFRNQRKTIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_FR_HXB2_	GNFRNQRKIV	KCFNCGKEGH	TARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_GA_OYI_	GNFKNQRKTIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_GB_CAM1_	GNFRNQRKTIV	KCFNCGKVGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCNE..
B_GB_GB8_A	GNFRSQRKTIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKECTE..
B_GB_MANC_	GNFRNQRKTIV	KCFNCGKEGH	IAKNCRAPRR	KGCWKCCKEG	HQMKDCTE..
B_KR_WK_AF	GNFRNQRRTV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_NL_3202A	GNFRNQRKIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_TW_TWCYS	GNFRNQRKTIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_US_BC_L0	GNFRNQRKTIV	KCFNCGKEGH	IARNCKAPRK	RGCWKCCKEG	HQMKDCTE..
B_US_DH123	GNFRNQRK.I	KCFNCGKEGH	ISKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_US_JRCSF	GNFRNQRKNV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKECTE..
B_US_MNCG_	GNFRNQRKII	KCFNCGKEGH	IAKNCRAPRK	RGCWKCCKEG	HQMKDCTE..
B_US_P896_	GNFRNQRKTIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_US_RF_M1	GNFRDQRKIV	KCFNCGKVGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCTNEG
B_US_SF2_K	GNFRNQRKTIV	KCFNCGKEGH	IAKNCRAPRK	KGCWRCCKEG	HQMKDCTE..
B_US_WEAU1	GNFRSPRKTI	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
B_US_WR27_	GNFRNQRKTIV	KCXXCCKEGH	XARXCKAPRK	RGCWKCCKEG	HQMXDXXE..
B_US_YU2_M	GNFRNQRKTIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
BF1_BR_93B	GNFRNQRKTI	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..

C_BR_92BR0	SNCKGPKRTI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQVKDCT..E
C_BW_96BW0	SNFKGPRRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
C_BW_96BW1	NNFKGPTRIV	KCFNCGKEGH	IARNCKAPRK	KGCWKCCKEG	HQMKDCT..E
C_BW_96BW1	GNFKGPRKIP	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCS..E
C_BW_96BW1	SNFKGPKRII	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
C_ET_ETH22	SNFKGPKRAI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
C_IN_93IN1	SNFKGSKRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
C_IN_93IN9	SNFKGSKRII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
C_IN_93IN9	SNFKGSKRTV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
C_IN_94IN1	SNFKGSKRII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
C_IN_95IN2	SNFKGSKRIV	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
CRF01_AE_C	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	RGCWKCCKEG	HQMKDCTE..
CRF01_AE_C	GNFKGQRK.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF01_AE_C	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF01_AE_T	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF01_AE_T	GNFKGPRR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF01_AE_T	GNFKGQTR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF01_AE_T	GNFKGQKR.I	KCFNCGKEGH	LARNCRALRK	KGCWKCCKEG	HQMKDCTE..
CRF01_AE_T	GNFKGQKR.I	KCFNCGREGH	LARNCRAPRK	QGCWKCCKEG	HQMKDCTE..
CRF01_AE_T	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF02_AG_F	GNFRGQRT.I	KCFNCGKEGH	LARNCKAPRK	KGCWKCCKEG	HQMKDCTE..
CRF02_AG_F	GNFRGQRT.I	KCFNCGKEGH	LARNCKAPRK	KGCWKCCKEG	HQMKDCTE..
CRF02_AG_G	GNFKGQK.RI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
CRF02_AG_N	GNFRGQRT.I	KCFNCGKEGH	LARNCKAPRK	RGCWKCCKEG	HQMKDCTE..
CRF02_AG_S	GNFRGQRP.I	KCFNCGKEGH	LARNCKAPRK	KGCWKCCKEG	HQMKDCTE..
CRF02_AG_S	GNFRGQRT.I	KCFNCGKEGH	LARNCKAPRK	RGCWKCCKEG	HQMKDCTE..
CRF03_AB_R	SNFRGPKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF03_AB_R	SNFRGPKR.I	KCFNCGKDGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCNE..
CRF04_cpx_	SKFKGQRRTI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF04_cpx_	SNFRGQKR.II	KCFNCGKEGH	LARNCRAPRK	RGRWKCCKEG	HQMKDCTE..
CRF04_cpx_	SNFKGPRRSI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCPE..
CRF05_DF_B	GNFKGPRKII	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF05_DF_B	GNFKGPRKSI	KCFNCGKEGH	TAKNCRAPRK	RGCWKCCKEG	HQMKDCIE..
CRF06_cpx_	SNFKGPKRSI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF06_cpx_	SNYKGPKRSI	KCFNCGREGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF06_cpx_	SNFKGP.RKI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF06_cpx_	SNFKGQRKNI	KCFNCGKEGH	TARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF11_cpx_	SNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
CRF11_cpx_	GNFKGQKR.I	KCFNCGKEGH	LARNCRAPRK	K.GCKCKEG	HQMKDCTE..
D_CD_84ZR0	SNFKGTRKIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
D_CD_ELI_K	GNFKGPRKII	KCFNCGKEGH	IAKNCRAPRK	KGCWRCCKEG	HQLKDCTE..
D_CD_NDK_M	GNFKGPRKSI	KCFNCGKEGH	TAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
D_UG_94UG1	GNFKGPKKII	KCFNCGKEGH	TAKNCRAPRK	KGCWKCCKEG	HQMKDCTE..
F1_BE_VI85	SNFKGQRRVV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
F1_BR_93BR	SNFKGQRRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
F1_FI_FIN9	SNFRGQRRIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
F1_FR_MP41	SNYKGPRRFI	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
F2_CM_MP25	SNFKGQRRIV	KCFNCGKEGH	IARNCRAPRK	RGCWKCCKEG	HQMKDCT..E
F2KU_BE_VI	GNFKGPRRDV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
G_BE_DRCBL	SNFKGPRRTI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKECTE..
G_NG_92NG0	SNFKGPRRII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKECTE..
G_SE_SE616	SNFKGPRRTI	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCTE..
H_BE_VI991	GNFKGPRRTV	KCSNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..G
H_BE_VI997	SNFKGPRKIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
H_CF_90CF0	GNFKGQRKIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
J_SE_SE702	GNFRDHKRIV	KCFNCGKQGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
J_SE_SE788	GNFRDHKRIV	KCFNCGKQGH	IAKNCRAPRK	KGCWKCCKEG	HQMKDCT..E
K_CD_EQTB1	GNFKGQRRII	KCFNCGKEGH	LARNCRAPRK	KGCWKCCKEG	HQMKDCS..E
K_CM_MP535	GNFKGHRKIV	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCT..E
N_CM_YBF30	GNFKGIRKPI	KCFNCGKEGH	LARNCKAPRR	GGCWKCCKEG	HQMKDCKNEG
O_CM_ANT70	GQNPIRKGTI	KCFNCGKEGH	IARNCRAPRK	KGCWKCCKEG	HQMKDCRN.G

O_CM_MVP51	GQNPNRKGPI	KCFNCGKEGH	IAKNCRAPRK	RGCWKCGQEG	HQMKDCKN.G
O_SN_99SE	GQNPSRKGPI	KCFNCGKEGH	LARNCRAPRK	KGCWKCGQEG	HQMKDCKN.G
O_SN_99SE	GQNPGRKGPI	KCFNCGKEGH	LARNCRAPRK	KGCWKCGQEG	HQMKDCRN.G
U_CD_83C	GNFKGPRRIV	KCFNCGKEGH	IAKNCRAPRK	KGCWKCGREG	HQMKDCT..E

	451			500
00BW0762_1	RQANFLGKIW	PSHKG.RPGN	FLQSR.....	PEP TAPPAESFK.
00BW0768_2	RQANFLGKIW	PSHKG.RPGN	FLQNRPEP..	 TAPPAESFK.
00BW0874_2	RQANFLGKIW	PSHKG.RPGN	FLQNRPEPSA	PPAESLRPEP SAPPAESLR.
00BW1471_2	RQANFLGKIW	PSQKG.RPGN	FLQNRPEP..	 SAPPAESFR.
00BW1616_2	RQANFLGKIW	PSHKG.RPGN	FLQSRPEPTA	PP....APVP TAPPAESFR.
00BW1686_8	RQANFLGKIW	PSHKG.RPGN	FLQNRPEP..	 SAPPAESFK.
00BW1759_3	RQANFLGKIW	PSHKG.RPGN	FLQ.....	SRPEP TAPPLESFK.
00BW1773_2	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P TAPPAESFR.
00BW1783_5	RQANFLGKIW	PSQKGGRPGN	FLQNRPA...	ESRLEP TAPPAESFR.
00BW1795_6	RQANFLGKIW	PSHKG.RPGN	FLQNRPE...	P TAPPAESFR.
00BW1811_3	RQANFLGRIW	PSHKG.RPGN	FLQNRPEPTA	P.....LEP TAPPAESFR.
00BW1859_5	RQANFLGKIW	PSHKG.RPGN	FLQNRPEP..	 TAPPAESFR.
00BW1880_2	RQANFLGKIW	PSHKG.RPGN	FLQSR.....	PEP TAPPAESFK.
00BW1921_1	RQANFLGKIW	PSHKG.RPGN	FLQSRPEP..	 TAPPAESFR.
00BW2036_1	RQANFLGKIW	PSNKG.RPGN	FLQNRTPPV	EP TAPPAESFR.
00BW2063_6	RQANFLGKIW	PSHKG.RPGN	FLQSRLE...	P TAPPAESF..
00BW2087_2	RQANFLGKIW	PSHKGGRPGN	FLQSRPEPTA	P.....PAEP TAPPAESFR.
00BW2127_2	RQANFLGKIW	PSHKG.RPGN	FLQNRPEPTA	P.....RPEP SAPPAESFR.
00BW2128_3	RQANFLGRIW	PSNKG.RPGN	FLQNRPEPTA	PPAE.NRPEP TAPPAESFR.
00BW2276_7	RQANFLGKLW	PSNKG.RPGN	FLQNRTEPTA	P.....LEP TAPPADSFK.
00BW3819_3	RQANFLGKIW	PSHKG.RPGN	FLQNRPE...	P.....TAP TAPPAESFR.
00BW3842_8	RQANFLGKIW	PSRGG.RPGN	FLQNRTEPTA	P.....PEP TAPPAESFR.
00BW3871_3	RQANFLGKIW	PSHKG.RPGN	FLQNRPEP..	 TAPPAESFR.
00BW3876_9	RQANFLGKIW	PSHKG.RPGN	FLQNRPE...	P TAPPAESFR.
00BW3886_8	RQANFLGKIW	PSHKG.RPGN	FLQNRPEPTA	P.....PAEP TAPPAESFR.
00BW3891_6	RQANFLGRIW	PSHKG.RPGN	FLQSRPE...	P TAPPAESFR.
00BW3970_2	RQANFLGRFW	PSQKG.RPGN	FLQ.....	SRSEP TAPPAESFR.
00BW5031_1	RQANFLGKIW	PSNKG.RPGN	FLQSRPEPTA	P.....PMP TAPPAESFR.
96BW01B21	RQANFLGKIW	PSHKG.RPGN	FLQNR.....	LEP SAPPAESFR.
96BW0407	RQANFLGKIW	PSHKG.RPGN	FLQ.....	SRPEP TAPPAESFR.
96BW0502	RQANFLGKIW	PSHKG.RPGN	FLQNRSEPA.	APTVP TAPPAESFR.
96BW06_J4	RQANFLGKIW	PSHKGGRPGN	FLQSRPEP..	 TAPPAESFR.
96BW11_06	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P TAPPAESL..
96BW1210	GQANFLGKIW	PSHKG.RPGN	FLQSR.....	PEP SAPPAESFR.
96BW15B03	RQANFLGKIW	PSHKG.RPGN	FLQNRTEP..	 TAPPAESFK.
96BW16_26	RQADFLGKIW	PSHKG.RPGN	FLQSRPE...	P TAPPAESFR.
96BW17A09	RQANFLGKIW	PSHKGGRPGN	FLQNRPEP..	 TAPPAESFR.
96BWM01_5	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P TAPPAEIL..
96BWM03_2	RQANFLGKIW	PSHKG.RPGN	FLQSRPEP..	 TAPPAERFR.
98BWMC12_2	RQANFLGRLW	PSHKG.RPGN	FPQNR.....	VEP TAPPAESLR.
98BWMC13_4	KQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P TAPPAESL..
98BWMC14_a	RQANFLGKIW	PSHKG.RPGN	FLQSRPEP..	 SAPPAESFR.
98BWM014_1	RQANFLGKIW	PSHKGGRPGN	FLQRRPEP..	 TAPPAESFR.
98BWM018_d	RQANFLGKIW	PSHKG.RPGN	FIQNRPAPT.	APPVEP TAPPAESFR.
98BWM036_a	RQANFLGRIW	PSHKG.RPGN	FLQSRPEPTA	P.....PAEP TAPPAESFR.
98BWM037_d	RQANFLGKIW	PSHKG.RPGN	FLQ.....	KRPEP TAPPAESFR.
99BW3932_1	RQVNFLGKIW	PSNKG.RPGN	FLQNRTPVPTA	PPAESFRIEP TAPPAESFR.
99BW4642_4	RQANFLGKIW	PSHKG.RPGN	FFQNRTEP..	 TAPPAESFR.
99BW4745_8	RQANFLGKIW	PSNKG.RPGN	FLQNRPEPTA	P.....LEP TAPPAESFR.
99BW4754_7	RQANFLGKIW	PSNKG.RPGN	FLQSR.....	PEP TAPPAESFK.
99BWMC16_8	RQANFLGKIW	PSNKG.RPGN	FLQNRPEPT.	APLEP TAPPAESFR.
A2_CD_97CD	RQANFLGKIW	PSNKG.RPGN	FPQSRTE...	P TAPPME....
A2_CY_94CY	RQANFLGKIW	PSNKG.RPGN	FPQSRTE...	P TAPPAENLR.
A2D_97KR	RQANFLGKIW	PSHSG.RPGN	FPQSRTE...	P TAPPAEDFG.
A2G_CD_97C				

A_BY_97BL0	RQANFLGRIW	PSSKG.RPGN	FPQSRPE...	PS	APP.AENFR.
A_KE_Q23_A	RQANFLGKIW	PSRKG.RPGN	FPQNRLE...	PT	APP.AETCG.
A_SE_SE659					
A_SE_SE725					
A_SE_SE753	RQANFLGRIW	PSSKG.RPGN	FPQSRLE...	PT	APP.AEIFG.
A_SE_SE853					
A_SE_SE889					
A_SE_UGSE8	RQANFLGKIW	PSHKG.RPGN	FPQSRPE...	PS	APP...AEM.
A_UG_92UG0	RQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PT	APPAAEIFG.
A_UG_U455_	RQANFLGKIW	PSNKG.RPGN	FPQSRPE...	P	TAPPAEIFG.
AC_IN_2130	RQANFLGKIW	PSHKG.RPGN	FLQ.....	NRPEP	TAPPAESFR.
AC_RW_92RW	RQANFLGKIW	PSNKG.RPGN	FPQSRLE...	EP	TAPPA.....
AC_SE_SE94					
ACD_SE_SE8	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	PT	APPA.ESFG.
ACG_BE_VI1	RQANFLGKIW	PSNKG.RPGN	FPQSRPEPTA	PP.....AEP	TAPPAESFG.
AD_SE_SE69	RQANFLGKIW	PSSKG.RPGN	FLQSRP....	EP	TAPPAESFG.
AD_SE_SE71					
ADHK_NO_97	RQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PS	APPA.ESFG.
ADK_CD_MAL	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	PT	APPA.ESFG.
AG_BE_VI11	RQANFLGKIW	PSSKG.RPGN	FPQSRLE...	PT	APPA.ESLG.
AG_NG_92NG	RQANFLGKIW	PSNKG.RPGN	FLQNRPE...	P	TAPPAESFG.
AGHU_GA_VI	RQANFLGKIW	PSNKG.RPGN	FLQNRPE...	PT	APPA.ESFG.
AGU_CD_Z32	RQANFLGKIW	PSNKG.RPGN	FLQNRPE...	P	TAPPAESFE.
AJ_BW_BW21	RQANFLGKIW	PSNKG.RPGN	FLQSRPE...	PT	APPA.ESFG.
B_AU_VH_AF	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_CN_RL42_	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_DE_D31_U	RQANFLGKIW	PSYKG.RPGN	FLQRRPE...	P	TAPPEESFR.
B_DE_HAN_U	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_FR_HXB2_	RQANFLGKIW	PSYKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_GA_OYI_	RQANFLGKIW	PSHKG.RPGN	FLQNRPE...	P	TAPPAESFG.
B_GB_CAM1_	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_GB_GB8_A	RQANFLGKIW	PSHKG.RPGN	FLQSRPEPIA	PP.....EP	TAPPEESFR.
B_GB_MANC_	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_KR_WK_AF	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	SAPPEESFR.
B_NL_3202A	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_TW_TWCYS	RQANFLGKIW	PSHKE.RPGN	FLQSRPE...	P	TAPPEESFR.
B_US_BC_L0	RQANFLGKIW	PSHKG.RPGN	FPQSRLE...	P	TAPPEESFR.
B_US_DH123	RQANFLGKIW	PSHKE.RPGN	FLQSRPE...	P	SAPPEESFR.
B_US_JRCSE	RQANFLGKIW	PSYKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_US_MNCG_	RQANFLGKIW	PSCKG.R.RN	FPQSRTE...	P	TAPPEESFR.
B_US_P896_	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_US_RF_M1	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_US_SF2_K	RQANFLGKIW	PSYKG.RPGN	FLQSRPE...	P	TAPPEESFR.
B_US_WEAU1	RQANFLGKIW	SSQKG.RPGN	FPQSRLE...	P	TAPPEESFR.
B_US_WR27_	RQAXFLGXIR	PSHKG.RPGX	FLQNRPE...	P	SAPPAESFR.
B_US_YU2_M	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPSEESVR.
BF1_BR_93B	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPAESFR.
C_BR_92BR0	RQANFLGKIW	PSHRG.RPGN	LLQNRTE...	EP	TAPPE.....
C_BW_96BW0	RQANFLGKIW	PSHKG.RPGN	FLQ.....	SRPEP	TAPPAESFR.
C_BW_96BW1	RRANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPAESF..
C_BW_96BW1	GQANFLGKIW	PSHKG.RPGN	FLQSR....	PEP	SAPPAESFR.
C_BW_96BW1	RQANFLGKIW	PSHKG.RPGN	FLQNRTEP..		TAPPAESFK.
C_ET_ETH22	RQANFLGRLW	PSNKG.RPGN	FLQSRP....	EP	TAPPEESLRPE
C_IN_93IN1	RQANFLGKIW	PSHKG.RPGN	FLQ.....	SRPEP	TAPPAESFR.
C_IN_93IN9	RQANFLGKIW	PSHKG.RPGN	FLQ.....	SRPEP	TAPPAESFR.
C_IN_93IN9	RQANFLGKIW	PSHKG.RPGN	FLQNRPEPTA	PP...ARPEP	TAPPAESFR.
C_IN_94IN1	RQANFLGKIW	PSHKG.RPGN	FLQ.....	SRPEP	TAPPAESFR.
C_IN_95IN2	RQANFLGKIW	PSHKG.RPGN	FLQ.....	SRPEP	TAPPAESFR.
CRF01_AE_C	RQANFLGKIW	PLNKG.RPGN	FPQSRLE...	PT	APPA.ESLG.
CRF01_AE_C	RQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PT	APPM.ESLG.
CRF01_AE_C	RQANFLGRIW	PSSKG.RPGN	FPQSRPE...	PT	APPA.ESLG.

CRF01_AE_T	RQANFLGKFW	PSNKG.RPGN	FPQSRPE...	PT	APPA.ENWG.
CRF01_AE_T	RQANFLGKIW	PSNKG.RPGN	FPQSRPE...	PT	APP..AEWG.
CRF01_AE_T	RQANFLGKIW	PSNKG.RPGN	FPQSKPE...	PT	APPA.ENWG.
CRF01_AE_T	RQANFLGKIW	PSNKG.RPGN	FPQSRPE...	PT	APPA.ENWG.
CRF01_AE_T	RQANFLGKIW	PSNKG.RPGN	FPQSRPE...	PT	APPA.ENWG.
CRF01_AE_T	RQANFLGKIW	PSNKG.RPGN	FPQSRPE...	PT	APPA.ENWG.
CRF02_AG_F	GQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PT	APPA.ESLG.
CRF02_AG_F	RQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PT	APPA.ESFG.
CRF02_AG_G	RQANFLGKIW	PSNKG.RPGN	FPQSRPE...	P.....	SAPPAESFG.
CRF02_AG_N	RQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PT	APPA.ESFG.
CRF02_AG_S	RQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PT	APPA.ESLG.
CRF02_AG_S	RQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PT	APPA.ESFG.
CRF03_AB_R	RQANFLGRIW	PSSKG.RPGN	FPQSRPE...	PS	APP.AENFG.
CRF03_AB_R	RQANFLGKIW	PSSKG.RPGN	FPQSRPE...	PS	APP.AENFG.
CRF04_cpx_	RQANFLGRMW	PSSKG.RPGN	FLQNRPE...	PT	APPA.ECLE.
CRF04_cpx_	RQANFLGRMW	PSSKG.RPGN	FLQSRPE...	PT	APPA.ESLE.
CRF04_cpx_	RQANSLGRMW	PSSKG.RPGN	FLQSRTE...	PT	APPA.ESFE.
CRF05_DF_B	RQANFLGKVV	PSHKG.RPGN	FLQSRP....	EP	SAPPAESFR.
CRF05_DF_B	GQANFLGRVW	LSHKG.RPGN	FLQSRP....	EP	SAPPAESFG.
CRF06_cpx_	RQANFLGKIW	PSNKG.RPGN	FLQNRPE...	P	TAPPIESFG.
CRF06_cpx_	RQANFLGKIW	PSNKG.RPGN	FLQNRPE...	P	TAPPAESFG.
CRF06_cpx_	RQANFLGRIW	PSSKG.RPGN	FLQNRPE...	P	TAPPAESFG.
CRF06_cpx_	RQANFLGKIW	PSHKG.RPGN	FLQNRPEQNR	P.....EP	SAPPAESFG.
CRF11_cpx_	RQANFLGKIW	PSSKG.RPGN	FLQSRPE...	PT	APPA.ESFG.
CRF11_cpx_	RQANFLGKIW	PSSKG.RPGN	FLQSRPE...	PT	APPA.ESFG.
D_CD_84ZR0	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPAE.FG.
D_CD_ELI_K	RQANFLGRIW	PSHKG.RPGN	FLQSRP....	EP	TAPPAESFG.
D_CD_NDK_M	RQANFLGKIW	PSHKG.RPGN	FLQSRP....	EP	TAPPAESFG.
D_UG_94UG1	RQANFLGKIW	PSHNG.RPGN	FLQSRPPA..	EP	TAPPAEIFG.
F1_BE_VI85	RQANFLGKIW	PSNKG.RPGN	FLQSRPE...	P	TAPPAESFG.
F1_BR_93BR	RQANFLGKIW	PSNKG.RPGN	FIQNRPE...	P	SAPPAESFR.
F1_FI_FIN9	RQANFLGKIW	PSNKG.RPGN	FLQSRPE...	P	TAPPAESLG.
F1_FR_MP41	RQANFLGKIW	PSNKG.RPGN	FLQNRPE...	P	TAPPAESFG.
F2_CM_MP25	RQANFLGKMW	PSNKG.RPGN	FLQNRPE...	P	TAPPAESFG.
F2KU_BE_VI	RQANFLGKIW	PSNKG.RPGN	FLQSRPE...	P	TAPPAESFG.
G_BE_DRCBL	RQANFLGKIW	PSNKG.RPGN	FLQNRPE...	P	TAPPAENFG.
G_NG_92NG0	RQANFLGKIW	PSNKG.RPGN	FLQNRTE...	P	TAPPAESFG.
G_SE_SE616	RQANFLGKIW	PSNKG.RPGN	FLQNRTE...	P	TAPPAESLG.
H_BE_VI991	RQANFLGKIW	PSSKG.RPGN	FPQKRLE...	P	TAPPAESFG.
H_BE_VI997	RQANFLGKIW	PSSKG.RPGN	FLQSRPE...	P	TAPPAESFG.
H_CF_90CF0	RQANFLGKIW	PSSKG.RPGN	FLQSRPE...	P	TAPPAESFG.
J_SE_SE702	RQANFLGKIW	PSSKG.RPGN	FLQSRPE...	P	TAPPAESLG.
J_SE_SE788	RQANFLGKIW	PSSKG.RPGN	FLQSRPE...	P	TAPPAESLG.
K_CD_EQTB1	RQANFLGKFW	PLNKE.RPGN	FLQNRPE...	P	TAPPAESFG.
K_CM_MP535	RQANFLGKIW	PSHKG.RPGN	FLQSRPE...	P	TAPPAESFG.
N_CM_YBF30	RQANFLGKSW	SPFKG.RPGN	FPQTTTRK..	EP	TAPPLESYG.
O_CM_ANT70	KQANFLGKYW	PP.GGTRPGN	YVQRPAAH...	P	SAPPMEEVK
O_CM_MVP51	RQANFLGKYW	PP.GGTRPGN	YVQKQVS...	P	SAPPMEEAVK
O_SN_99SE_	RQANFLGKYW	PP.GGTRPGN	YAQRQVS...	P	SAPPMTEEMK
O_SN_99SE_	KQANFLGKYW	PP.GGTRPGN	YAQRQVS...	P	SAPPMTEEMK
U_CD___83C	RQANFLGKIW	PSNKG.RPGN	FLQNRPE...	P	TAPPAESFG.

501

550

00BW0762_1	FE...	..ETNPPTP..	...KQE....	PKDRE.....	PLTSLKSLFG
00BW0768_2	FE...	..ETTPAP..	...KQE....	LKDR.....E	PLTALKSLFG
00BW0874_2	FE...	..ETTPAL..	...KRE....	LKER.....E	PLISLKSLFG
00BW1471_2	FE...	..ETTPAP..	...KQE....	PKDR.....E	PLTSLKSLFG
00BW1616_2	F....	..GETTPSP..	...RQE....	AKDRE.....	PLISLKSLFG
00BW1686_8	FE...	..ETTPAP..	...KQE....	PKDR.....E	PLTSLKSLFG
00BW1759_3	FE...	..ETTPAP..	...KQE....	PKDRE.....	TLTSLRSLFG
00BW1773_2	FE...	..ETTPAP..	...KQE....	PKDRE.....	PLTSLKSLFG

00BW1783_5	FE...	..ETTPVQ..	...KQE....	TKDRE.....	PLTSLKSLFG
00BW1795_6	F....	..EETTPSP..	...KQE....	LKDKE.....	PLTSLKSLFG
00BW1811_3	FE...	..ETTPAS..	...KQE....	KKDRE.....	TLTSLRSLFG
00BW1859_5	FE...	..ETTPAP..	...KQE....	QKDR.....E	PLTSLKSLFG
00BW1880_2	FE...	..ETTPAP..	...KQE....	PKDRE.....	PLTSLKSLFG
00BW1921_1	FE...	..ETTPAP..	...KQE....	PKDR.....E	PLTSLKSLFG
00BW2036_1	FE...	..ETTPAP..	...KQE....	LKDR.....E	PLISLKSLFG
00BW2063_6		..EETTPAP..	...KQE....	MKDKE.....	PLISLKSLLG
00BW2087_2	FE...	..ETTPAS..	...KQD....	LKDRE.....	PLTSLKSLFG
00BW2127_2	FE...	..ETTHAP..	...KQE....	LKDRE.....	ALTSLKSLFG
00BW2128_3	FE...	..ETTPAP..	...KQE....	PKNRE.....	PLTSLKSLFG
00BW2276_7	FE...	..ETTPEL..	...KQG....	PKDR.....E	PLTSLKSLFG
00BW3819_3	FE...	..EITPAP..	...KQE....	TKDRE.....	PLTSLKSLFG
00BW3842_8	FE...	..ETTPAP..	...KQE....	PKDRGPY.RE	PLISLKSLFG
00BW3871_3	FE...	..ETTPVP..	...KQE....	PTDR.....E	PLTSLKSLFG
00BW3876_9	FE...	..ETTPTL..	...KQE....	LKDRE.....	PLTSLKSLFG
00BW3886_8	FE...	..ETTPVP..	...KQE....	QKDRE.....	ALTSLKSLFG
00BW3891_6	FE...	..EITPVP..	...KQE....	PKDR.....E	PLTSLKSLFG
00BW3970_2	FE...	..ETTPAP..	...KQE....	PKDRE.....	PLISLKSLFG
00BW5031_1	FG...	..ETTPAP..	...KQE....	MKERE.....	PLISLKSLFG
96BW01B21	FE...	..ETTPAP..	...KQE....	PKDRE.....	PLTSLRSLFG
96BW0407	FE...	..ETTPGQ..	...KQE....	SKDRE.....	TLTSLKSLFG
96BW0502	FE...	..ETTPAP..	...KQE....	PKDREPY.RE	PLTALRSLFG
96BW06_J4	FE...	..ETTPAL..	...KQE....	PKDK.....E	PLTSLKSPFG
96BW11_06		..EETTPAP..	...KQE....	TKDRE.....	PLISLKSLFG
96BW1210	FE...	..ETTPAQ..	...KQE....	PKDREP....	PLASLKSLFG
96BW15B03	FE...	..ETTPAP..	...KQE....	PKDR.....E	PLISLKSLFG
96BW16_26	FG...	..ETTPAP..	...KQE....	PKDRE.....	PLTSLRSLFG
96BW17A09	FE...	..ETTPAP..	...KQE....	PKDR.....E	PLTSFKSLFG
96BWM01_5		..EETTPAP..	...KQE....	MKDKEPY.KE	PLISLRSLFG
96BWM03_2	PE...	..PTAPPAE..	...RQE....	SKDR.....E	PLISLKSLFG
98BWMC12_2	LE...	..ETTPAS..	...KQE....	MKDR.....E	PLISLKSLFG
98BWMC13_4		..EETTPAP..	...KQE....	PKDKE.....	PLISLKSLFG
98BWMC14_a	FE...	..ETTPAP..	...KQE....	QKDR.....E	PLTSLKSLFG
98BWM014_1	FESRP	..EPTAPPAES.	...FRQE....	PKDR.....E	PLTALKSLFG
98BWM018_d	FE...	..ETTPAL..	...KQE....	PKDREA....	PLTSLKSLFG
98BWM036_a	FE...	..ETNLAP..	...KQE....	PKDRE.....	PLTSLKSLFG
98BWM037_d	FE...	..ETTPAP..	...RQE....	AKDKE.....	PLNSLKSLFG
99BW3932_1	FE...	..ETTPAP..	...KQE....	LKDRE.....	ALTSLKSLFG
99BW4642_4	FE...	..ETTPAP..	...KQE....	PKDR.....E	PLTSLKSLFG
99BW4745_8	FE...	..GATPTP..	...KQE....	PRDR.....E	PLTSLKSLFG
99BW4754_7	FE...	..ETTPAQ..	...KQE....	SKDRE.....	PLTSLKSLFG
99BWMC16_8	FE...	..ETNPAP..	...KQE....	LKNRE.....	TLTSLRSLFG
A2_CD_97CD		..EEITSSL..	...KQE....	NREPS....T	PAISLKSLFG
A2_CY_94CY		..MGEEITSSL	...KQEL...E	TREPY....N	PAISLKSLFG
A2D_97KR		..MGEEITPLQ	...KQEL...K	NREQH....T	PAISLKSLFG
A2G_CD_97C					
A_BY_97BL0		..MGEEIT....	PSLK.Q...E	QKDRE..QYP	PSISLKSLFG
A_KE_Q23_A		..MGEEITV....	SPLK.Q...E	QKDRE..QAQ	PLVSLKSLFG
A_SE_SE659					
A_SE_SE725					
A_SE_SE753		..MREEIA....	SPPK.Q...E	Q..KG..QDP	PLVSLKSLFG
A_SE_SE853					
A_SE_SE889					
A_SE_UGSE8		..MGEEIA....	SPPK.Q...E	Q.....NNP	PSVSLKSLFG
A_UG_92UG0		..MREEIV....	SPPK.Q...E	QNDRD..QNP	PSVSLKSLFG
A_UG_U455_		..MGEEKMTSPA	...KQEL...K	DREQ.....T	PLVSLKSLFG
AC_IN_2130	FE...	..ETTPAL..	...KQE....	QKDRE.....	PLTSLKSLFG
AC_RW_92RW	ENFG	..MGEEIASPL.	...K.QE....	QKDRE.....	PLISLKSLFG
AC_SE_SE94					
ACD_SE_SE8		..FGEEITP...	S.QK.Q...E	QKDKE...LY	PLASLKSLFG

ACG_BE_VI1		KEDAIDSS..	..PKQE....	PRDKG..LYP	PLTSLKSLFG
AD_SE_SE69		FGEEIAP...	.SQKQE...Q	KDK....ELY	PLASLKSLFG
AD_SE_SE71					
ADHK_NO_97		IGEEIT....	SYQK.Q...E	QKDRE..PPP	PLVSLKSLFG
ADK_CD_MAL		FGEEIK....	PSQK.Q...E	QKDKE..L.Y	PLASLKSLFG
AG_BE_VI11		MEEEIT....	PSQK.Q...E	PRDTG..LYP	PLTSLKSLFG
AG_NG_92NG		FGEEIAP...	S.LK.Q...E	PREKE..SPP	L.TSLKSLFG
AGHU_GA_VI		FGEEIA....	PSPR.P...E	PREKE..R.Y	PLTSLKSLFG
AGU_CD_Z32		TKEEITS...	S.PK.Q...E	PRDKE..LYP	PLASLKSLFG
AJ_BW_BW21		FGEEITA....	PSPK.Q...E	GKDKE..L.Y	PLTSLKSLFG
B_AU_VH_AF		FGREETTP..	.SQKQE....	PIDK...ELY	PLASLRSLFG
B_CN_RL42_		FGREETTP..	.SQKQE....	PIDK...ELY	PLASLKSLFG
B_DE_D31_U		FGREETATP..	.FQKQE....	PIDK...ELY	PLASLRSLFG
B_DE_HAN_U		FGEEATAP...	.SQKQE....	PIDK...ELY	PLASLKSLFG
B_FR_HXB2_		SGVETTP..	.PQKQE....	PIDK...ELY	PLTSLRSLFG
B_GA_OYI_		FGREETTP..	.PQKQE....	PIDK...GLY	PLTSLRSLFG
B_GB_CAM1_		FGEEKTTP..	.SQKQE....	PIDK...ELY	PLASLRSLFG
B_GB_GB8_A		FGGETTP..	.SQKQE....	PINK...EPY	PLASLRSLFG
B_GB_MANC_		FGREETTP..	.AQKQE....	PIDK...ELY	PLASLRPLFG
B_KR_WK_AF		FGREETTP..	.SQKQE....	PIDK...ELY	PLASLRSLFG
B_NL_3202A		FGREETTP..	.SQKQE....	PRDK...ELY	PLASLRSLFG
B_TW_TWCYS		FGEQTTTP..	.SQKQE....	PIDK...DLY	PLASLESFLG
B_US_BC_L0		FGREETTP..	.PQKQERE..	..DK...EMY	PLASLRSLFG
B_US_DH123		FGREETATP..	.SQKQE....	..PK...ELY	PLASLKSLFG
B_US_JRCSF		FGREETATP..	.SQKQE.QKQE	PIDK...ELY	PLTSLRSLFG
B_US_MNCG_		FGREETTP..	.YQKQEKQKQE	TIDK...DLY	PLASLKSLFG
B_US_P896_		FGREETTP..	.SQKQE....	PIDK...ELY	PLASLRSLFG
B_US_RF_M1		FGREETTP..	.SQKQE....	KIDK...ELY	PLASLKSLFG
B_US_SF2_K		FGEEKTTP..	.SQKQE....	PIDK...ELY	PLTSLRSLFG
B_US_WEAU1		FREETTP..	.SQKQE....	PIDK...ELY	PLTSLKSLFG
B_US_WR27_		FGXETTP..	.SQKQE....	PIDK...ELY	PLASLRSLFV
B_US_YU2_M		FGREETTP..	.SQKQE....	PIDK...ELY	PLASLRSLFG
BF1_BR_93B		FGEEVTTP..	.SQKQE....	PIDK...EMY	PLASLRSLFG
C_BR_92BR0	ESFR	FGREETTPS..	..RKQE....	TIDKEL....	PLTSLKSLFG
C_BW_96BW0	FE...	..ETTPVP..	...KQE....	PKDRE....	PLTSLKSLFG
C_BW_96BW1		..ETTPAP..	...KQE....	TKDRE....	PLISLKSLFG
C_BW_96BW1	FE...	..ETTPAQ..	...KQE....	PKDREP....	PLASLKSLFG
C_BW_96BW1	FE...	..ETTPAP..	...KQE....	PKDR....E	PLISLKSLFG
C_ET_ETH22	PTAPPESFR	FEEATPSPK..	..Q..E....	LKDRE....	ALTSLKSLFG
C_IN_93IN1	FE...	..ETTPAP..	...KQE....	PKDRE....	PLTSLKSLFG
C_IN_93IN9	FE...	..ETPPAP..	...KQE....	PKDRE....	PLTSLRSLFG
C_IN_93IN9	FE...	..ETTPAL..	...KQE....	PKDRE....	PLTSLKSLFG
C_IN_94IN1	FE...	..ETPPAP..	...KQE....	PKERE....	PLTSLRSLFG
C_IN_95IN2	FE...	..ETTPAP..	...KQE....	PKDRE....	PLTSLRSLFG
CRF01_AE_C		MGEETIT....	SFPK.Q...E	QKDKE..HPS	PLVSLKSLFG
CRF01_AE_C		MGEETIT....	SFPK.Q...E	QKDKK..QPP	PLVSLKSLFG
CRF01_AE_C		MGEETIT....	SFSR.Q...E	QKDRE..HPP	PLVSLKSLFG
CRF01_AE_T		MGEETT....	.SLLKQ...E	QKDKE..HHP	PLVSLKSLFG
CRF01_AE_T		MGEETIT....	SLPK.Q...E	QKDKD..PPP	.LVSLKSLFG
CRF01_AE_T		MGEE.....		QKDKE..HPP	PSVSLKSLFG
CRF01_AE_T		MGEETT....	SSLK.Q...E	QKDKE..PPP	PLISLKSLFG
CRF01_AE_T		MGEETITGEEI	TSLPKQ...E	QKDKE..HPP	PLVSLKSLFG
CRF01_AE_T		MGEETIT....	SFLK.Q...E	QKDKE..HPP	PSVSLKSLFG
CRF02_AG_F		MGEETIT....	SPPK.Q...E	ARDQG..LYP	PLASLKSLFG
CRF02_AG_F		MGEETIT....	SPPK.Q...E	PRDQG..LYP	PLASLKSLFG
CRF02_AG_G		TREEITSS..	..PQQE....	PRDKG..LYP	PLTSLKSLFG
CRF02_AG_N		MGEIIP....	PSPQ.Q...E	PRDKG..LYP	PLTSLKSLFG
CRF02_AG_S		IGEEIT....	SSQK.Q...E	PGDKG..LYP	PLASLKSLFG
CRF02_AG_S		MGEETIT....	SSPK.Q...E	PGDKG..LYP	PLTSLKSLFG
CRF03_AB_R		MGEETIT....	PSLK.Q...E	QKDRE..QHP	PSISLKSLFG
CRF03_AB_R		MGEETIT....	PSLK.Q...E	QKDRG..QHP	PSISLKSLFG

CRF04_cpx_		RKEETTS...	S.LK.Q...	E	PRDKE..LYP	.LTSLKSLFG
CRF04_cpx_		MKEETTS...	S.PK.Q...	E	PRDKE..LYP	.LTSLKSLFG
CRF04_cpx_		MKEETTS...	S.PK.Q...	E	QRDKE..LYP	.ITSLKSLFG
CRF05_DF_B		FGEEIAS...	.SPKQE...	Q	KDEG...LYP	PLASLKSLFG
CRF05_DF_B		FGEEITP...	.SPKQE...	Q	KDEG...KYP	PLASLKSLFG
CRF06_cpx_		FGEEIAP...	S.PK.Q...	E	SKEKEEKGLY	PLASLKSLFG
CRF06_cpx_		FGEEIAP...	S.PE.Q...	K	PKEKE...LY	PLTSLRSLFG
CRF06_cpx_		FGEEIAP...	S.LK.Q...	E	PKEKEKE.LY	PLASLKSLFG
CRF06_cpx_		FGEEIAP...	S.PK.Q...	E	PKEKE...LY	PLASLKSLFG
CRF11_cpx_		FGEEIAP...	.SPK.Q...	E	PKEKEK.ELY	PLTSLKSLFG
CRF11_cpx_		FGEEITP...	.SPK.Q...	E	PKEK...ELY	PITSLKSLFG
D_CD_84ZR0		FGEEITP...	.SQKQEQK..		DKDK...ELY	PLASLKSLFG
D_CD_ELI_K		FGEEITP...	.SQKQE...	Q	KDK...ELY	PLTSLKSLFG
D_CD_NDK_M		FGEEITP...	.SQKQE...	Q	KDK...ELY	PLASLKSLFG
D_UG_94UG1		LGEEITP...	.PQKQE...	Q	KDK...ELY	PLTSLKSLFG
F1_BE_VI85	FR...	.EEITPSP..	...KQE...		QKDGL..YP	PLASLKSLFG
F1_BR_93BR	FG...	.EETTPSP..	...KQE...		QKDEGL..YP	PLASLKSLFG
F1_FI_FIN9	IR...	.EEVTPSP..	...RQE...		QKEEGQ..YP	PLASLKSLFG
F1_FR_MP41	FK...	.EEITPSP..	...KQE...		QKDEGQGLYP	PLASLKSLFG
F2_CM_MP25	FG...	.EEIAPSP..	...KQE...		QKDKEQ..VP	PLISLKSLFG
F2KU_BE_VI	FG...	.EEINPSP..	...RQE...		TKDKGQ..EP	PLTSLKSLFG
G_BE_DRCBL		FGEEIAP...	S.PK.Q...	E	QKEKE..LYP	L.SSLKSLFG
G_NG_92NG0		FGEEIAP...	S.PK.Q...	E	PKEKE..LYP	L.TSLKSLFG
G_SE_SE616		FGEEIAP...	S.PK.Q...	E	MKEKE..LYP	...SLKSLFG
H_BE_VI991	FG...	.EEITPSP..	...RQE...		LKEQE....P	PLTSLRSLFG
H_BE_VI997	FG...	.EEMTSSP..	...KQE...		LKDKE....P	PFASLKSLFG
H_CF_90CF0	FG...	.EEMTPSP..	...KQEQ...		LKDKE....P	PLASLRSLFG
J_SE_SE702	FG...	.EEIPSP..	...KQE...		PKDKE...LY	PLTSLRSLFG
J_SE_SE788	LG...	.EEIPSP..	...KQE...		PKDKE...LY	PLTSLKSLFG
K_CD_EQTB1	FG...	.EKITPSL..	...RQE...		MKDQEQ..GP	PLTSLKSLFG
K_CM_MP535	FG...	.EEITPSP..	...RQE...		TKDKEQ..SP	PLTSLKSLFG
N_CM_YBF30	FQ...	.EEKSTQ..	GKEMQE...	N	QERTENSLYP	PLTSLRSLFG
O_CM_ANT70		.GQENQEQ..	...KGG...		PNE.....LY	PFASLKSLFG
O_CM_MVP51		.EQENQSQ..	...KGD...		QEE.....LY	PFASLKSLFG
O_SN_99SE_		.GQENQEQ..	...KED...		QNE.....LY	PFASLRSLFG
O_SN_99SE_		.GQENQEQ..	...KGD...		QNE.....LY	PFASLKSLFG
U_CD_83C	FG...	.EETTPSP..	...KQE...		PRDKESL.YP	PLTSLKSLFG

551

00BW0762_1	SDPLSQ
00BW0768_2	SDPLSQ
00BW0874_2	NDPLSQ
00BW1471_2	SDPLSQ
00BW1616_2	SDPLSQ
00BW1686_8	SDPLSQ
00BW1759_3	SDPLSQ
00BW1773_2	SDPLSQ
00BW1783_5	SDPLSQ
00BW1795_6	SDPLSQ
00BW1811_3	SDPLSQ
00BW1859_5	SDPLSQ
00BW1880_2	NDPLSQ
00BW1921_1	SDPLSQ
00BW2036_1	SDPLSQ
00BW2063_6	NDPLSQ
00BW2087_2	SDPLSQ
00BW2127_2	SDPLSQ
00BW2128_3	SDPWSQ
00BW2276_7	SDPLSQ
00BW3819_3	SDPLSQ
00BW3842_8	SDPLSQ

00BW3871_3	SDPLSQ
00BW3876_9	SDPLSQ
00BW3886_8	SDPLSQ
00BW3891_6	SDPLSQ
00BW3970_2	SDPLSQ
00BW5031_1	SDPLSQ
96BW01B21	SDPLSQ
96BW0407	NDPLSQ
96BW0502	SGPLSQ
96BW06_J4	SDPLSQ
96BW11_06	SDPLSQ
96BW1210	NDPLSQ
96BW15B03	SDPLSQ
96BW16_26	NDPLSQ
96BW17A09	SDPLSQ
96BWM01_5	SDPLSQ
96BWM03_2	SDPLSQ
98BWMC12_2	NDPLSQ
98BWMC13_4	SDPLSQ
98BWMC14_a	NDPLSQ
98BWM014_1	SDPLSQ
98BWM018_d	SDPLSQ
98BWM036_a	SDPLSQ
98BWM037_d	SDPLSQ
99BW3932_1	SDPLSQ
99BW4642_4	SDPLSQ
99BW4745_8	SDPLSQ
99BW4754_7	NDPLSQ
99BWMC16_8	GDPLSQ
A2_CD_97CD	NDLLSQ
A2_CY_94CY	NDPLLQ
A2D___97KR	NDPLLQ
A2G_CD_97C	
A_BY_97BL0	NDPLSQ
A_KE_Q23_A	NDLLSQ
A_SE_SE659	
A_SE_SE725	
A_SE_SE753	NDLLSQ
A_SE_SE853	
A_SE_SE889	
A_SE_UGSE8	NDLLSQ
A_UG_92UG0	NDLLSQ
A_UG_U455_	NDPLSQ
AC_IN_2130	SDPLSQ
AC_RW_92RW	NDPLSQ
AC_SE_SE94	
ACD_SE_SE8	NDP...
ACG_BE_VI1	NDP...
AD_SE_SE69	NDP...
AD_SE_SE71	
ADHK_NO_97	NDPLSQ
ADK_CD_MAL	NDQLSQ
AG_BE_VI11	NDP...
AG_NG_92NG	NDP...
AGHU_GA_VI	SDP...
AGU_CD_Z32	SDP...
AJ_BW_BW21	SDP...
B_AU_VH_AF	NDPSSQ
B_CN_RL42_	NDPSSQ
B_DE_D31_U	NDPSSQ
B_DE_HAN_U	SDPSSQ

B_FR_HXB2_	NDPSSQ
B_GA_OYI_	NDPSSQ
B_GB_CAM1_	NDPSSQ
B_GB_GB8_A	NDPSSQ
B_GB_MANC_	NDPSSQ
B_KR_WK_AF	NDPSSQ
B_NL_3202A	NDPSSQ
B_TW_TWCYS	NDPSSQ
B_US_BC_L0	NDPSSQ
B_US_DH123	NDP...
B_US_JRCSF	NDPSSQ
B_US_MNCG_	NDPLSQ
B_US_P896_	NDPSSQ
B_US_RF_M1	NDPSSQ
B_US_SF2_K	NDPSSQ
B_US_WEAU1	NDPSSQ
B_US_WR27_	NDPSSQ
B_US_YU2_M	SDPSSQ
BF1_BR_93B	NDPSSQ
C_BR_92BR0	SDPLST
C_BW_96BW0	SDPLSQ
C_BW_96BW1	SDPLSQ
C_BW_96BW1	NDPLSQ
C_BW_96BW1	SDPLSQ
C_ET_ETH22	NDHLLQ
C_IN_93IN1	SDLLSQ
C_IN_93IN9	SDPLSQ
C_IN_93IN9	SDPLSQ
C_IN_94IN1	SDPLSQ
C_IN_95IN2	SDPLSQ
CRF01_AE_C	NDPLSQ
CRF01_AE_C	NDPLSQ
CRF01_AE_C	NDPLSQ
CRF01_AE_T	NDPSSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF02_AG_F	NDP...
CRF02_AG_F	NDP...
CRF02_AG_G	NDP...
CRF02_AG_N	NDP...
CRF02_AG_S	NDP...
CRF02_AG_S	NDPYSQ
CRF03_AB_R	DDPLSQ
CRF03_AB_R	NDPLSQ
CRF04_cpx_	SDPLSQ
CRF04_cpx_	NHPLSQ
CRF04_cpx_	SDPLSR
CRF05_DF_B	NDPLSQ
CRF05_DF_B	NDPLSQ
CRF06_cpx_	SDP...
CRF06_cpx_	NDP...
CRF06_cpx_	NDP...
CRF06_cpx_	SDP...
CRF11_cpx_	SDP...
CRF11_cpx_	SDPLSQ
D_CD_84ZR0	NDPLSQ
D_CD_ELI_K	NDPLSQ
D_CD_NDK_M	NDPSSQ

D_UG_94UG1	NDPLSQ
F1_BE_VI85	NDP...
F1_BR_93BR	NDP...
F1_FI_FIN9	NDP...
F1_FR_MP41	SDP...
F2_CM_MP25	SDQ...
F2KU_BE_VI	SDPLLQ
G_BE_DRCBL	NDQ...
G_NG_92NG0	SDP...
G_SE_SE616	SDP...
H_BE_VI991	NDQ...
H_BE_VI997	NDPLSQ
H_CF_90CF0	SDPLLQ
J_SE_SE702	SDPLSQ
J_SE_SE788	SDPLSQ
K_CD_EQTB1	SDPLSQ
K_CM_MP535	NDPLSQ
N_CM_YBF30	NDPSSQ
O_CM_ANT70	TDQ...
O_CM_MVP51	TDQ...
O_SN_99SE	TDQ...
O_SN_99SE	TDQ...
U_CD_83C	SDPSLQ

Table 12. HIV Env Sequence Alignment
GCG Multiple Sequence File.
Written by Omega 1.1

Name: 00BW0762_1	Len: 962	Check: 4645	Weight: 1.00
Name: 00BW0768_2	Len: 962	Check: 9565	Weight: 1.00
Name: 00BW0874_2	Len: 962	Check: 7745	Weight: 1.00
Name: 00BW1471_2	Len: 962	Check: 9593	Weight: 1.00
Name: 00BW1616_2	Len: 962	Check: 792	Weight: 1.00
Name: 00BW1686_8	Len: 962	Check: 3744	Weight: 1.00
Name: 00BW1759_3	Len: 962	Check: 9808	Weight: 1.00
Name: 00BW1773_2	Len: 962	Check: 3500	Weight: 1.00
Name: 00BW1783_5	Len: 962	Check: 9684	Weight: 1.00
Name: 00BW1795_6	Len: 962	Check: 8410	Weight: 1.00
Name: 00BW1811_3	Len: 962	Check: 2068	Weight: 1.00
Name: 00BW1859_5	Len: 962	Check: 5692	Weight: 1.00
Name: 00BW1880_2	Len: 962	Check: 1901	Weight: 1.00
Name: 00BW1921_1	Len: 962	Check: 5923	Weight: 1.00
Name: 00BW2036_1	Len: 962	Check: 7035	Weight: 1.00
Name: 00BW2063_6	Len: 962	Check: 4853	Weight: 1.00
Name: 00BW2087_2	Len: 962	Check: 2085	Weight: 1.00
Name: 00BW2127_2	Len: 962	Check: 4015	Weight: 1.00
Name: 00BW2128_3	Len: 962	Check: 5884	Weight: 1.00
Name: 00BW2276_7	Len: 962	Check: 8913	Weight: 1.00
Name: 00BW3819_3	Len: 962	Check: 9390	Weight: 1.00
Name: 00BW3842_8	Len: 962	Check: 8867	Weight: 1.00
Name: 00BW3871_3	Len: 962	Check: 7069	Weight: 1.00
Name: 00BW3876_9	Len: 962	Check: 4761	Weight: 1.00
Name: 00BW3886_8	Len: 962	Check: 7681	Weight: 1.00
Name: 00BW3891_6	Len: 962	Check: 379	Weight: 1.00
Name: 00BW3970_2	Len: 962	Check: 8001	Weight: 1.00
Name: 00BW5031_1	Len: 962	Check: 4902	Weight: 1.00
Name: 96BW01B21	Len: 962	Check: 5774	Weight: 1.00
Name: 96BW0407	Len: 962	Check: 4260	Weight: 1.00
Name: 96BW0502	Len: 962	Check: 4658	Weight: 1.00
Name: 96BW06_J4	Len: 962	Check: 9749	Weight: 1.00
Name: 96BW11_06	Len: 962	Check: 4328	Weight: 1.00
Name: 96BW1210	Len: 962	Check: 3855	Weight: 1.00
Name: 96BW15B03	Len: 962	Check: 9133	Weight: 1.00
Name: 96BW16_26	Len: 962	Check: 5	Weight: 1.00
Name: 96BW17A09	Len: 962	Check: 6458	Weight: 1.00
Name: 96BWMO1_5	Len: 962	Check: 9487	Weight: 1.00
Name: 96BWMO3_2	Len: 962	Check: 8766	Weight: 1.00
Name: 98BWMC12_2	Len: 962	Check: 2722	Weight: 1.00
Name: 98BWMC13_4	Len: 962	Check: 2526	Weight: 1.00
Name: 98BWMC14_a	Len: 962	Check: 7761	Weight: 1.00
Name: 98BWMO14_1	Len: 962	Check: 93	Weight: 1.00
Name: 98BWMO18_d	Len: 962	Check: 279	Weight: 1.00
Name: 98BWMO36_a	Len: 962	Check: 134	Weight: 1.00
Name: 98BWMO37_d	Len: 962	Check: 9669	Weight: 1.00
Name: 99BW3932_1	Len: 962	Check: 3527	Weight: 1.00
Name: 99BW4642_4	Len: 962	Check: 1175	Weight: 1.00
Name: 99BW4745_8	Len: 962	Check: 8117	Weight: 1.00
Name: 99BW4754_7	Len: 962	Check: 5709	Weight: 1.00
Name: 99BWMC16_8	Len: 962	Check: 285	Weight: 1.00
Name: A2_CD_97CD	Len: 962	Check: 2892	Weight: 1.00
Name: A2_CY_94CY	Len: 962	Check: 8628	Weight: 1.00
Name: A2D_97KR	Len: 962	Check: 471	Weight: 1.00
Name: A2G_CD_97C	Len: 962	Check: 939	Weight: 1.00
Name: A_BY_97BL0	Len: 962	Check: 4291	Weight: 1.00
Name: A_KE_Q23_A	Len: 962	Check: 1190	Weight: 1.00

WO 2005/012502

PCT/US2004/009510

Name: A_SE_SE659	Len: 962	Check: 6674	Weight: 1.00
Name: A_SE_SE725	Len: 962	Check: 4925	Weight: 1.00
Name: A_SE_SE753	Len: 962	Check: 2482	Weight: 1.00
Name: A_SE_SE853	Len: 962	Check: 1860	Weight: 1.00
Name: A_SE_SE889	Len: 962	Check: 2102	Weight: 1.00
Name: A_SE_UGSE8	Len: 962	Check: 5063	Weight: 1.00
Name: A_UG_92UG0	Len: 962	Check: 6685	Weight: 1.00
Name: A_UG_U455_	Len: 962	Check: 8657	Weight: 1.00
Name: AC_IN_2130	Len: 962	Check: 7784	Weight: 1.00
Name: AC_RW_92RW	Len: 962	Check: 4676	Weight: 1.00
Name: AC_SE_SE94	Len: 962	Check: 2949	Weight: 1.00
Name: ACD_SE_SE8	Len: 962	Check: 1464	Weight: 1.00
Name: ACG_BE_VI1	Len: 962	Check: 2980	Weight: 1.00
Name: AD_SE_SE69	Len: 962	Check: 8959	Weight: 1.00
Name: AD_SE_SE71	Len: 962	Check: 7056	Weight: 1.00
Name: ADHK_NO_97	Len: 962	Check: 487	Weight: 1.00
Name: ADK_CD_MAL	Len: 962	Check: 2555	Weight: 1.00
Name: AG_BE_VI11	Len: 962	Check: 6342	Weight: 1.00
Name: AG_NG_92NG	Len: 962	Check: 1272	Weight: 1.00
Name: AGHU_GA_VI	Len: 962	Check: 1974	Weight: 1.00
Name: AGU_CD_Z32	Len: 962	Check: 4356	Weight: 1.00
Name: AJ_BW_BW21	Len: 962	Check: 9995	Weight: 1.00
Name: B_AU_VH_AF	Len: 962	Check: 5833	Weight: 1.00
Name: B_CN_RL42_	Len: 962	Check: 4092	Weight: 1.00
Name: B_DE_D31_U	Len: 962	Check: 5486	Weight: 1.00
Name: B_DE_HAN_U	Len: 962	Check: 3480	Weight: 1.00
Name: B_FR_HXB2_	Len: 962	Check: 6939	Weight: 1.00
Name: B_GA_OYI_	Len: 962	Check: 9780	Weight: 1.00
Name: B_GB_CAM1_	Len: 962	Check: 9716	Weight: 1.00
Name: B_GB_GB8_C	Len: 962	Check: 4180	Weight: 1.00
Name: B_GB_MANC_	Len: 962	Check: 9762	Weight: 1.00
Name: B_KR_WK_AF	Len: 962	Check: 6641	Weight: 1.00
Name: B_NL_3202A	Len: 962	Check: 7168	Weight: 1.00
Name: B_TW_TWCYS	Len: 962	Check: 3591	Weight: 1.00
Name: B_US_BC_L0	Len: 962	Check: 7266	Weight: 1.00
Name: B_US_DH123	Len: 962	Check: 6905	Weight: 1.00
Name: B_US_JRCSF	Len: 962	Check: 9381	Weight: 1.00
Name: B_US_MNCG_	Len: 962	Check: 9951	Weight: 1.00
Name: B_US_P896_	Len: 962	Check: 5855	Weight: 1.00
Name: B_US_RF_M1	Len: 962	Check: 6075	Weight: 1.00
Name: B_US_SF2_K	Len: 962	Check: 1434	Weight: 1.00
Name: B_US_WEAU1	Len: 962	Check: 5451	Weight: 1.00
Name: B_US_WR27_	Len: 962	Check: 4262	Weight: 1.00
Name: B_US_YU2_M	Len: 962	Check: 5841	Weight: 1.00
Name: BF1_BR_93B	Len: 962	Check: 5506	Weight: 1.00
Name: C_BR_92BR0	Len: 962	Check: 8769	Weight: 1.00
Name: C_BW_96BW0	Len: 962	Check: 6197	Weight: 1.00
Name: C_BW_96BW1	Len: 962	Check: 8144	Weight: 1.00
Name: C_BW_96BW1	Len: 962	Check: 1160	Weight: 1.00
Name: C_BW_96BW1	Len: 962	Check: 2736	Weight: 1.00
Name: C_ET_ETH22	Len: 962	Check: 8219	Weight: 1.00
Name: C_IN_93IN1	Len: 962	Check: 4068	Weight: 1.00
Name: C_IN_93IN9	Len: 962	Check: 3674	Weight: 1.00
Name: C_IN_93IN9	Len: 962	Check: 1581	Weight: 1.00
Name: C_IN_94IN1	Len: 962	Check: 9352	Weight: 1.00
Name: C_IN_95IN2	Len: 962	Check: 6988	Weight: 1.00
Name: CRF01_AE_C	Len: 962	Check: 8684	Weight: 1.00
Name: CRF01_AE_C	Len: 962	Check: 3342	Weight: 1.00
Name: CRF01_AE_C	Len: 962	Check: 5017	Weight: 1.00
Name: CRF01_AE_T	Len: 962	Check: 9124	Weight: 1.00
Name: CRF01_AE_T	Len: 962	Check: 2718	Weight: 1.00

WO 2005/012502

PCT/US2004/009510

Name: CRF01_AE_T	Len: 962	Check: 2104	Weight: 1.00
Name: CRF01_AE_T	Len: 962	Check: 8495	Weight: 1.00
Name: CRF01_AE_T	Len: 962	Check: 4076	Weight: 1.00
Name: CRF01_AE_T	Len: 962	Check: 948	Weight: 1.00
Name: CRF02_AG_F	Len: 962	Check: 9298	Weight: 1.00
Name: CRF02_AG_F	Len: 962	Check: 9278	Weight: 1.00
Name: CRF02_AG_G	Len: 962	Check: 4373	Weight: 1.00
Name: CRF02_AG_N	Len: 962	Check: 8955	Weight: 1.00
Name: CRF02_AG_S	Len: 962	Check: 252	Weight: 1.00
Name: CRF02_AG_S	Len: 962	Check: 5147	Weight: 1.00
Name: CRF03_AB_R	Len: 962	Check: 2239	Weight: 1.00
Name: CRF03_AB_R	Len: 962	Check: 2671	Weight: 1.00
Name: CRF04_cpx_	Len: 962	Check: 4892	Weight: 1.00
Name: CRF04_cpx_	Len: 962	Check: 8070	Weight: 1.00
Name: CRF04_cpx_	Len: 962	Check: 5453	Weight: 1.00
Name: CRF05_DF_B	Len: 962	Check: 174	Weight: 1.00
Name: CRF05_DF_B	Len: 962	Check: 2694	Weight: 1.00
Name: CRF06_cpx_	Len: 962	Check: 7351	Weight: 1.00
Name: CRF06_cpx_	Len: 962	Check: 5073	Weight: 1.00
Name: CRF06_cpx_	Len: 962	Check: 661	Weight: 1.00
Name: CRF06_cpx_	Len: 962	Check: 8440	Weight: 1.00
Name: CRF11_cpx_	Len: 962	Check: 2217	Weight: 1.00
Name: CRF11_cpx_	Len: 962	Check: 8216	Weight: 1.00
Name: D_CD_84ZR0	Len: 962	Check: 4843	Weight: 1.00
Name: D_CD_ELI_K	Len: 962	Check: 8403	Weight: 1.00
Name: D_CD_NDK_M	Len: 962	Check: 5813	Weight: 1.00
Name: D_UG_94UG1	Len: 962	Check: 9407	Weight: 1.00
Name: F1_BE_VI85	Len: 962	Check: 2982	Weight: 1.00
Name: F1_BR_93BR	Len: 962	Check: 8919	Weight: 1.00
Name: F1_FI_FIN9	Len: 962	Check: 6761	Weight: 1.00
Name: F1_FR_MP41	Len: 962	Check: 478	Weight: 1.00
Name: F2_CM_MP25	Len: 962	Check: 9292	Weight: 1.00
Name: F2KU_BE_VI	Len: 962	Check: 567	Weight: 1.00
Name: G_BE_DRCBL	Len: 962	Check: 6261	Weight: 1.00
Name: G_NG_92NG0	Len: 962	Check: 4508	Weight: 1.00
Name: G_SE_SE616	Len: 962	Check: 6733	Weight: 1.00
Name: H_BE_VI991	Len: 962	Check: 7498	Weight: 1.00
Name: H_BE_VI997	Len: 962	Check: 8345	Weight: 1.00
Name: H_CF_90CF0	Len: 962	Check: 2490	Weight: 1.00
Name: J_SE_SE702	Len: 962	Check: 4446	Weight: 1.00
Name: J_SE_SE788	Len: 962	Check: 1662	Weight: 1.00
Name: K_CD_EQTB1	Len: 962	Check: 7406	Weight: 1.00
Name: K_CM_MP535	Len: 962	Check: 512	Weight: 1.00
Name: N_CM_YBF30	Len: 962	Check: 1733	Weight: 1.00
Name: O_CM_ANT70	Len: 962	Check: 75	Weight: 1.00
Name: O_CM_MVP51	Len: 962	Check: 3290	Weight: 1.00
Name: O_SN_99SE_	Len: 962	Check: 6963	Weight: 1.00
Name: O_SN_99SE_	Len: 962	Check: 6278	Weight: 1.00
Name: U_CD_83C	Len: 962	Check: 9044	Weight: 1.00

1

50

00BW0762_1	MRVMGI	MRNC.QQWWI	WV.ILGFWML	MVCN.VIGNL	WVTVYYGVPV
00BW0768_2	MRVREI	LRNC.QQWWT	WG.SLGFWMV	MIYS.VVGEL	WVTVYYGVPV
00BW0874_2	MRAMGT	QRNC.RQWWI	WG.ILGFWML	MTCS.GVG.E	MVTVYYGVPV
00BW1471_2	MRVMGI	LRSC.QQWWI	WG.ILGFWML	MICS.VLGNL	WVTVYYGVPV
00BW1616_2	MRVMGI	QRNC.QRWWI	WG.ILGFWMI	Y..N.VVGNL	WVTVYYGVPV
00BW1686_8	MRVKGI	QRNW.PQWWI	WG.SLGFWML	MFYS.VMGNL	WVTVYYGVPV
00BW1759_3	MRVRGI	PRNW.QQWWI	WG.ILGFCMI	ITCK.VVGNL	WVTVYYGVPV
00BW1773_2	MRVREI	LRSY.QHWWM	WS.ILGLWIL	IISN.VVGNL	WVTVYYGVPV
00BW1783_5	MRVMGI	KRNC.PPWWI	WG.ILGFWML	MICN.VMGNL	WVTVYYGVPV
00BW1795_6	TRVMGI	RRNW.QQWWI	WG.ILGFWML	IICN.VMGNM	WVTVYYGVPV

00BW1811_3	MRVRGT	LKNY.QRWWI	WG.ILGLWIL	LINI.VVGNL	WVTVYYGVPV
00BW1859_5	MRVRGI	PRNW.QQWWI	WG.ILGFWIL	MICN.VVGNL	WVTVYYGVPV
00BW1880_2	MRVTGI	MRNC.QQWWI	WV.ILGFWML	MICN.VIGNL	WVTVYYGVPV
00BW1921_1	MRVRGI	QRNW.QQWWI	WG.SLGFWVM	A.CS.VVGNL	WVTVYYGVPV
00BW2036_1	MRVRGI	SRNW.QQWWI	WG.ILGFWMF	MICS.VLGNL	WVTVYYGVPV
00BW2063_6	MRVMGI	MRNW.PPWWI	WG.ILGFWML	MICN.VMGNL	WVTVYYGVPV
00BW2087_2	MRVTGM	WKNC.QQWWI	WG.ILGFWML	MICS.VLGNL	WVTVYYGVPV
00BW2127_2	MRVRGI	PRNW.QQWWI	WG.ILGF..	MIYS.MMGNL	WVTVYYGVPV
00BW2128_3	MRVRGI	LKNC.QQWWI	WI.ILGFWLL	IITN.VVGKL	WVTVYYGVPV
00BW2276_7	MRVRGI	LRNW.QQWWI	WG.ILGFWMV	MICS.VWGNL	WVTVYYGVPV
00BW3819_3	MRVRGI	LRNW.QQWWI	WG.ILSFVVL	MICS.RGEDR	WVTVYYGVPV
00BW3842_8	MRVRGI	LRNW.QQWWI	WV.ILGF..	.IYS.VAGNL	WVTVYYGVPV
00BW3871_3	MRVRGI	QRNW.QQWWI	WG.SLGFWML	MIYN.VMGSL	WVTVYYGVPV
00BW3876_9	MRVREI	LRNW.KQLWT	WG.ILGFWVL	IICS.AGGNL	WVTVYYGVPV
00BW3886_8	MRVRGI	LRIW.QQWWI	WA.SLGFWML	IICN.EKGKL	WVTVYYGVPV
00BW3891_6	MRVRGI	LRNY.QQWWI	WG.ILGFWML	MMCN.VMGDL	WVTVYYGVPV
00BW3970_2	MRVKGI	MRNC.QQWWI	WG.ILGFWML	LICN.GEGNL	WVTVYYGVPV
00BW5031_1	MRVMGT	QRNC.QQWWI	WG.ILGFWML	MIYN.VGGNL	WVTVYYGVPV
96BW01B21	MRVRGI	LRNY.PQWWI	WG.ILGFWMI	MLCN.VMGNL	WVTVYYGVPV
96BW0407	MRVMGI	QRNC.QQWWI	WG.ILGFWMI	FNGM.GSW..	VT.VYYGVPV
96BW0502	MRVMGI	LKNY.QQWWM	WG.ILGFWML	IISS.VVGNL	WVTVYYGVPV
96BW06_J4	MRVKGI	PRNW.QQWWI	WG.SLGFWII	C..S.VMGNL	WVTVYYGVPV
96BW11_06	MRVMEI	MRNC.QQWWI	WG.ILGFWML	MICN.VMGKS	WVTVYYGVPV
96BW1210	MRVRGI	LRNY.LQWWI	WG.ILGFWML	MVCS.K.ENM	WVTVYG.VPV
96BW15B03	MRVRGI	LRNW.QQWWI	WG.TLGFW..	.ICS.GLGNL	WVTVYDGVPV
96BW16_26	MIVRGI	LKTC.QQWWI	WI.ILGFWIL	IINN.VVGHL	WVTVYDGVPV
96BW17A09	MRVMGI	LRNC.QQWWI	WG.ILGFWML	MICS.VLGNL	WVTVYYGVPV
96BWM01_5	MRVMGI	KKNW.QPWWI	WG.VLGFWTL	MICS.VMGNL	WVTVYYGVPV
96BWM03_2	MRVRGT	QRNW.QRWWI	WS.ILAFWIL	INCN.GEEKL	WVTVYYGVPV
98BWMC12_2	MRVMGI	QKNC.QRWWI	WG.ILGFWMI	MSYS.VLGNL	WVTVYYGVPV
98BWMC13_4	MRVMGI	KMNW.QQWWI	WG.ILGFWML	MICS.VMGNL	WVTVYYGVPV
98BWMC14_a	MRVKGI	LRNW.LQWWI	WG.SLGFWML	C..S.VMGNM	WVTVYYGVPV
98BWM014_1	MRVMGT	LRNC.QQWWT	WG.ILGFWML	MICS.VGGNL	WVTVYYGVPV
98BWM018_d	MRVMGI	QKNC.QHWWI	WG.ILGFWML	MICN.GK.DL	WVTVYYGVPV
98BWM036_a	MRVRGI	LRNC.PQWWI	WG.ILGFWML	MTCN.MEGNL	WVTVYYGVPV
98BWM037_d	MRVRGI	LRNY.QQWWI	WG.ILGFWML	MICN.VVGNL	WVTVYYGVPV
99BW3932_1	MRVRGI	PRNW.QQWWI	WS.ILG....	.FCS.VVGQL	WVTVYYGVPV
99BW4642_4	MRVKGI	LRNW.QQWWI	WG.ILGFWML	MICN.VVGNL	WVTVYYGVPV
99BW4745_8	MRVRGI	LRDY.QQWWI	WS.ILGFW..	.ICN.GMGNL	WVTVYYGVPV
99BW4754_7	MRVMGI	KRNC.QQWWI	WG.ILGFWML	MI...CNGNL	WVTVYYGVPV
99BWMC16_8	MRVMEI	WRNC.PPWWI	WG.ILGFWML	MICN.GG.NR	WVTVYYGVPV
A2_CD_97CD	TRVMGT	QRNC.QKWE	WG.ILVFGMI	MMCK.AAD.L	WVTVYYGVPV
A2_CY_94CY	MRVMGT	QRNY.QHLWR	GG.ILILGML	IMCK.ATD.L	WVTVYYGVPV
A2D___97KR	MRVRGI	QRNY.QHLWK	WG.ILILGML	MISK.ATEDL	WVTVYYGVPV
A2G_CD_97C	MRVKGM	QRNW.QNLWK	WG.ALILGLV	IICS.ASNL	WVTVYYGVPV
A_BY_97BL0	MKARGM	XRNY.QHLWR	XG.TMLFWM	IMCK.AAEDL	VX.VYXVPV
A_KE_Q23_A	MRVMGI	QRNC.QHLLT	WG.IMILGTI	IFCS.AVENL	WVTVYYGVPV
A_SE_SE659	MRVMGI	QRNC.QHLLR	WG.TIILGLI	IICS.VADKL	WVTVYYGVPV
A_SE_SE725	MRVMGT	QMNW.QHLLR	WG.TIILGMI	MICS.TADNL	WVTVYYGVPV
A_SE_SE753	MRAMGI	QRNC.QHLLR	WG.TMILGLV	IICS.VAGNL	WVTVYYGVPV
A_SE_SE853	MRVKGI	QRNS.QHLLR	WG.TMILGMI	IICS.TADKL	WVTVYYGVPV
A_SE_SE889	MRVMGT	QMNW.QNLWR	WG.TMILGII	IICS.AAENL	WVTVYYGIPV
A_SE_UGSE8	MRVMGT	QRNC.QHLLN	WG.IMILGMI	IICS.TAENL	WVTVYYGVPV
A_UG_92UG0	MRVMGI	ERNY.PCWWT	WG.IMILGMI	IICN.TAENL	WVTVYYGVPI
A_UG_U455_	MRVMGI	QRNY.PCLWR	WG.TMILGLI	IICN..AQQL	WVTVYYGVPV
AC_IN_2130	MRVRGI	LRNY.QQWWI	WG.SLGFWML	MVCN.VVGNL	WVTVYYGVPV
AC_RW_92RW	MRVMGT	LMNY.QNLWG	WG.TMILGML	TICS.AANNL	WVTVYYGVPV
AC_SE_SE94	MRVTGT	QRSC.QPWWI	WG.ILGFWML	IICS.ATDKL	WVTVYYGVPV
ACD_SE_SE8	MRVMGI	QRNW.QHLLR	WG.TMILGMI	LICS.AVDKL	WVTVYYGVPV
ACG_BE_VI1	MRVKGI	QRNY.QQWWT	WG.SLGLWML	LICN.VMGNL	WVTVYYGVPV
AD_SE_SE69	MRVRGI	EMN.YQNLWR	WG.TLLLGLM	MT.CSVTGR	WVTVYYGVPV

AD_SE_SE71	MRVMGI	QRNC.QNLLT	WG.TMILGMI	IICS.VAENL	WVTVYYGVPV
ADHK_NO_97	MKVMGT	QRN.YPNWWR	WG.VLILGML	LICS.TTGNL	WVTVYYGVPV
ADK_CD_MAL	MRVREI	QRN.YQNWWR	WG.MMLLGML	MT.CSIAEDL	WVTVYYGVPV
AG_BE_VI11	MRVRGT	QMSWP.HLWN	GG.ILILGLV	IICS.ASNL	WVTVYYGVPV
AG_NG_92NG	MRVKGT	QRNWQ.HLWT	WW.TLILGLV	IICS.ASNL	WVTVYYGVPV
AGHU_GA_VI	MRVMET	QRN.YPRLWR	WG.TIILGML	MICN.AKENL	WITVYYGVPV
AGU_CD_Z32	MKVKGI	QRNC.QHLWK	WG.TFILGLV	IICS.AAENL	WVTVYYGVPV
AJ_BW_BW21	MRVMET	LMNCT.NLWR	WG.LMIFGML	MTCS.ATGNM	WVTVYYGVPV
B_AU_VH_AF	MKVKET	KRN.WQRLWR	WG.IMLLGML	MICS.ATEKL	WVTVYYGVPV
B_CN_RL42	MRVTGI	RKN.YQHLWR	WG.TMLLGML	MICN.AAENL	WVTVYYGVPV
B_DE_D31_U	MKVKEI	RKN.YQHLWR	WG.TMLLGML	MICS.ATEKL	WVTVYYGVPV
B_DE_HAN_U	MKVKET	RKN.YQRLWR	GG.TLLLGML	MISS.VAGNL	WVTVYYGVPV
B_FR_HXB2	MRVKEK	YQHLWRWGWR	WG.TMLLGML	MICS.ATEKL	WVTVYYGVPV
B_GA_OYI	MTARGET	RKN.YQRLWR	WG.TMLLGML	MICS.AAENL	WVTVYYGVPV
B_GB_CAM1	MRAKGI	RKN.CQRLWR	WG.TMLLGML	MICS.AADKL	WVTVYYGVPV
B_GB_GB8_C	MKAKGT	RKN.YQHLWK	WG.IMLLGML	MICS.ATEKL	WVTVYYGVPV
B_GB_MANC	MKVKEI	RKN.YQNLWR	WG.TLFLGML	MICS.AEEKL	WVTVYYGVPV
B_KR_WK_AF	MRVKGI	RKN.YQHWWR	WG.IMLLGMW	MICS.AAEKL	WVTVYYGVPV
B_NL_3202A	MKVKET	RKN.YQHLWR	WG.TMLLGML	MICS.AAEQL	WVTVYYGVPV
B_TW_TWCYS	MRVRGT	RMN.CQHLWR	WG.TMLLGML	MISS.AAENL	WVTVYYGVPV
B_US_BC_L0	MRVKEI	RKN.YQHLWR	WG.TMLFGIL	MIYS.AAGNL	WVTVYYGVPV
B_US_DH123	MRVMGI	RKN.YQHLWK	GG.TLLLGIL	MICS.AAEQL	WVTVYYGVPV
B_US_JRCSF	MRVKGI	RKN.YQHLWK	GG.ILLGTL	MICS.AVEKL	WVTVYYGVPV
B_US_MNCG	MRVKGI	RRN.YQHWWG	WG.TMLLGLL	MICS.ATEKL	WVTVYYGVPV
B_US_P896	MRVKEI	RKN.WQHLR.	GG.ILLGML	MICSAAKEKT	WVTIYYGVPV
B_US_RF_M1	MRVMEM	RKN.CQHLWK	WG.TMLLGML	MICS.AAEDL	WVTVYYGVPV
B_US_SF2_K	MKVKGT	RRN.YQHLWR	WG.TLLLGML	MICS.ATEKL	WVTVYYGVPV
B_US_WEAU1	MRVKGI	RKN.YQHLWK	WG.IMLLGIL	MICS.AAENL	WVTVYYGVPV
B_US_WR27	MRVKGI	RKN.CQHLWR	WG.IMLLGML	MICN.ATEQL	WVTVYYGVPV
B_US_YU2_M	MRATEI	RKN.YQHLWK	GG.TLLLGML	MICS.AAEQL	WVTVYYGVPV
BF1_BR_93B	MRVRGM	QRN.WQHLGK	WG.LLFLGIL	IICN..AENL	WVTVYYGVPV
C_BR_92BR0	MRVEGI	QRNW.KQWWI	WG.ILGFWMV	MIYN.VRGNL	WVTVYYGVPV
C_BW_96BW0	MRVMGI	QRNC.QQWWI	WG.ILGFWMI	INGM.GSW..	VT.VYYGVPV
C_BW_96BW1	MRVMGI	MRNC.QPWWI	WG.ILGFWML	MICN.VMGKS	WVTVYYGVPV
C_BW_96BW1	MRVRGI	LRNY.LQWWI	WG.ILGFWML	MVCS.K.ENM	WVTVYG.VPV
C_BW_96BW1	MRVRGI	LRSW.QQWWI	WG.TLGFW..	.ICS.GLGNL	WVTVYDGPV
C_ET_ETH22	MKVMGI	QRNC.QQWWI	WG.ILGFWML	MICN.GMGNL	WVTVYYGVPV
C_IN_93IN1	MRVRGT	LRNY.QQWWI	WG.VLGFWML	MICN.GGGNL	WVTVYYGVPV
C_IN_93IN9	MRVRGT	LRNY.QQWWI	WG.VLGFWML	MICN.VGGNL	WVTVYYGVPV
C_IN_93IN9	MRVRGI	LRNY.QQWWI	WG.ILGFWML	MICN.VVGNL	WVTVYYGVPV
C_IN_94IN1	MRVRGT	LRNY.QQWWI	WG.VLGFWML	MICN.GGKDL	WVTVYYGVPV
C_IN_95IN2	MRVRGI	LRNY.QQWWI	WG.VLGFWML	MICN.VVGNL	WVTVYYGVPV
CRF01_AE_C	MGVKGT	QMNW.PHLWK	WG.TLILGLV	IICS.ASDTL	WVTVYYGVPV
CRF01_AE_C	MRVKGT	RRNW.PNLWK	WG.TLILGLV	IICS.ASDNL	WVTVYYGVPV
CRF01_AE_C	MRVKGT	QMNW.PNLWK	WG.TLILGLV	IMCS.ASDNL	WVTVYG.VPV
CRF01_AE_T	MRVKET	QMNW.PNLWK	WG.TLILGLV	IICS.ASDNL	WVTVYYGVPV
CRF01_AE_T	MRVKET	QMN..PNLWK	WG.TLILGLV	IICS.ASDDL	WVTVYYGVPV
CRF01_AE_T	MRVKET	QINW.PNLWK	WG.TLILGLV	IMCS.ASNL	WVTVYYGVPV
CRF01_AE_T	MRVKET	QMSW.PNLWK	WR.TLILGLV	IICS.ASDNL	WVTVYYGVPV
CRF01_AE_T	MRVKET	QMNW.PNLWK	WG.TLILGLV	IICS.ASDNL	WVTVYYGVPV
CRF01_AE_T	MRVKET	QMNW.PNLWK	WG.TLILGLV	IICS.ASENL	WVTVYYGVPV
CRF02_AG_F	MRVMGM	QRNY.PLLWK	WG.TIIFWIM	IICN..AEKL	WVTVYYGVPV
CRF02_AG_F	MRVMGI	QRNY.PLFWK	WG.MIIFWIM	IICN..AEKL	WVTVYYGVPV
CRF02_AG_G	MRVRGM	QRNC.QNLWR	WA..HDFWIL	IICN.AAENL	WVTVYYGVPV
CRF02_AG_N	MRVMGI	QKNY.PLLWR	WG.TNIFWIM	IICN..AEQL	WVTVYYGVPV
CRF02_AG_S	MRVMGI	QKNY.PLLWR	WG.MIIFWIM	TICS..AGNL	WVTVYYGVPV
CRF02_AG_S	MRVMGI	LKSC.PPFWR	WGMIMLLWIL	IICN..AENL	WVTVYYGVPV
CRF03_AB_R	MRVKEI	RKH....LWR	WG.TLFLGML	MICS.ATENL	WVTVYYGVPV
CRF03_AB_R	MRVKEI	RKH....LWR	WG.TLLLGML	MICS.ATENL	WVTVYYGVPV
CRF04_cpx	MRVMGM	QRN.YPHLWE	WG.TLILGLV	IICS.ASNL	WVTVYYGVPV
CRF04_cpx	MRVMGI	QRN.YPHLWE	WG.TLILGLV	IMCS.ASKDM	WVTVYYGVPV

CRF04_cpx_	MTVMGT	QRN.CPRLWT	WG.TFILWL	IICS.ASNL	WVTVYGV
CRF05_DF_B	MRVRGM	QRN.WPHLGK	WG.LLFLGIL	IICS.ATDKF	WVTVYGV
CRF05_DF_B	MRVRGM	QRN.WQHLGK	WG.LLFLGIL	IICS.AADNL	WVTVYGV
CRF06_cpx_	MRVKGI	QTSWQ.HLWK	WG.TLILGLV	IICS.ASKNM	WVTVYGV
CRF06_cpx_	MRVRGI	QKNWQ.HLWK	WG.TLILGLV	IICS.ASNL	WVTVYGV
CRF06_cpx_	MRVKGI	QMNWQ.HLWK	WG.TLILGLA	IICS.ATSNL	WVTVYGV
CRF06_cpx_	MTVKGI	QRNWQ.HLWK	WG.TLILGLV	IICS.ASQNM	WVTVYGV
CRF11_cpx_	MRARGT	QKNWH.DLWR	WG.LMISGML	MICN.ATDNL	WVTVYGV
CRF11_cpx_	MRVKET	QRNWH.NLWR	WG.LMIFGML	MICN.AEK.M	WVTVYGV
D_CD_84ZR0	MRVKGI	KRN.YQPLWK	WG.IMLLGML	MMTYSAADNL	WVTVYGV
D_CD_ELI_K	MRARGI	ERN.CQNWWK	WG.IMLLGIL	MT.CSAADNL	WVTVYGV
D_CD_NDK_M	MRAREK	ERN.CQNLWK	WG.IMLLGML	MT.CSAAEDL	WVTVYGV
D_UG_94UG1	MRVRET	KRN.YQHLWK	WG.TMLLGML	MI.CSVTGKS	WVTVYGV
F1_BE_VI85	MRVRGM	QRN.WQHLGK	WG.LLFLGIL	IICN.AADNL	WVTVYGV
F1_BR_93BR	MRVRGM	QRN.WQHLGK	WG.LLFLGTL	IICN.AAENL	WVTVYGV
F1_FI_FIN9	MRVRGM	QRN.WQHLGK	WG.LLFLGML	IICK.AADDL	WVTIYGV
F1_FR_MP41	MRVRVM	QRN.WQHLGK	WG.LLFLGIL	IICS.AADNL	WVTVYGV
F2_CM_MP25	MRVREM	QRN.WQHLGR	WG.LLFLGIL	IICS.AADKL	WVTVYGV
F2KU_BE_VI	MRVRER	RRN.WQPLGR	WG.ILFLGIF	IICN.AAEDL	WVTVYGV
G_BE_DRCBL	MRVKGI	QRNWQ.HLWN	WG.ILILGLV	IICS.AEK.L	WVTVYGV
G_NG_92NG0	MRVKGI	QRNWQ.HLWK	WG.TLILGLV	IICS.ASDNL	WVTVYGV
G_SE_SE616	MRVTGI	QRNW..HLWK	WG.TLILGLV	IICS.ASNL	WVTVYGV
H_BE_VI991	TRVMET	QRN.YPSLWR	WG.TLILGML	LICS.VVGNL	WVTVYGV
H_BE_VI997	TRV	MRN.YPQWWR	GG.ILLLGML	LIYS.AAGNL	WVTVYGV
H_CF_90CF0	TRVMET	QRN.YPSLWR	WG.TLILGML	LICS.AAQNL	WVTVYGV
J_SE_SE702	TRVMET	QTSWL.SLWR	WG.LMIFGML	MICS.ARENL	WVTVYGV
J_SE_SE788	TRVMET	QKNWQ.TLWR	GG.LMIFGML	MICK.AKEDL	WVTVYGV
K_CD_EQTB1	MRAREI	QRN.WQHLGK	RG.ILFLGIL	IICS.AANNL	WVTVYGV
K_CM_MP535	MRVRGM	QRN.WQTLGN	WG.ILFLGIL	IICS.NADKL	WVTVYGV
N_CM_YBF30	MKVMGM	QSGWMGMKSG	WLLFYLLVSL	IKVIG.SEQH	WVTVYGV
O_CM_ANT70	MKAMEK	RNK...K..L	WTLYLAMALI	TPCLSLR.QL	YATVYAGVPV
O_CM_MVP51	MKVMKK	NNR...K..S	WSLYIAMALL	IPCLSYSKQL	YATVYSGVPV
O_SN_99SE_	MTVTMKVMEQ	RNR...K..L	GILCIVMALI	TPCLSYN.QH	YATVYAGVPV
O_SN_99SE_	MTVTMKVMEK	RNR...K..L	GILCMVMALI	TPCLSHN.QH	YATVYAGVPV
U_CD__83C	MRVKEI	QRN.YQHLWK	WS.LIILGMI	MICK.AIEKS	WVTVYGV

	51				100
00BW0762_1	WREAKTTLFC	ASDAKAYDRE	VHNVWATHAC	VPTDPNPQEL	VLENVTENFN
00BW0768_2	WKEAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEI	VLGNVTENFN
00BW0874_2	WKEAKTTLFC	ASDAKAYERE	VHNVWATHAC	VPTDPDPQEM	VLENVTENFN
00BW1471_2	WREAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEM	YLVNVTENFN
00BW1616_2	WKEAKTTLFC	ASDAKAYDRE	VHNVWATHAC	VPTDPNPQEI	GLENVTENFN
00BW1686_8	WKEAKTTLFC	ASDAKAYEKE	VHNIWATHAC	VPTDPNPQEI	VLENVTENFN
00BW1759_3	WRETKTTLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPNPQEL	VLGNVTESFN
00BW1773_2	WKEAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEI	PLKNVTENFN
00BW1783_5	WKEAKTTLFC	ASDAKAYEKE	AHNIWATHAC	VPTDPNPQEM	FLENVTONFN
00BW1795_6	WREAKAPLFC	ASDAKAYDRE	VHNVWATHAC	VPTDPNPQEM	VLKNVTENFN
00BW1811_3	WKEAKTTLFC	ASDAKGYDRE	VHNVWATHAC	VPTDPNPQEL	VLGNVTENFN
00BW1859_5	WKEAKTTLFC	ASDAKAYERE	VHNVWATHAC	VPTDPNPQEI	VLENVTENFN
00BW1880_2	WKEAKATLFC	ASEAKAYESE	VHNVWATHAC	VPTDPNPQEI	VLENVTENFN
00BW1921_1	WKEAKTTLFC	ASDAKAYETE	VHNVWATHAC	VPTDPNPQEM	ALENVTENFN
00BW2036_1	WREAKTTLFC	ASDAKAYETE	VHNIWATHAC	VPTDPNPQEI	VLGNVTENFN
00BW2063_6	WREAKATLFC	ASDAKAYERE	VHNVWATHAC	VPTDPNPQEI	VLENVTENFN
00BW2087_2	WKEAKTTLFC	ASDAKAYERE	VHNVWATHAC	VPTDPNPQEM	ELKNVTENFN
00BW2127_2	WKEAKAPLFC	ASDAKAYEKE	AHNVWATHAC	VPTDPNPQEI	ELKNVTENFN
00BW2128_3	WKEAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEM	VLENVTENFN
00BW2276_7	WKEAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEL	VLENVTENFN
00BW3819_3	WREAKATLFC	ASDAKAHERE	VHNVWATHAC	VPTDPNPQEM	VMENVTENFN
00BW3842_8	WKEAKTTLFC	ASDAKGYETE	VHNVWATHAC	VPTDPDPQEM	VLGNVTENFN
00BW3871_3	WREAKTTLFC	ASDAKAYERE	VHNVWATHAC	VPTDPNPQEM	LLKNVTENFN
00BW3876_9	WKEAKTTLFC	ASDAKVYEKE	VHNVWATHAC	VPTDPNPQEM	VLDNVTENFN
00BW3886_8	WKEAKTTLFC	ASDAKAYERE	VHNVWATHAC	VPTDPNPQEI	VLGGVTENFN
00BW3891_6	WREAKTTLFC	ASDAKGYEKE	VHNVWATHAC	VPTDPDPQEM	VLENVTENFN
00BW3970_2	WKEAKTTLFC	ASDAKGYERE	VHNIWATHAC	VPTDPNPQEM	FLHNVTENFN
00BW5031_1	WKEAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPSPQEI	VLENVTETFN
96BW01B21	WKEAKTTLFC	ASDAKAYEKG	VHNVWATHAC	VPTDPNPQEV	FLENVTENFN
96BW0407	WKEAKATLFC	ASDARAYDRE	VHNVWATHAC	VPTDPNPQEV	ILENVTENFN
96BW0502	WKEAKTTLFC	TSDAKAYETE	VHNVWATHAC	VPTDPNPQEI	VLENVTENFN
96BW06_J4	WKEAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPSPQEI	VLGNVTENFN
96BW11_06	WREAKATLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEV	VLENVTENFN
96BW1210	WKEAKTTLFC	ASDAKAYEGE	VHNVWATHAC	VPTDPNPQEL	VLGNVTENFN
96BW15B03	WREASNTLFC	ASYAKAYEKE	VHNVWATHAC	VPTDPNPQEI	ELDNVTENFN
96BW16_26	WKEAKTTLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPNPQEI	ILKNVTENFN
96BW17A09	WREAKTTLFC	ASDAKAFESE	VHNVWATHAC	VPTDPNPQEM	VLENVTENFN
96BWM01_5	WREAKTTLFC	ASDAKAYEAE	VHNVWATHAC	VPTDPNPQEI	ELKNVTENFN
96BWM03_2	WKEAKTTLFC	ASDAKAYERE	VHNVWATHAC	VPTDPNPQEM	VLENVTENFN
98BWMC12_2	WREAKTTLFC	ASDAKAYERE	VHNIWATHAC	VPTDPNPQEM	VLENVTENFN
98BWMC13_4	WKEAKAPLFC	ASDAKVYEKE	VHNVWATHAC	VPTDPNPQEL	VLENVTENFN
98BWMC14_a	WREATTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEI	VLGNVTENFN
98BWM014_1	WKEAKTTLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPSPQEM	FLANVTENFN
98BWM018_d	WREAKATLFC	ASNAKAYEKE	VHNVWATHAC	VPTDPNPQEM	VLENVTENFN
98BWM036_a	WKEAKATLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPDPQEI	VLENVTESFN
98BWM037_d	WKEAKTTLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPNPQEM	VLENVTENFN
99BW3932_1	WKEAKATLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQKL	VLGNVTENFN
99BW4642_4	WKEAKTTLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPNPQEI	VLENVTENFN
99BW4745_8	WREAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEL	VLKNVTENFN
99BW4754_7	WREAKTTLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPNPQEI	VLENVTENFN
99BWMC16_8	WREAKATLFC	ASDAKAYERE	VHNVWATHAC	VPTDPDPQEI	ALENVTENFN
A2_CD_97CD	WRDADTTLFC	ASDAKAYATE	KHNVWATHAC	VPTDPNPQEV	NLANVTEDFN
A2_CY_94CY	WKDADTILFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEI	NLENVTENFN
A2D_97KR	WRDAETTLFC	ASDAKAYDTE	AHNVWATHAC	VPTDPNPQEI	NLENVTENFN
A2G_CD_97C	WEDANTPLFC	ASDAKSYSTE	RHNVWATHAC	VPTDPNPQEM	ILENVTESFN
A_BY_97BL0	XXDAATTLFC	ASDAKAXDKE	VHNVWATHAC	VPTDPDPQEI	ILGNVTEKFD
A_KE_Q23_A	WRDADTTLFC	ASDAKAYETE	KHNVWATHAC	VPTDPNPQEI	HLDNVTEKFN
A_SE_SE659	WKDAETTLFC	ASDAKAYDPE	VHNVWATHAC	VPTDPNPQEM	HLENVTEESN
A_SE_SE725	WKDAETTLFC	ASDAQAYKTE	MHNVWATHAC	VPTDPNPQEL	HLKNVTEEFN
A_SE_SE753	WKDAETTLFC	ASDAKAYETE	KHNVWATHAC	VPTDPNPQEI	HLKNVTEKFN

A_SE_SE853	WKDAETTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEL	HLANVTEEFN
A_SE_SE889	WRDAETTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEI	NLENVTEEFN
A_SE_UGSE8	WKDAETTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEL	YLENVTEDFN
A_UG_92UG0	WKDANTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPSPQEL	KMENVTEEFN
A_UG_U455_	WKDAVTTLFC	ASDAKAYDAE	VHNVWATHAC	VPTDPNPQEI	DLVNVTEEFN
AC_IN_2130	WKDAETTLFC	ASDAKAYETE	KHNVWATHAC	VPTDPNPQEI	HLENVTEDFN
AC_RW_92RW	WKDAETTLFC	ASDAKAYDPE	KHNVWATHAC	VPIDPDPQEI	HLENVTEEFN
AC_SE_SE94	WKEAKTTLFC	ASDAKAYEAE	VHNVWATHAC	VPTDPNPHEI	NLENVTENFN
ACD_SE_SE8	WKDAETTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	LLGNVTEDFN
ACG_BE_VI1	WKEAKTTLFC	ASDAKAYETE	AHNVWATHAC	VPTDPSPQEI	ELKNVTENFN
AD_SE_SE69	WRDAETTLFC	ASDAKAYDAE	VHNVWATHAC	VPTDPNPQEI	NLENVTEEFN
AD_SE_SE71	WKDAETTLFC	ASDAKAYETE	KHNVWATHAC	VPTDPNPQEI	HLENVTENFN
ADHK_NO_97	WKEAKTTLFC	ASDAKAYDTE	MHNVWATHAC	VPTDPSPQEI	VLENVTENFN
ADK_CD_MAL	WKEATTTLFC	ASDAKSYETE	VHNIWATHAC	VPTDPNPQEI	ELENVTEGFM
AG_BE_VI11	WEDADTTLFC	ASDAKAYSTE	SHNVWATHAC	VPTDSNPQEI	PLENVTENFN
AG_NG_92NG	WEDADTPLFC	ASDAKAYSTE	RHNVWATHAC	VPTDPNPQEI	TLENVTETFN
AGHU_GA_VI	WRDAKTSLFC	ASDAKSYSTE	SHNVWATHAC	VPTDPNPQEI	NLENVTENFN
AGU_CD_Z32	WKDAETTLFC	ASDAKAYDTE	KHNVWATHAC	VPTDPNPQEL	SLGNVTEKFN
AJ_BW_BW21	WKEAKTTLFC	ASNAKAYSTE	GHNIWATHAC	VPTDPNPQEI	ILENVTENFN
B_AU_VH_AF	WKEATTTLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPNPQEI	LLENVTEEFN
B_CN_RL42_	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	VLGNVTENFN
B_DE_D31_U	WKEATTTLFC	ASDAKAYDKE	VHNVWATHAC	VPTDPDPQEV	VLENVTEDFN
B_DE_HAN_U	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	VMGNVTENFN
B_FR_HXB2_	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	VLVNVTENFN
B_GA_OYI_	WKEATTTLFC	ASDARAYATE	VHNVWATHAC	VPTDPNPQEV	VLGNVTENFD
B_GB_CAM1_	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	VLENVTENFN
B_GB_GB8_C	WKEATTTLFC	ASDAKAYDTE	KHNVWATHAC	VPTDPNPQEV	VLGNVTENFN
B_GB_MANC_	WKEATTTLFC	ASDAKAHGTE	VHNVWATHAC	VPTDPNPQEI	VLENVTEYFD
B_KR_WK_AF	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPIDPNPQEV	FLENVTENFN
B_NL_3202A	WKEATTTLFC	ASDAKASDTE	VHNVWATHAC	VPTDPNPQEV	ALENVTEDFD
B_TW_TWCYS	WKEATTTLFC	ASDAKTYDTE	VHNVWATHAC	VPTDPNPQEV	TLGNVTENFN
B_US_BC_L0	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	IPIDPNPQEI	VLENVTEDFN
B_US_DH123	WKEANTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEI	LLENVTEDFN
B_US_JRCSE	WKETTTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	VLENVTEDFN
B_US_MNCG_	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	ELVNVTENFN
B_US_P896_	WREATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	VLGNVTENFN
B_US_RF_M1	WKEATTTLFC	ASEAKAYKTE	VHNVWATHAC	VPTDPNPQEV	LLENVTENFN
B_US_SF2_K	WKEATTTLFC	ASDARAYDTE	VHNVWATHAC	VPTDPNPQEV	VLGNVTENFN
B_US_WEAU1	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	VLENVTENFN
B_US_WR27_	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	RLENVTEDFN
B_US_YU2_M	WKEATTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEV	KLENVTENFN
BF1_BR_93B	WKEATTTLFC	ASDAKAYEKE	AHNVWATHAC	VPTDPNPQEV	VLENVTENFD
C_BR_92BR0	WKEAKTTLFC	ASDAKAYDAE	VHNVWATHAC	VPTDPNPQEM	VLENVTENFN
C_BW_96BW0	WKEAKATLFC	ASDARAYDRE	VHNVWATHAC	VPTDPNPQEV	NLENVTENFN
C_BW_96BW1	WREAKATLFC	ASDAKAYERE	VHNVWATHAC	VPTDPNPQEV	VLENVTENFN
C_BW_96BW1	WKEAKTTLFC	ASDAKAYEGE	VHNVWATHAC	VPTDPNPQEL	VLGNVTENFN
C_BW_96BW1	WREASNTLFC	ASYAKAYEKE	VHNVWATHAC	VPTDPNPQEI	ELDNVTENFN
C_ET_ETH22	WKDASPTLFC	ASDAKAYDTE	VHNVWGTFAC	VPTDPSPQEL	GLENVTENFN
C_IN_93IN1	WKEAKTTLFC	ASDAKAYERE	VHNVWATHAC	VPTDPNPQEI	VLGNVTENFN
C_IN_93IN9	WKEAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEI	PLGNVTENFN
C_IN_93IN9	WKEAKTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEM	VLENVTENFN
C_IN_94IN1	WKEAKTTLFC	ASDAKAYGKE	VHNVWATHAC	VPTDPNPQEI	SLENVTENFN
C_IN_95IN2	WKEANTTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEI	VMENVTENFN
CRF01_AE_C	WRDADTTLFC	ASDANAQETE	VHNVWATHVC	VPTDPNPQEI	HLENVTENFD
CRF01_AE_C	WRDADTTLFC	ASDAKAHVTE	VHNVWATHAC	VPTDPNPQEI	YLENVTENFD
CRF01_AE_C	WRDADTTLFC	ASDAKAHETE	VHNIWATHAC	VPTDPNPQEI	DLENVTENFN
CRF01_AE_T	WKDADTTLFC	ASDAKAHETE	VHNVWTTHAC	VPTDPNPQEI	HLENVTENFN
CRF01_AE_T	WRDADTTLFC	ASDAKAHETE	VHNVWATHAC	VPTDPNPQEI	HLENVTENFN
CRF01_AE_T	WRDADTTLFC	ASDAKAHETE	VHNVWATHAC	IPIDPNPQEM	HLENVTENFN
CRF01_AE_T	WRDADTTLFC	ASDAKAHEAE	VHNVWATHAC	VPTDPNPQEI	HLENVTENFN
CRF01_AE_T	WRDADTTLFC	ASDAKAHETE	VHNVWATHAC	VPTDPNPQEI	HLENVTENFN

CRF01_AE_T	WRDADTTLFC	ASDAKAHETE	VHNVWATHAC	VPTDPNPQEI	HMENVTENFN
CRF02_AG_F	WRNAETTLFC	ASDAKAYDAE	VHNVWATHAC	VPTDPNPQEI	HLKNVTEKFN
CRF02_AG_F	WRDAETTLFC	ASDAKAYDVE	VHNVWATHAC	VPTDPNPQEI	HLKNVTEKFN
CRF02_AG_G	WKTADTTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEI	HLENVTEEFN
CRF02_AG_N	WKTAETTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEI	HLENVTEKFN
CRF02_AG_S	WRDAETTLFC	ASDAKAYDTE	VHNVWATHAC	VPTDPNPQEI	HLGNVTEDFN
CRF02_AG_S	RRDAETTLFC	ASDAKAYETE	VHNVWATHAC	VPTDPSPQEM	HLENVTENFN
CRF03_AB_R	WKEATTTLFC	ASDAKAYSKE	VHNVWATYAC	VPTDPSPQEI	PLKNVTENFN
CRF03_AB_R	WKEATTTLFC	ASDAKAYSKE	VHNVWATYAC	VPTDPSPQEI	PLENVTENFN
CRF04_cpx_	WRDAETTLFC	ASEAKAYEKE	VHNIWATHAC	VPTDPNPQEV	ALINVTENFN
CRF04_cpx_	WRDAETKPXC	ASDAKAYDKE	IHNIWATHAC	VPTDPNPQEL	ALTNVTENFN
CRF04_cpx_	WRDAETSPFC	ASDAKAYDKE	VHNIWATHAC	VPTDPNPQEI	FLKNVTEDFN
CRF05_DF_B	WREAKTTLFC	ASDAKGYDKE	VHNVWATHAC	VPTDPNPQEV	VLGNVTENFN
CRF05_DF_B	WKEANTTLFC	ASDAKGYEKE	AHNVWATHAC	VPTDPNPQEL	ALENVTENFN
CRF06_cpx_	WEDADTTLFC	ASDAKAYSPD	KHNVWATHAC	VPTDPNPQEI	SLKNVTENFN
CRF06_cpx_	WEDADTTLFC	ASDAKAYSAE	KHNVWATHAC	VPTDPNPQEI	PLKNVTENFN
CRF06_cpx_	WEDADTTLFC	ASDAKAYSAE	KHNVWATHAC	VPTDPNPQEI	PLENVTENFN
CRF06_cpx_	WEDADTTLFC	ASDAKAYSAE	KHNVWATHAC	VPTDPNPQEI	KLENVTENFN
CRF11_cpx_	WKDANTTLFC	ASDAQAYSPE	KHNVWATHYC	VPTDPNPQEI	LLGNVTENFN
CRF11_cpx_	WRDADTTLFC	ASDARTYSTE	KHNVWATHSC	VPTDPNPREL	SENVVTENFN
D_CD_84ZR0	WKEATTTLFC	ASDAKSYKTE	AHNIWATHAC	VPTDPNPQEI	ELKNVTENFN
D_CD_ELI_K	WKEATTTLFC	ASDAKSYETE	AHNIWATHAC	VPTDPNPQEI	ALENVTENFN
D_CD_NDK_M	WKEATTTLFC	ASDAKAYKKE	AHNIWATHAC	VPTDPNPQEI	ELENVTENFN
D_UG_94UG1	WKEATTTLFC	ASDAKAYKAE	AHNIWATHAC	VPTDPNPQEI	KLENVTENFN
F1_BE_VI85	WKEATTTLFC	ASDAKAYERE	AHNVWATHAC	VPTDPNPQEV	FLKNVTENFD
F1_BR_93BR	WKEATTTLFC	ASDAKSYEKE	AHNVWATHAC	VPTDPNPQEV	VLENVTERFN
F1_FI_FIN9	WKEANTTLFC	ASDAKSYEKE	VHNVWATHAC	VPTDPNPQEV	ALN.VTENFN
F1_FR_MP41	WKEATTTLFC	ASDAKGYERE	VHNVWATHAC	VPTDPNPQEI	WLKNVTENFD
F2_CM_MP25	WKEATTTLFC	ASDAKAYERE	VHNVWATYAC	VPTDPSPQEL	VLGNVSEKFN
F2KU_BE_VI	WKEANTTLFC	ASDAKPYDTE	VHNVWVTHAC	VPTDPNPQEV	FLQNVTENFN
G_BE_DRCBL	WEDANAPLFC	ASDAKAHSTE	SHNIWATHAC	VPTDPSPQEI	NMRNVTENFN
G_NG_92NG0	WEDADTPLFC	ASDAKSYSSE	KHNVWATHAC	VPTDPNPQEI	AIENVTENFN
G_SE_SE616	WEDADTTLFC	ASDAKSYSAE	SHNVWATHAC	VPTDPNPQEI	IMENVTEYFN
H_BE_VI991	WKEAKTTLFC	ASDAKAYDTE	RHNVWATHAC	VPTDPNPQEM	VLENVTETFN
H_BE_VI997	WKEAKTTLFC	ASDAKAYEPE	KHNVWATHAC	VPTDPSPQEM	VLANVTENFN
H_CF_90CF0	WKEAKTTLFC	ASDAKAYETE	KHNVWATHAC	VPTDPNPQEM	VMENVTESFN
J_SE_SE702	WRDAKTTLFC	ASDAKAYSTE	KHNVWATHAC	VPTDPNPQEM	SLPNVTENFN
J_SE_SE788	WKDAKTTLFC	ASDAKAYSTE	KHNVWATHAC	VPTDPSPQEM	NLPNVTENFN
K_CD_EQTB1	WKEATTTLFC	ASDAKAYETE	VHNVWATHAC	VPTDPNPQEV	VLENVTENFN
K_CM_MP535	WKEATPTLFC	ASDAKAYEKE	VHNVWATHAC	VPTDPNPQEV	EMENVTENFN
N_CM_YBF30	WREAETTLFC	ASDAKAHSTE	AHNIWATQAC	VPTDPNPQEV	LLPNVTEKFN
O_CM_ANT70	WEDATPVLFC	ASDANLTSTE	KHNIWASQAC	VPTDPTPYEY	PLHNVTDDEFN
O_CM_MVP51	WEEAAPVLFC	ASDANLTSTE	QHNIWASQAC	VPTDPNPHEF	PLGNVTDNFD
O_SN_99SE_	WEEATPVLFC	ASDANLTSTE	QHNIWASQAC	VPTDPSPYEY	PLTKVTDNFN
O_SN_99SE_	WEEATPVLFC	ASDVNLTSTE	QHNIWASQAC	VPTDPSPYEY	PLKNVTDNFN
U_CD_83C	WKDAETTLFC	ASDAKAYEKE	SHNVWATHAC	VPTDPSPQEL	VLGNVTENFN

	101				150
00BW0762_1	MWKNYMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCINV	TN.....
00BW0768_2	MWKNDMVDQM	HEDVISLWDQ	SLKPCVKLTP	LCITLNCTSI	NG.....
00BW0874_2	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCTNY	NN.....
00BW1471_2	MWKNDMVDQM	HEDIISIWDQ	SLKPCVKLTP	LCVTLYCTNV	TKR.....
00BW1616_2	MWKNDMVNQM	HEDIISLWDQ	SLKPCVRLTP	LCVTLNCNNV	TT.....
00BW1686_8	MWKNDVVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLQCSNV	S.....
00BW1759_3	MWKNDMVDQM	HEDIISLWDQ	SLNPCVKLTP	LCVTLKCSNV	N.....
00BW1773_2	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLECEDA	N.....
00BW1783_5	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLTCSNL	NI.....
00BW1795_6	MWKNDMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLDCCTNA	TI.....
00BW1811_3	MWKNDMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCINA	TN.....
00BW1859_5	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLTCKDY	N.....
00BW1880_2	MWENGMDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCSNA	KA.....
00BW1921_1	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCSDA	K.....
00BW2036_1	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCTNA	NV.....
00BW2063_6	MWENDMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCRNA	NN.....
00BW2087_2	MWKNDMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCTV	NCT.....
00BW2127_2	MWKNDMVEQM	HEDIIRLWDE	SLKPCVRLTP	LCVTLRCSNA	GSG.....
00BW2128_3	MWKNDMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLSCNAT	N.....
00BW2276_7	MWKNDMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCSIV	N.....
00BW3819_3	MWKNDMVDQM	HEDIISLWDE	SLKPCVKLTQ	LCVTLECSNV	N.....
00BW3842_8	MWENDMVDQM	HEDVISLWDE	SLKPCVKLTP	LCVTLNCTNY	NG.....
00BW3871_3	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLMCSNV	T.....
00BW3876_9	MWENDMVDQM	HEDIISLWDQ	SPKPCVKLTP	LCVTLKCTDA	T.....
00BW3886_8	MWKNDMVDQM	HEDVISLWDE	SLKPCVKLTP	LCVTLKCGNV	NN.....
00BW3891_6	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCVSI	N.....
00BW3970_2	MWKNDMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLECKNV	TTN.....
00BW5031_1	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCSAA	NN.....
96BW01B21	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCSTY	N.....
96BW0407	MWKNDMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCTKV	NGT.....
96BW0502	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCRNV	N.....
96BW06_J4	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTS	LCVTLHCSNV	N.....
96BW11_06	MWENDMVNQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT..	
96BW1210	MWKNDMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCSNN	VTR.....
96BW15B03	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCTNY	ST.....
96BW16_26	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLDCCTNV	T.....
96BW17A09	MWKNDMVDQM	HEDIISLRDQ	SLKPCVKLTP	LCVTLNCTNA	TN.....
96BWM01_5	MWENDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCKNV	TS.....
96BWM03_2	MWKNDMADQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCTEA	KV.....
98BWMC12_2	MWKNDMVDQM	HEDIIRLWDQ	SLKPCVKMTP	LCVTLNCFANF	NAS.....
98BWMC13_4	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCSNV	TV.....
98BWMC14_a	MWRNDMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLDCCTNV	T.....
98BWM014_1	MWENDMVDQM	HQDIISLWDE	SLKPCVKLTP	LCVTLNCRNA	NLN.....
98BWM018_d	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLTCTNA	TKNVTN....
98BWM036_a	MWKNDMVDQM	HEDVISIWDQ	SLKPCVKLTP	LCVTLNCSNV	N.....
98BWM037_d	MRDNDMVDQM	HEDIINLWDQ	SLKPCVRLTP	LCVTLNCKDA	SVN.....
99BW3932_1	MWKNDMVDQM	HEDMIRLWDQ	SLKPCVKLTP	LCVTLKCREV	N.....
99BW4642_4	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCTNV	N.....
99BW4745_8	MWKNDMVDQM	HEDVISLWDQ	SLKPCVKLTP	LCVTLICSNN	I.....
99BW4754_7	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCKNV	TV.....
99BWMC16_8	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCKNV	TKNVTK....
A2_CD_97CD	MWKNNMVEQM	HADIISLWDQ	SLKPCVKLTP	LCVTLNCSNA	NTTNT.....
A2_CY_94CY	MWKNNMVEQM	QEDIISLWDQ	SLKPCVKLTP	LCVILNCSNA	NTSTH.....
A2D_97KR	MWKNGMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCSRV	KNTIS.....
A2G_CD_97C	MWKNDMVEQM	HVDIISLWDQ	SLKPCVKLTP	FCVTLNCTNA	TFPNA.....
A_BY_97BL0	MXKNNXVEQM	QTDIISL.DQ	SLKPCVKLTP	LCVTLNCAEP	NSTRS.....
A_KE_Q23_A	MWKNNMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLHCTNV	TSV.....
A_SE_SE659	MWKNNMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLNCTNV	
A_SE_SE725	MWKNSMVEQM	HTDIISLWDE	SLKPCVKLTP	LCVTLNCTNA	
A_SE_SE753	MWKNYMVEQM	HTDIISLWDQ	SLEPCVKLTP	LCVTLECHYN	ITV.....

A_SE_SE853	MWKNSMVEQM	HTDIISLWDQ	SLIPCVKLTP	LCVTLECN DY	NYN.....
A_SE_SE889	MWKNNMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLNCS SV	TN.....
A_SE_UGSE8	MWKNNMVEQM	HTDIISLWDQ	SLKPCVQLTP	LCVTLNCS NN	VT A.....
A_UG_92UG0	MWKNNMVEQM	HTDIISLWDQ	SLKPCVQLTP	LCVTLD CSYN	ITN.....
A_UG_U455	MWKNNMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLDCH NI	TIN.....
AC_IN_2130	MWKNSMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLNCS NV	NG.....
AC_RW_92RW	MWKNNMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLECN NI	TNVNN.....
AC_SE_SE94	IWKNYMVEQM	HQDIISLWDQ	SLKPCVKLTP	LCVTLNCR DA	TV.....
ACD_SE_SE8	MWKNNMVEQM	HTDIISLWDQ	SLQPCVKLTP	LCVTLNCT NV	TIT.....
ACG_BE_VI1	MWKNDMVDQM	HQDIISLWDE	SLKPCVKLTP	LCVTLNCS NV	TAIN.....
AD_SE_SE69	MWKNNMVEQM	HTDIISLWDQ	SLKPCVQLTP	LCVTLNCS NV	TNKIN.....
AD_SE_SE71	MWKNNMVKQM	HTDIISLWDQ	SLQPCVKLTP	LCVTLHCND T	..N.....
ADHK_NO_97	MWENNMVDQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLNCT DP	AN.....
ADK_CD_MAL	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT NV	NGTAVNG.TN
AG_BE_VI11	MWKNNMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLTCT NV	NCTNN.....
AG_NG_92NG	MWKNNMVEQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCT NV	NCNSN...VT
AGHU_GA_VI	MWKNNMVEQM	HTDIISLWDQ	SLKPCVQITP	LCVTLECSKI	N.....
AGU_CD_Z32	MWKNNMVEQM	HEDVISLWDQ	SLKPCVKLTP	LCVTLS CS DI	R.....
AJ_BW_BW21	IWKNDMVEQM	QEDIISVWDE	SLKPCVKLTP	LCVTLNCT NA	TVSNT.....
B_AU_VH_AF	MWKNNMVEQM	HEDIISLWDQ	SLKPCVQLTP	LCVTLNCT DE	LT.....
B_CN_RL42	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT NL	K.....
B_DE_D31_U	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT DL	K.....
B_DE_HAN_U	MWKNDMVEQM	QEDIISLWDQ	SLKPCVKLTP	LCVTLKCT DY	N.....
B_FR_HXB2	MWKNDMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVSLKCT DL	K.....
B_GA_OYI	MWKNNMVEQM	QEDIISLWDQ	SLKPCVKLTP	LCVTLDCT DV	NTTSSS....
B_GB_CAM1	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLIC TNV	NN.....
B_GB_GB8_C	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT DL	R.....
B_GB_MANC	MWKNNMVEQM	QEDVISLWDQ	SLKPCVKLTP	LCVTLDCT DY	VG.....
B_KR_WK_AF	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCT DL	NDTNTN....
B_NL_3202A	MWKNNMVEQM	HEDIINLWDQ	SLKPCVKLTP	LCVTLNCT DF	G.....
B_TW_TWCYS	MWKNNMADQM	QEDIISLWDE	SLKPCVELTP	LCVTLKCN DT	
B_US_BC_L0	MWKNNMVEQM	QEDIISLWDQ	SLKPCVKLTP	LCVTLNCT DE	LKNA.....
B_US_DH123	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLHCT DL	K.....
B_US_JRCSE	MWKNNMVEQM	QEDVINLWDQ	SLKPCVKLTP	LCVTLNCK DV	
B_US_MNCG	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT DL	R.....
B_US_P896	MWKNNMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCT NL	
B_US_RF_M1	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT DA	NLN.....
B_US_SF2_K	MWKNNMVEQM	QEDIISLWDQ	SLKPCVKLTP	LCVTLNCT DL	G.....
B_US_WEAU1	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT NV	NVTN.....
B_US_WR27	MWKNNMXEQM	HEDIIXLWDQ	SLKPCVKLTP	LCVTLNCT DV	
B_US_YU2_M	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT DL	
BF1_BR_93B	MWKNNMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLRCS NA	TT.....
C_BR_92BR0	MWENDMVEQM	HQDIISLWDQ	SLKPCVKLTP	LCVTLHCS NR	T.....
C_BW_96BW0	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT EV	NGTSDSS...
C_BW_96BW1	MWENDMVNQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCT NV	TV.....
C_BW_96BW1	MWKNDMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCS NN	VTR.....
C_BW_96BW1	MWKNDMVDQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLKCT NY	ST.....
C_ET_ETH22	MWKNDMVEQM	HQDIISLWDQ	GLKPCVKLTP	LCVTLNCS NA	KNNTKVT...
C_IN_93IN1	MWKNDMVDQM	HEDVISLWDQ	SLKPCVKLTP	LCVTLECR NV	S.....
C_IN_93IN9	MWKNDMVNQM	HEDVISLWDQ	SLKPCVKLTP	LCVTLECK NV	K.....
C_IN_93IN9	MWKNDMVNQM	HEDVISLWDQ	SLKPCVKLTP	LCVTLECS EY	NGTSKAN...
C_IN_94IN1	MWKSDMVNQM	HEDVISLWDQ	SLKPCVKLTP	LCVTLECG NV	T.....
C_IN_95IN2	MWKNDMVNQM	HEDVISLWDQ	SLKPCVKLTP	LCVTLECR NV	NST.....
CRF01_AE_C	MWKNNMVEQM	QEDVISLWDQ	SLKPCVKLTP	LCVTLHCT KA	KLNDTYN...
CRF01_AE_C	MWKNNMVEQM	QEDVISLWDQ	SLQPCVKLTP	LCVTLHCT KA	SFTNATS...
CRF01_AE_C	MWKNNMVEQM	QEDVISL.DQ	SLKPCVKLTP	LCVTLDCT KA	DFYTTKF...
CRF01_AE_T	MWKNNMVEQM	QEDVISLWDQ	SLQPCVKLTP	LCVTLHCT TA	KLTNVTN...
CRF01_AE_T	MWKNNMVEQM	QEDVISLWDQ	SLKPCVKLTP	LCVTLNCT NA	NLTNVNN...
CRF01_AE_T	MWKNNMVEQM	QEDVISLWDQ	SLKPCVKLTP	LCVTLNCT NA	NWTNANV...
CRF01_AE_T	MWKNNMVEQM	QEDVISLWDQ	SLKPCVKLTP	LCVTLNCT TA	NFTNENL...
CRF01_AE_T	MWKNNMVEQM	QEDVISLWDQ	SLKPCVKLTP	LCVTLNCT NA	NLTNGSS...

CRF01_AE_T	MWKNKMAEQM	QEDVISLWDQ	SLKPCVKLTP	LCVTLNCTNV	NATNVSN...
CRF02_AG_F	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLDCHNV	NSS.....
CRF02_AG_F	MWKNNMVEQM	HEDIISLWDQ	SLKPCVELTP	LCVTLDCCNV	SS.....
CRF02_AG_G	MWKNNMVEQM	HVDIISLWDQ	SLKPCVKLTP	LCVTLDCCNF	KN.....
CRF02_AG_N	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLDCHNF	NN.....
CRF02_AG_S	MWKNSMVEQM	HEDIISLWDQ	SLKPCVQLTP	LCVTLHCQDN	LT.....
CRF02_AG_S	MWKNNMVEQM	HVDIISLWDQ	SLKPCVKLTP	LCVTLECHNY	NYT.....
CRF03_AB_R	MGKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCTDL	KK.....
CRF03_AB_R	MGKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCTEV	KT.....
CRF04_cpx_	MWKNDMVEQM	HEDIISLWNE	GLKPCAKLTS	LCVTFTCINA	T.....
CRF04_cpx_	MWENSTVEQM	HEDIISLWDE	GLKPCVKLTP	LCVALNCSNA	TIIINS....
CRF04_cpx_	MWKNNMVEQM	HEDIISLWEE	GLKPCVKLTP	LCVALNCGDA	TIK.....
CRF05_DF_B	MWKNDMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLNCTDF	KA.....
CRF05_DF_B	MWKNNMVEQM	HADIISLWDQ	SLKSCVKLTP	LCVTLNCTDA	TS.....
CRF06_cpx_	MWKNNMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLTCTNA	TLGNKTLGNN
CRF06_cpx_	MWENHMVEQM	HEDIISLWDE	SLKPCVKLTP	LCVTLICCTNI	NITSTNS...
CRF06_cpx_	MWKNNMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCTNV	TDHGIN....
CRF06_cpx_	MWKNNMVEQM	HEDIISLWEE	SLKPCVKLTP	LCVTLNCTNV	NAT.....K
CRF11_cpx_	MWKNNMVEQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCAEV	TS.....
CRF11_cpx_	MWKNNMVEQM	HEDVISLWDE	SLKPCVKLTP	LCVALNCTDA	R.....
D_CD_84ZR0	MWKNNMVDQM	HEDIISLWDQ	SLKPCVKLTP	RCVTLNCTDA	SRN.....S
D_CD_ELI_K	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCSDE	LRNNG....T
D_CD_NDK_M	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCTDE	LRN.....S
D_UG_94UG1	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLNCTNW	VTD.....
F1_BE_VI85	MWKNNMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLNCTNA	TN.....
F1_BR_93BR	MWENNMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLDCCNI	AT.....
F1_FI_FIN9	MWENDMVEQM	HKDIISLWDQ	SLKPCVKLTP	LCVTLNCTNA	TT.....
F1_FR_MP41	MWKNNMVEQM	HEDIISLWDQ	SLKPCVKLTP	LCVTLHCSDV	NI.....
F2_CM_MP25	MWKNNMVDQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCTKA	II.....
F2KU_BE_VI	MWKNNMVEQM	HADIISLWDQ	GLQPCVKLTP	LCVTLNCSSEK	IN.....
G_BE_DRCBL	MWKNNMVEQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCTEI	N...N.....
G_NG_92NG0	MWKNNMVEQM	QEDIISLWEE	SLKPCVKLTP	LCITLNCTNV	N.....
G_SE_SE616	MWKNNMVEQM	HEDIISLWDE	SLKPCVKLTP	LCVTLNCTDV	TNKGKNR.NN
H_BE_VI991	MWVNDMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLDCCSV	NA.....
H_BE_VI997	MWDNDMVEQM	QTDIISLWDQ	SLKPCVKLTP	LCVTLDCCSI	TR.....
H_CF_90CF0	MWENNMVEQM	HTDIISLWDQ	SLKPCVKLTP	LCVTLNCTNV	RN.....
J_SE_SE702	MWKNDMVDQM	QEDIISVWDE	SLKPCVKITP	LCVTLNCSDV	NSNNS.....
J_SE_SE788	MWKNDMVDQM	QEDIISVWDE	SLKPCVKITP	LCVTLNCSNI	TSNSN.....
K_CD_EQTB1	MWKNNMVEQM	HTDIISLWDE	SLKPCVKLTP	LCVTLTCTNV	TN.....
K_CM_MP535	MWKNNMVEQM	HTDIISLWDE	SLKPCVELTP	LCVTLNCTDY	KG.....
N_CM_YBF30	MWENKMADQM	QEDIISLWEE	SLKPCVKLTP	LCVTMLCND	YGEER.....
O_CM_ANT70	IWKNYMVEQM	QEDIISLWDQ	SLKPCVQMTF	LCVQMECTN	
O_CM_MVP51	IWKNYMVDQM	HEDIISLWEE	SLKPCCKMTF	LCVQMNCVD	
O_SN_99SE	IWKNYMVEQM	QEDIISLWEE	SLKPCVQMTF	LCVQMNCCTNY	VQ.....
O_SN_99SE	IWENYMVEQM	QEDIISLWEE	SLKPCVQMTF	LCVQMNCCTN	
U_CD_83C	MWKNKMVEQM	HEDIISLWDQ	SLKPCVKLTF	LCVTLNCTDV	KN.....

	151		200
00BW0762_1	ATNVT N.....	A.....	DYKNCSFNIT TELRDKRKKE
00BW0768_2	THK V.....	TNDTLYG...	EIKNCSFNVT TEIRDKKRKE
00BW0874_2	T E.....	GNTTYGG...	EMRNCSFNIT TELRDKKRQE
00BW1471_2	VI T.....	YNNNMTE...	EMKNCSFNIT TELRDKKTKE
00BW1616_2	YNGTY S.....	D.....	GMKSCSFNIT TELRDKRRKE
00BW1686_8	..RDATSSSS E.....	GMREGMR...	EIKNCSFNVT TELRDKRKNV
00BW1759_3	KT.NFT D.....	TTNG.....	EIKNCSFNIT TEVRDRKKNE
00BW1773_2	NGS... ..	...VTQG...	EIKNCSFNVT TELRDKIQKV
00BW1783_5	.TNSTAFNTT I.....	K.....E...	EMKNCSFNMT TEIRDKRKE
00BW1795_6	.TDNTIDEGM G.....	N.....	EIKNCSFNIT TELRDKKKRE
00BW1811_3	SN T.....	YNNSIKE...	EMKNCSFNIT TEIRDKKKQV
00BW1859_5	KN I.....	TYD..NN...	EIKNCSFNIT TELRDKRKNV
00BW1880_2	NSNAS M.....	EG.....	EIENCSFNIT TELRDKRKQE
00BW1921_1	..GNGN... ..	RTG...	EIKNCSFNVT TELRDKRQOV
00BW2036_1	.TYHNVTY... ..	...NNTD...	EVKNCSFNMT TELRDKKQNV
00BW2063_6	.NASNTYQVT T.....	PT..PPD...	IMKNCSFNIP TELRDKSKKE
00BW2087_2	NNVTSC L.....	CNNTICE...	QMRNCSFNAT TEIRDKKQKV
00BW2127_2	NATANR A.....	THNPMEG...	EIKNCSFNAT TEIKDRKKQV
00BW2128_3	KSNSTN V.....	TSENKEG...	EMKNCSFNIT TELINKKQRE
00BW2276_7	GT		DMTNCTFNAT TEIKDKKRKV
00BW3819_3	VTSSV N.....	ITKNNMD...	EIKNCSFNVT TEVRDKKKQV
00BW3842_8	.TKNNDTK.. I.....	YNDTMYG...	EIKNCSFNMT TELRDKKEKM
00BW3871_3	..RNQTKNQ N.....	YTYEGIG...	EIKNCSFNMT TELRDKKKNV
00BW3876_9	VND T.....	VNHSMKE...	ETKNCSFNAT TEIRDKKRKV
00BW3886_8	.DTEN... ..	MKE...	EMRNCTFNIT TEIRDKEKQM
00BW3891_6	AT.SNG T.....	VTIN..G...	EIKNCSFNVT TELRDKRKNE
00BW3970_2	VTINN V.....	TANNNTS...	DMKNCSFNAT TEVTDKIRKE
00BW5031_1		TVAEMKG...	EIKNCSFHIS TEMRDKRQKE
96BW01B21	G..TYT D.....	N...YQE...	KIKNCSFNIT TEIRDKKQSG
96BW0407	NGTSN N.....	SSVPMEE...	EMKNCSFNIT TELRDKKQOG
96BW0502	ATNNIM I.....	D.NSNKG...	EMKNCSFNVT TELRDKRQEV
96BW06_J4	..GSNN.ANS S.....	YSNDMKE...	EIKNCSFNMT TELRDKKQKV
96BW11_06	.NDTLHQNLIT D.....		.MKNCSFNVT TELRDKRKHE
96BW1210	N.....	YNNKNG...	EIKNCSFNAT TEIRDKQQKV
96BW15B03	.NYSNTMN.. S.....	YNNNTTE...	EIKNCTFNMT TELRDKKQOV
96BW16_26	SNATMG N.....	TLENGEG...	EMKNCSFNMT TEIRDKKKQV
96BW17A09	TN N.....	VTSSMIG...	GMKNCSFNIT TELRDKRKKE
96BWM01_5	.KDINTSNAE M.....	K.....A...	EMKNCSFNIT TELRDKKKQE
96BWM03_2	.NMKKDT... ..	MKE...	EIKNCSFKVN TELKDKKHKE
98BWMC12_2	STS T.....	GTNSMNG...	QIKNCSFNIT TELRDKRKQE
98BWMC13_4	.NTTYNNAID G.....	ET..IDK...	EMKNCSFNIT TELRDKKKQE
98BWMC14_a	..VDANSTYV I.....	HVGNITT...	EMKNCSFNMT TELRDKNKNV
98BWM014_1	STRKS.	.NPSMQG...	DIKNCSFNIT TEIRDKRRKV
98BWM018_d	N N.....	DTTYNIE...	EMRNCSFNIT TEIRDKRRQE
98BWM036_a	N. T.....	IDGAMKE...	GMKNCSFNIT TEVRDKKKNQ
98BWM037_d	YTNAAT G.....	WPTETE...	KLQNCNFNVT TVIRDKKHKE
99BW3932_1	ATK N.....	GNITMKG...	EIKNCSFNAT TEIKHKKKEM
99BW4642_4	AT N.....	VNRTMTE...	EIKNCSFNIT MELRDRKQKV
99BW4745_8	TI.TNT T.....	IYKYTTS...	DIRNCPFNVV TELKDKRRKE
99BW4754_7	NTTVT V.....	TNNTMDT...	VMKNCSFNVT TELRDKRKQE
99BWMC16_8	N.....	LNNNMKE...	EIKNCSFNIT TELRDKKQNV
A2_CD_97CD	...NS.....	TE	EIKNCSYNMP TELKDKTQKV
A2_CY_94CY	...SNSSSTQ S.....	PINE	EIKNCSYNTT TILRDKTQKV
A2D___97KR	...STQS... ..	PDSN	NTMNCNFETT TELRDKKQKV
A2G_CD_97C	...TGNN... S.....	IFTE	EMKNCSYNIT TELRDKTKTV
A_BY_97BL0	...NNSSVNS N.....	SSDSLFX...	XMKNCSFNMT TELRDKRKTV
A_KE_Q23_A	...NTTGDR.	E	GLKNCSFNMT TELRDKRQKV
A_SE_SE659	...NS..TRV V.....	NITDKE...	EIKNCSFNMT TELVDKQQOV
A_SE_SE725	...NG..TQN V.....	NITN.V....	GMRNCSFNMT TELRDKKQKG
A_SE_SE753	...KNITVSS N.....	NNISISNSTE	DMRNCSFNMT TELRDKQQKV

A_SE_SE853	...VTNSSHS Y.....	NVTNMQ....	EMKNCSFNVT	TELDRKRQKV
A_SE_SE889	SSVT N.....	ITSDMAG...	EIKNCSFNMT	TEIRDKRQKV
A_SE_UGSE8	...NTNSTSA N.....	LTDSVKG...	EMRNCSFNIT	TELDRKKKKV
A_UG_92UG0	...NITNSIT N.....	SSVNMRE...	EIKNCSFNMT	TELDRKNRKV
A_UG_U455_	...NTN.NNT N.....	ITDGVR...E	EMKNCSFNMT	TELDRKKQKV
AC_IN_2130	...NSTGWGK	E	EIKNCSFNIT	TELDRKRQKV
AC_RW_92RW	...TVN....	ITDDMKG...	EIKNCSFNMT	TELDRKKQRV
AC_SE_SE94	...TPNNATH N.....	DSM..V...G	DMKNCPFNMT	TELDRKRRKE
ACD_SE_SE8	...TNATDSN N.....	..ASLQDMAK	EMTNCSFNMT	TELDRKKQRV
ACG_BE_VI1	SNGTAI N.....	ITESIKG...	EMKNCSFKAT	TEIKDKKKKE
AD_SE_SE69	ETSMN G.....		EIKNCSFNMT	TELDRKEQQV
AD_SE_SE71	...VTNATNI T.....	NANTITG...	EMKNCSFNMT	TELMDKKRKV
ADHK_NO_97	...HTD'TTNN	...TSIQPSQ	PSANCSFNVT	TAIRDKQQKV
ADK_CD_MAL	AGSNRTNAEL KM.....	EIG	EVKNCSFNIT	PVGSDKR.QE
AG_BE_VI11	.STREIRGKN CSLD.....	TEVG	ELKNCSFNIT	TELDRKKKTE
AG_NG_92NG	STGNSAGTNA TCNI.....	EEAN	NLKNCSFNIT	TEIRDKKKTE
AGHU_GA_VI	...ITNNSTD KANV.....	...TNN..DA	EMRNCSFNIT	TEIRDRKKKE
AGU_CD_Z32	NSTES N.....	ITAEMQG...	EIKNCSYNMT	TELDRKQRKI
AJ_BW_BW21	GCTNN NCT.....	VS	EMKECHFNT	GGGR..RKKE
B_AU_VH_AF	...NVTFTNS RHVTNS....	.SYVGSMEKG	EMKNCSFNIT	TSIRDKRHKE
B_CN_RL42_	...NATNTSS T.....	MEGG	EIKNCSFNIT	TSIKTKVK.D
B_DE_D31_U	...NATNTNN SSWT.....	...MTGEMKG	EIKNCSFNIT	TSIRDKVQKE
B_DE_HAN_U	...NATNSS.	...WGRMEKG	EIQNCSFKVT	TNIRDKVQKE
B_FR_HXB2_	...NDTNTNS SS.G.....	...RMIMEKG	EIKNCSFNIS	TSIRGKVQKE
B_GA_OYI_	.LRNATNTTS SS.....	...WETMEKG	ELKNCSFNIT	TSIRDKMQEQ
B_GB_CAM1_	...TRTNSSD WDRR.....	...EGEKMKG	EIKNCSFNVT	TSIRNKVRKE
B_GB_GB8_C	...NDTNTNN SIME.....	GG	EMKNCSFNIT	TSIRDKMQKE
B_GB_MANC_	...NATNTTS TNNTAS....	.GSWGAMR.G	EIKNCSFNIT	TNIRDKVHKE
B_KR_WK_AF	.N.SSTSENN TNPTIS....	..GGEGMGEG	EMKNCSFNVT	TNIRDKVQKE
B_NL_3202A	...NATNTTS S..S.....	...GVIEIEKG	EIKNCSFKIN	TNMKDKAQIE
B_TW_TWCYS	...TMSKNDS N.....	...TLTMEKG	EIKNCSFNVT	TSLRNKVQKE
B_US_BC_L0	.TNTTSTNTP SGS.....	...WKKMERG	EIKNCSFNVL	G...DKKQKA
B_US_DH123	...NGTNLKN GTK.....	..IIGKSMRG	EIKNCSFNVT	KNIIDKVKKE
B_US_JRCSE	...NATNTTS	.SSEGMMERG	EIKNCSFNIT	KSIRDKVQKE
B_US_MNCG_	...NTTNTNN STANN.NS..	.NSEGTIKGG	EMKNCSFNIT	TSIRDKMQKE
B_US_P896_	...NITKNTT N....PTS..	.SSWGMMEKG	EIKNCSFYIT	TSIRNKVKKE
B_US_RF_M1	GTNVTS SSG.....	...GTMMENG	EIKNCSFQVT	TSRRDKTQKK
B_US_SF2_K	...KATNTNS SN.....	...WKEEIKG	EIKNCSFNIT	TSIRDKIQKE
B_US_WEAU1	.LKNETNTNS SSG.....	...GEKMEEG	EMKNCSFNVT	TLIRNKRKTE
B_US_WR27_	...WNATSTS KNTTITNS..	.SNERPMEKG	EMKNCSFSIT	TSIRDKVQKE
B_US_YU2_M	.R.NATNTTS SS.....	...WETMEKG	EIKNCSFNIT	TSIRDKVQKE
BF1_BR_93B	...NST.... .QND.....	...TLKEEPG	AIQNCSFNMT	TEVRDKQLKV
C_BR_92BR0	IDYN.. N.....	RTDNMGG...	EIKNCSFNMT	TEVRDKREKV
C_BW_96BW0	V..PANGTSN	SSVSMKE...	EMRNCSFNIT	TELDRKNKQE
C_BW_96BW1	.NDTLHQNFT D.....		.MKNCSFNVT	TELDRKRKHE
C_BW_96BW1	NS..NA T.....	YNNKNNG...	EIKNCSFNAT	TEIRDKQQKV
C_BW_96BW1	.NYSNTMNAT S.....	YNNNTTE...	EIKNCTFNMT	TELDRKKQQV
C_ET_ETH22	N N.....	SINSAND...	EMKNCSFNIT	TELDRKKRKA
C_IN_93IN1	RNVSSY N.....	TYNGSVE...	EIKNCSFNAT	PEVRDRKQRM
C_IN_93IN9	NDSTHN E.....	TYTESVK...	EIKNCSFNAT	TEIRDRKQTV
C_IN_93IN9	ATNNVNATSN G.....	NATSNGE...	EIQQCFFNVT	TEMRDKKQRV
C_IN_94IN1	QNGTYN D.....	...ESNK...	EITNCTFNIT	TEIRGRKQKV
C_IN_95IN2	GNGTHS K.....	TYNESMK...	EIKNCSFNAT	TVIKDKKQTV
CRF01_AE_C	GTAKL....N	DTIGD	EVRNCSFNVT	TELDRKKQEV
CRF01_AE_C	DRIK.....	MED	AVRNCSFNMT	TELQDKKQEV
CRF01_AE_C	NTTEK....P E.....	..IEISEMQK	EVSNCPFNIT	TELDRKEQEV
CRF01_AE_T	I.....TNVP N.....	..IG..NITD	EVRNCSFNMT	TEIRDKKQKV
CRF01_AE_T	NVS N.....	..IIG.NITD	EVRNCSFNMT	TELDRKKQQV
CRF01_AE_T	TNVN..NNVT N.....	..IVG.NITE	EVRNCSFNMT	TELIDKKQKV
CRF01_AE_T	TKADNMTNVS N.....	..ITIGNITD	EVRNCTFNMT	TDLIDKKQKV
CRF01_AE_T	K.....TNVS N.....	..IIG.NITD	EVRNCTFNMT	TELTDKKQKV

CRF01_AE_T	T.....TEAP	N.....	...IV...GTD	EVKNCSFNVT	TELRDKTQQV
CRF02_AG_F	...NSSTSNS	SNSSTPINRT	IDSDMQE...	EIKNCSFNMT	TELRDKKQKV
CRF02_AG_F	...NSSISVK	S.....	ISRDMQG...	EIKNCSFNMT	TELRDKKQKV
CRF02_AG_G			ISDGSNS...	EIKNCSFNMT	TELRDKKQKV
CRF02_AG_N	SYSNSS	N.....	LTSDMNG...	EIKNCSFNIT	TEVRDKKKKM
CRF02_AG_S	SS.GN.		ISENMQG...	EIKNCSFNMT	TELRDKKQKV
CRF02_AG_S	...RNN....		SKINEVQ...	EMKNCSFNMT	TVLKDKKKKM
CRF03_AB_R	...EVTSTNT	S.....	SIKMM	EMKNCSFNIT	TDLRDKVKKE
CRF03_AB_R	...NDTSTNA	S.....	GIEMM	KN..CSFNIT	TDLRDKVKKE
CRF04_cpx_	TTNSTN	GTVI.....	...K.....E	GIKNCSFDIT	TEIRDKKKKE
CRF04_cpx_	TNNSTTNSTG	NSTV.....	...KS...TA	EIKNCSFNIT	TEVRDKQKKE
CRF04_cpx_	STNSTP	NVTT.....	...T.....N	EMKNCSFNIT	TEIRDKKKKA
CRF05_DF_B	...NSTANST	TNST.....	...TLKEETG	AVQNCSFNMT	TEVNDKKLKV
CRF05_DF_B	...ATTTSKN	ISATPTSN..	PNDTLKEEQG	AIQNCTFNIT	TEVKDKNKR
CRF06_cpx_	STNSTLGNN	TIVD.....	DISK	EIKNCSFNIT	TEIRDKTKKE
CRF06_cpx_	T.KNI	TVES.....	GE	EIKNCSFNVT	TEIRDKQKEE
CRF06_cpx_	NNN	TVEG.....	KE	EIKNCSFNVT	TEIKDKKKKE
CRF06_cpx_	YSNETVGKSL	TVKD.....	RE	EIKNCSFNIT	TEVRDQKKTE
CRF11_cpx_			YNTT	EMKNCSFNVT	TELIDRRKQE
CRF11_cpx_			DNAT	DIKNCTFNIT	TELEDKKKNE
D_CD_84ZR0	TDNNSTLPTV	KP.....	GE.	.MKNCSFNIT	TVVTDKRQV
D_CD_ELI_K	MGNNVTTEEK	G.....		.MKNCSFNVT	TVLKDKKQV
D_CD_NDK_M	KGNGKVEEEE	K.....		.RKNCSFNVR	DKREQV
D_UG_94UG1	TTNTT	G.....		.MANCSFNIT	TEIRDKKQV
F1_BE_VI85	...NSQ....		...EK...PG	AMQNCSFNMT	TEVRDKKLKL
F1_BR_93BR	...NGTNDTI	AIND.....	...TLKEDPE	AIQNCSFNIT	TEIRDKQLKV
F1_FI_FIN9	...TNDTLS.	DQSS.....	...TLKEEPG	AIQNCSFNMT	TEVEDKKQKV
F1_FR_MP41	...TSNATTT	NDTS.....	...TP.EESG	AIQNCSFNMT	TEVKDKKLRV
F2_CM_MP25	...NVTSSNN	TTLA.....	...PNVTISE	EMKNCSFNIT	TEIRDKQKKE
F2KU_BE_VI	...INSTDLT	NWANKTNNWA	NETTLLNITT	GMRNCSFNIT	TMLKDKKKKQ
G_BE_DRCBL	NS	TRNI.....	TEEY	RMTNCSFNMT	TELRDKKKAE
G_NG_92NG0	.SANHTEANN	..TV.....	ENKE	EIKNCSFKIT	TERGGKKKEE
G_SE_SE616	STDNSTETNN	S.TV.....	DNPG	EIKNCSFNVT	TEIRDKKKKE
H_BE_VI991	...TNVTKSN	NSTD.....	...INIGEIQ	EQRNCSFNVT	TAIRDKNQKV
H_BE_VI997	...NDTNSSS	TVNA.....	...TSSPSAN	ELTNCSFNVT	TVIRDKQQRV
H_CF_90CF0	...NTSNSTS	SMEA.....	...GG.....	ELTNCSFNVT	TVLRDKQQKV
J_SE_SE702	TDSNS	SASN.....	NSPE	IMKNCSFNVT	TEIRNKRKQE
J_SE_SE788	TTSNS	SVS.....	SPD	IMTNCSFNIT	TEIRNKRKQE
K_CD_EQTB1	...NRTNANK	NDT.....	NINATVTSTD	EIKNCSFNIT	TELKDKKKRV
K_CM_MP535	...TNSTN..		NATSTVVSPA	EIKNCSFNIT	TEIKDKKKKE
N_CM_YBF30	...NNTNMTT	R.....	...EPDIGYK	QMKNCSFNAT	TELTDKKKQV
O_CM_ANT70	IAG...		TTNEN	LMKKCEFNV	TVIKDKKEKK
O_CM_MVP51	...LQTNKTG		LLN...ETIN	EMRNCSFNVT	TVLTDKKEQK
O_SN_99SE_	..GNYTNSS		INNDTSSPEN	LVKQCEFNV	TVVKDKKEKK
O_SN_99SE_	..VNDETNS		VKNDTSSSEN	LMKKCEFNV	TVLKDKKEQK
U_CD_83C	STNN	N.....	TEEA	TITNCSFKVP	TELKDKTETV

	201		250
00BW0762_1	YALFYRLDIV QLGE.....	NNAN SE.....	YRLI
00BW0768_2	HALFYRLDIV PLDEKDK...	SSN SN.....	YRLI
00BW0874_2	SALFYRLDIV PLNGS.....	ERNK SE.....	YRLI
00BW1471_2	RALFYRLDIV PLNESDN...	NSY RE.....	YRLI
00BW1616_2	YAIFHSLDIV PLEN.....	SEN SE.....	YRLI
00BW1686_8	YALFYKLDIV PLEE.....	NDI ST.....	YRLI
00BW1759_3	HALFYRLDIV PLEGE.....	NNTN NE.....	YRLI
00BW1773_2	HALFYRLDIV QLD.....	N SS.....	YRLI
00BW1783_5	YALFYKLDIV PLEGNS...	E.. ..	YRLI
00BW1795_6	YALFYRLDIV SLDNENN...	KT. AE.....	YRLI
00BW1811_3	YALFYKPDIV PLDGS.....	NS SE.....	YRLI
00BW1859_5	YALFYKIDIV PLND....N.	NSN.N SM.....	YRLI
00BW1880_2	YALFYRLDVV PLDSPS....	NATN SR.....	YRLI
00BW1921_1	YALFYRLDVV QLN.....	SE.....	YRLI
00BW2036_1	YALFYKLDIV PLNGNSG...	SE.....	YRLI
00BW2063_6	YALFYKLDIV PLGNTNG...	T.. .E.....	YRLI
00BW2087_2	YALFYKLDIV SLDD.....	NN S.....	YRLI
00BW2127_2	YALFYRLDVV PLDND.....	SA TN.....	YRLI
00BW2128_3	YALFYKLDIV PLNNS.....	SDNSS GE.....	YRLI
00BW2276_7	QALFYKLDIV PLNSTGE...	NNN TE.....	YRLI
00BW3819_3	YALFYRLDVV PLNGK.....	NS SS.....	YRLI
00BW3842_8	HALFYRLDIV PLEDNSG...	NSS SN.....	YRLI
00BW3871_3	YALFYKLDIV PLND.....	N.. NE.....	YRLI
00BW3876_9	NALFYKLDVV PLHE.....	GN S.....	YRLI
00BW3886_8	YALFYRLDIV PLHDSSS...	DG. SE.....	YVLI
00BW3891_6	HALFYRLDIV PLNG.....	KNQS NE.....	YRLI
00BW3970_2	NALFYTLDIV PLDENQ....	N.....	YRLI
00BW5031_1	FALFNILDIV PLNNEN....	NTKN SD.....	YRLI
96BW01B21	YALFYKFDVV PLN.....	GNNT SE.....	YILI
96BW0407	RALFYSLDIV QPNN.....	S TE.....	YRLI
96BW0502	HALFYRLDVV PLQG.....	NN NE.....	YRLI
96BW06_J4	YALFYRLDVV PLGD.....	N.. SS.....	YRLI
96BW11_06	YALFYRLDIV PLNNKNE...	S.. SE.....	YRLI
96BW1210	YALFYRLDIV PLDN.....	NS SE.....	YRLI
96BW15B03	YALFYKLDIV PLNSNS....	SE.....	YRLI
96BW16_26	YALFYRLDVV PLNGE.....	NSNSS GE.....	YRLI
96BW17A09	SALFYRLDIV PLNENNS...	SSN SE.....	YRLI
96BWM01_5	YALFYKLDIV PLTNDAS...	EN. SE.....	YRLI
96BWM03_2	YALFYKLDIV PLDGNNE...	DGN KQ.....	YRWI
98BWMC12_2	SALFYRLDIV PLK.....	ENS SE.....	YRLI
98BWMC13_4	QALFYRLDIV PLDNANG...	T.. SE.....	YRLI
98BWMC14_a	YALFYRLDIV PLGE.....	D.. SS.....	YRLI
98BWM014_1	YALFYKLDIV ELDG.....	NS SN.....	YVLI
98BWM018_d	SALFYKLDIV PLD.....	NSS SK.....	YILI
98BWM036_a	YVLFYKLDIV PLNGNG....	SN SE.....	YRLI
98BWM037_d	YALFYRPDIV PLNEG.....	N.....	YRLI
99BW3932_1	YALFYRLDIV PLKN.....	S SE.....	YRLI
99BW4642_4	NALFYKLDIV PLNE....K.	ANNSY SY.....	YRLI
99BW4745_8	YALFYRIDIV PLDE.....	NNNS SE.....	YRLI
99BW4754_7	HALFYRLDIV PLETK.....	NSNE SA.....	YRLI
99BWMC16_8	YALFYKVIV PLSE.....	NST SE.....	YRLI
A2_CD_97CD	YSLFYELDVV LLNRSKN...	SSY ST.....	YRLI
A2_CY_94CY	YSLFYRLDVV QLDESENKNT	SGSN TL.....	YRLI
A2D___97KR	QALFYELDIV QLNSSDSND.	TLN RQ.....	YRLI
A2G_CD_97C	RSLFYTLDIV QINKDNN...	.T.....	YRLI
A_BY_97BL0	HSLFYKLDIV STSNNDX...	.Q.....	YRLI
A_KE_Q23_A	YSLFYRLDIV PINEN.....	QG..... SE.....	YRLI
A_SE_SE659	HSLFYRLDIV QMNEN.....	RGNSSNSSY NE.....	YRLI
A_SE_SE725	YSLFYKLDIV QINDN.....	GNNSNNS.. SE.....	YRLI
A_SE_SE753	YSLFYRLDLV KIDEN.....	KSN...SSN SK.....	YRLI

WO 2005/012502

PCT/US2004/009510

A_SE_SE853	TSLFYKLDVV	PIGGN.....	DTNS	TQ.....	YRLI
A_SE_SE889	HALFYRLDVV	PMDN.....	NNS.	.L.....	YRLI
A_SE_UGSE8	YSLFYKLDIV	KINKNKSFRG	.KNSSGNSSS	DR.....	YRLI
A_UG_92UG0	YSLFYKLDVV	QINNG.....	NNSS	NL.....	YRLI
A_UG_U455	YSLFYRLDIV	QINKTD.....	N	NS.....	YRLI
AC_IN_2130	YSLFYRLDVV	PIEEGQGNS.	SNSGY	KE.....	YRLI
AC_RW_92RW	YSLFYRLDIV	QINSNSN...	NSSH	NQ.....	YRLI
AC_SE_SE94	HALFYRLDIV	PLDEGNSNSN	ESNNNY	SD.....	YRLI
ACD_SE_SE8	YSLFYKLDVV	QINSN.....	QNNS	SQ.....	YRLI
ACG_BE_VI1	YALFYRLDIL	PLNKENK...	GSS	GK.....	YRLI
AD_SE_SE69	HSLFYKLDVV	QMGNSNS...		SQ.....	YRLI
AD_SE_SE71	YSLFYKLDVV	QINENQ....	..YNSSNNSN	KE.....	YRLI
ADHK_NO_97	HALFYRVDLV	SIDN.....	NDNN	TQ.....	YRLI
ADK_CD_MAL	YATFYNLDLV	QIDDSDN...	S	S.....	YRLI
AG_BE_VI11	HALFYRPDVV	PINNDN....			SSYMLI
AG_NG_92NG	YALFYRLDVV	PIDGNNNV..		.S.....	NNYRLI
AGHU_GA_VI	YALFYKHDLV	PITN.....	ET	KT.....	FILI
AGU_CD_Z32	YSLFYRLDIV	PIEENSNG.	N..S	SE.....	YRLI
AJ_BW_BW21	YALFYKEDIA	LIKDRPN...	NS	NY.....	SEYILV
B_AU_VH_AF	FALFYKLDVV	QIDGS.....	N	TS.....	YRLI
B_CN_RL42	YALFYKVDVV	PIGND.....	S	TS.....	YRLI
B_DE_D31_U	YAHFYKLDVV	PIDND.....	N	TS.....	YRLI
B_DE_HAN_U	SALFYKTDVV	PIDNNKTS..	NRDNT	TS.....	YMLI
B_FR_HXB2	YAFFYKLDII	PIDND.....	T	TS.....	YKLT
B_GA_OYI	YALFYKLDVL	PIDKN.....	D	TK.....	FRLI
B_GB_CAM1	YALFYKLDVV	PIDKAN....		TS.....	YTLI
B_GB_GB8_C	YALLYKLDIV	SIGSD.....	N	TS.....	YILT
B_GB_MANC	YALFYKLDVV	PIEKK.....	N	TS.....	FRLI
B_KR_WK_AF	YALFYKLDII	PIDN.....		TS.....	YALR
B_NL_3202A	YALFYKLDVV	PIDNN..N..	TNTSY	TS.....	YRLI
B_TW_TWCYS	YASFYRLDLV	QTDEN.....	S	TS.....	YRLI
B_US_BC_L0	YALFYKLDVV	PIDNDKNS..		TK.....	YRLI
B_US_DH123	YALFYRHDVV	PIDRN.....	I	TS.....	YRLI
B_US_JRCSF	YALFYKLDVV	PID.....	NKNN	TK.....	YRLI
B_US_MNCG	YALLYKLDIV	SIDND.....	S	TS.....	YRLI
B_US_P896	YALFNRLDVV	PIE.....	NTNN	TK.....	YRLI
B_US_RF_M1	YALFYKLDVV	PIEKGNI SPK	N.NTSNNTSY	GN.....	YTLI
B_US_SF2_K	NALFRNLDVV	PIDN..AS..	TTTNY	TN.....	YRLI
B_US_WEAU1	YALFYKLDVM	PIDHDNTS..			YTLI
B_US_WR27	HALFYRLDVV	PIDK.....	NNTN	TS.....	YRLI
B_US_YU2_M	YALFYNLDVV	PIDN.....		AS.....	YRLI
BF1_BR_93B	HALFYRLDIV	PISNDNSSND	NSS	RE.....	YRLI
C_BR_92BR0	HALFYRLDIV	PLKNE.....	SSNTS	GD.....	YRLI
C_BW_96BW0	RARFYRLDIV	QLNNN.....	SNS	NE.....	YRLI
C_BW_96BW1	YALFYRLDIV	PLNNKNE...	SN.	SE.....	YRLI
C_BW_96BW1	YALFYRLDIV	PLDN.....	NS	SE.....	YRLI
C_BW_96BW1	YALFYKLDIV	PLNSNS....		SE.....	YRLI
C_ET_ETH22	YALFYKLDIV	PLN.....	NGS	TD.....	YRLI
C_IN_93IN1	YALFYGLDIV	PLN...KKN.	SSENS	SE.....	YRLI
C_IN_93IN9	YALFYRLDIV	QLN...SDD.	KKNSS	EY.....	YRLI
C_IN_93IN9	HALFYRLDLV	PLDNENKSS.	FSNSS	KT.....	YRLI
C_IN_94IN1	YALFYKLDIV	PIS...ETS.	NQS..		RLI
C_IN_95IN2	YALFYKLDIV	PLDNEEQEN.	DSNSS	GY.....	YRLI
CRF01_AE_C	HALFYVPDIV	RIG...EK.	..NKNSSGNS	SE.....	YILI
CRF01_AE_C	HALFYTSDDV	QISSSVQNNN	NSNTSGQNNS	HK.....	FRLI
CRF01_AE_C	YALFYRSDLV	PIE.....	..RNSGENNG	SS.....	YRLI
CRF01_AE_T	HALFYKLDIV	QIEDK.....	NDS	SK.....	YGLI
CRF01_AE_T	HALFYKLDIV	QMN.....	KNS	SE.....	YRLI
CRF01_AE_T	YALFYKLDIV	QMN.....	SNS	SE.....	YRLI
CRF01_AE_T	YALFYKLDIV	PIG.....	NNN	NM.....	YRLI
CRF01_AE_T	HALFYKLDIV	QIEDK.....	KTS	SE.....	YRLI

CRF01_AE_T	QALFYKLDIV	QMGG.....	NDS	GE.....	YRLI
CRF02_AG_F	SALFYRLDVV	QINES.....	GN	SQ.....	YRLI
CRF02_AG_F	SALFYRLDVV	QINES.....	SN	SQ.....	YRLI
CRF02_AG_G	NALFYRVDVV	QMNS.....		.Q.....	YRLI
CRF02_AG_N	HALFYRLDVV	QINEN.....	NG	SQ.....	YRLI
CRF02_AG_S	YALFYRYDVV	QINETG....	DN	IQ.....	YRLI
CRF02_AG_S	AALFYKIDIV	PIDKN.....	A	TY.....	YRLI
CRF03_AB_R	YALFYKLDVV	QIDND.....		.S.....	YRLI
CRF03_AB_R	HALFYKLDVV	QIDND.....		.S.....	YRLI
CRF04_cpx_	YALFYRIDIV	PINARVPING	...SNRNNST	EE.....	YMLI
CRF04_cpx_	HALFYRLDVV	PINNNVPINN	...TSNTSEY	RE.....	YRLM
CRF04_cpx_	YALFYRLDIV	PINDNNSTN.	...SRRSSNT	SD.....	YMLI
CRF05_DF_B	HALFYRLDIV	PISSD....D	SSN	SS.....	YRLI
CRF05_DF_B	HALFYRLDIV	SINS.....	SRK	E.....	YRLI
CRF06_cpx_	YALFYRPDIV	PIGDD.....	S	NN.....	SDYRLI
CRF06_cpx_	YALFYRLDVV	PINDG.....	S	NN.....	NSYRLI
CRF06_cpx_	RALFYTLDVV	PINDN.....	G	NN.....	STYRLI
CRF06_cpx_	YALFYRPDVM	QVDG.....	K	NS.....	STYRLI
CRF11_cpx_	YALFYKLDIV	PINDNNN...	SS	NV.....	SDYRLI
CRF11_cpx_	RALFYRLDVV	PINDS.....	SS	NI.....	GQYRLI
D_CD_84ZR0	HALFYRLDVV	QIDNEGKNE.	INDTY	GT.....	YRLI
D_CD_ELI_K	YALFYRLDIV	PIDNDSS...	TNS	TN.....	YRLI
D_CD_NDK_M	YALFYKLDIV	PIDNNNR...	TNS	TN.....	YRLI
D_UG_94UG1	QALFYKLDVV	KINDNDS...	DN	TS.....	YRLI
F1_BE_VI85	SALFYRLDIV	PIGNN...N.	S	SE.....	YRLI
F1_BR_93BR	HALFYKLDIV	QINKD..DN.		RT.....	YRLI
F1_FI_FIN9	HALFYRLDIE	PISNN...N.	SR	EE.....	YRLI
F1_FR_MP41	NALFYKLDII	PINNS.....	SS	SD.....	YRLI
F2_CM_MP25	YALFYKLDVV	QINNS.....		NTS.....	YRLI
F2KU_BE_VI	YALFYREDIV	PINIKKNKKT	NSN	SKKNNTSNN	SIENSKYRLI
G_BE_DRCBL	YALFYRTDVV	PINEMNNENN	GT	NS.....	TWYRLT
G_NG_92NG0	YALFYKLDVV	PISNGN....		.K.....	TSYRLI
G_SE_SE616	YAFFYRLDVV	PINN.....		.A.....	TNYRLT
H_BE_VI991	HALFYRADIV	QIDEGER...	NKSD	NH.....	YRLI
H_BE_VI997	HALFYRLDVV	PIDETSNNN.	NSNS	TK.....	YRLI
H_CF_90CF0	HALFYRLDVV	PIDNNS....		TQ.....	YRLI
J_SE_SE702	YALFYRQDVV	PIN.....	S	DN.....	KSYILI
J_SE_SE788	YALFYRQDVV	PID.....	S	NN.....	KNYILI
K_CD_EQTB1	SALFYKLDIV	QIKQSEINQS	ESE		DRLI
K_CM_MP535	SALFYRLDVL	PLN.GEGNS	STE		YRLI
N_CM_YBF30	YSLFYVEDVV	PINAYN....			KTYRLI
O_CM_ANT70	QALFYVSDLM	ELNETSSTNK	T	NS.....	KMYTLT
O_CM_MVP51	QALFYVSDLS	KVNDNAVN.		.G.....	TTYMLT
O_SN_99SE_	QALFYVSDLM	KINEANDT..		.K.....	DMYTLI
O_SN_99SE_	QALFYVSDLM	KVNENND...			TMYTLI
U_CD_83C	HTLFYKLDVV	PLNVTN....	N	SS.....	..ISSTYRLI

	251				300
00BW0762_1	NCNTSTITQA	CPKVNFDPIP	IHYCAPAGYA	ILKCNTKTFD	GTGPCTNVST
00BW0768_2	NCNTSAVTQA	CPKVSFEPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCHNVST
00BW0874_2	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW1471_2	NCNTSTITQA	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW1616_2	KCNTSTITQA	CPKVNFDPIP	IHYCAPAGYA	ILKCRNKTFN	GTGPCNNVST
00BW1686_8	NCNTSSISQA	CPKVSFGPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCQNVST
00BW1759_3	NCNTSAVTQA	CPKVTFDPIP	VHYCAPAGYA	ILKCNNKTFN	GAGPCNNVST
00BW1773_2	NCNTSAITQA	CPKVSFDPIP	IHYCTPAGYA	ILKCNNQTFN	GTGPCNDVSS
00BW1783_5	NCNTSAITQA	CPKVSFEPIP	IHFCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW1795_6	NCNTSTITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GKGPCNNVST
00BW1811_3	NCNTSALTQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCTNVST
00BW1859_5	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNDKTFN	GTGPCQNVST
00BW1880_2	NCNTSAITQA	CPKINFDPPI	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW1921_1	NCNTSAITQA	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW2036_1	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCQNVST
00BW2063_6	NCNTSTITQS	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW2087_2	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVSI
00BW2127_2	NCNTSAITQA	CPKISFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW2128_3	NGNTSALTQA	CPKVSFDPIP	IHYCTPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW2276_7	NCNTSAITQA	CPKITFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCQNVSP
00BW3819_3	NCNTSAVTQS	CPKISFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW3842_8	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GIGPCQNVSI
00BW3871_3	NCNTSAISQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCQNVST
00BW3876_9	HCNTSTITQA	CPKVSFEPIP	IHYCAPAGYA	ILKCNDKTFN	GTGPCLNVST
00BW3886_8	NCNTSTITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNKTFN	GTGPCQNVST
00BW3891_6	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW3970_2	NCNTSKVTQA	CPKVSFDPIP	LHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
00BW5031_1	SCNTSTITQA	CPKVSWDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
96BW01B21	NCNTSAISQA	CPKVSFDPIP	IHYCTPAGFA	ILKCNNKTFN	GTGPCNNVST
96BW0407	NCNTSTITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GLGPRNNVST
96BW0502	NCNTSAITQA	CPKVSFDPIP	IHYCTPAGYA	ILKCNNQTFN	GTGPCNNVSS
96BW06_J4	NCNTSTISQA	CPKISFDPIP	IHYCAPGGYA	ILKCNNKTFN	GTGPCQNVST
96BW11_06	NCNTSTITQS	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
96BW1210	NCNTSTITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCTNVST
96BW15B03	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCKNNKTFN	GTGPCQNVST
96BW16_26	NCNTSALTQA	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
96BW17A09	NCNTSTITQA	CPKVSFDPIP	IHYCAPAGYA	ILECNNKTFN	GTGPCTNVST
96BWM01_5	NCDTSTITQS	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
96BWM03_2	NCNTSSITQA	CPKVSFDPIP	IHYCAPAGFA	ILKCNNKTFN	GTGPCQNVST
98BWMC12_2	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
98BWMC13_4	NCNTSTITQS	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GIGPCNNVST
98BWMC14_a	NCNTSAISQA	CPKISFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCQNVST
98BWM014_1	NCNTSTVKQA	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
98BWM018_d	NCNTSVITQA	CPKVTFEPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
98BWM036_a	NCNTSTLTQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
98BWM037_d	NCNTSTVTQA	CPKVSFDPIP	IHYCTPAGYA	ILKCNNKTFN	GKGPCNNVST
99BW3932_1	NCYTSAIAQT	CPKISFDPIP	IHYCAPAGYA	ILKCYNKTFN	GTGPCKNVST
99BW4642_4	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCQNVST
99BW4745_8	NCNTSAVTQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
99BW4754_7	NCNTSAITQA	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTRPCNNVST
99BWMC16_8	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNDVST
A2_CD_97CD	SCNTSVITQA	CPKVSFEPIP	IHYCAPAGYA	ILKCKDKEFN	GKGSCSNVSS
A2_CY_94CY	NCNTSTITQA	CPKVTFEPIP	IHYCAPAGFA	ILKCKDPRFN	GTGSCCKNVSS
A2D_97KR	HCDTSTITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCKDPEFN	GTGPCKNVSS
A2G_CD_97C	KCNTSTITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCKDPKFN	GTGPCENVSS
A_BY_97BL0	NCNTSAMTQA	CPXVTFEPIP	IYYCAPAGFA	ILKCKDTNFT	RTXPCKNVST
A_KE_Q23_A	NCNTSAITQA	CPKVSFEPIP	IHYCTPAGFA	ILKCKDEGFN	GTGLCKNVST
A_SE_SE659	NCNTSAITQA	CPKVTFEPIP	IHYCAPAGFA	ILKCKDDAFN	GTGRCKNVST
A_SE_SE725	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCRDKEFN	GTGPCNNVST
A_SE_SE753	NCNTSAITQA	CPKVTFEPIP	IHYCAPAGFA	ILKCNKTFN	GTGLCNCNVST

A_SE_SE853	NCNTSAITQA	CPKVTFEPIP	IHYCAPAGFA	ILKCRDENFN	GTGPCKNVST
A_SE_SE889	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNDKEFN	GTGPCKNVSS
A_SE_UGSE8	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNEDEFN	GTGTCRNVST
A_UG_92UG0	NCNTSALTQA	RPKVTFEPIP	IHYCAPAGYA	ILKCNDKEFN	GTGLCKNVST
A_UG_U455	NCNTSTITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCKDPEFN	GKGPCRVNST
AC_IN_2130	NCNTSAITQA	CPKVTFNPIP	IHYCTPAGYA	ILKCNEETFN	GTGPCKNVST
AC_RW_92RW	NCNTSAITQA	CPKVSFEPIP	INYPAPAGFA	ILKCKDKKFN	GTGPCKNVST
AC_SE_SE94	NCNTSAITQA	CPKVSFDPIP	IHYCAPPGFA	ILKCKDAKFN	GIGPCNNVST
ACD_SE_SE8	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNNKEFN	GTGPCKNVST
ACG_BE_VI1	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGFA	ILKCNNKTFN	GTGPCNNVST
AD_SE_SE69	NCNTSAIKQA	CPKVTFEPIP	IHYCAPAGFA	ILKCKDTEFN	GTGPCKNVST
AD_SE_SE71	NCNTSAITQA	CPKVTFEPIP	IHYCAPAGFA	ILKCKDEKFN	GTGPCKNVST
ADHK_NO_97	NCNTSVITQA	CPKVTFEPIP	IHYCAPAGFA	ILKCNNKTFS	GTGPCKNVST
ADK_CD_MAL	NCNTSVITQA	CPKVTFDPIP	IHYCAPAGFA	ILKCNDKKFN	GTEICKNVST
AG_BE_VI11	NCNSSTIKQA	CPKVSFEPIP	IHYCAPAGFA	ILRCRDKKFN	GTEPCKNVST
AG_NG_92NG	NCNVSTIKQA	CPKVSFDPLP	IHYCAPAGFA	ILKCRGKNFT	GTGQCKNVSS
AGHU_GA_VI	HCNTSTITQA	CPKVSFDPIP	IHYCAPAGFA	ILKCKDKAFN	GTGPCKNVST
AGU_CD_Z32	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCRDEEFE	GKGPCRVNST
AJ_BW_BW21	KCNTTVIKQA	CPKVSFQPIP	IHYCAPAGFA	ILQCNDKKFN	GTGPCKNVST
B_AU_VH_AF	NCNTSVITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNNKTFN	GKGPCANIST
B_CN_RL42	NCNTSVITQA	CPKVSFEPIP	IHYCTPAGFA	ILKCNNKKFN	GTGPCTNVST
B_DE_D31_U	SCNTSVITQA	CPKVSFEPIP	IHYCTPAGFA	MLKCKDKRFN	GKGQCKNVST
B_DE_HAN_U	HCNRSVITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNDKKFN	GKGPCKNVST
B_FR_HXB2	SCNTSVITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNNKTFN	GTGPCTNVST
B_GA_OYI	HCNTSTITQA	CPKISFEPIP	MHYCTPAGFA	ILKCNDKKFN	GTGPCTNVST
B_GB_CAM1	HCNTSVITQA	CPKVSFEPIP	IHYCTPAGFA	ILKCNDKKFN	GKGPCCTNVST
B_GB_GB8_C	ECNASVITQA	CPKISFEPIP	IHFAPAGFA	ILKCNNKTFD	GKGPCCTNVST
B_GB_MANC	SCNTSTITQA	CPKVSFEPIP	IHYCTPAGFA	ILKCNNKKFD	GKGQCTNVST
B_KR_WK_AF	HCNTSVITQA	CPKVSFEPIP	IHYCAPAGFA	ILQCNDKKFN	GTGPCS NVST
B_NL_3202A	SCNTSVITQA	CPKVSFEPIP	IHFCTPAGFA	LLKCNDKKFN	GTGPCKNVST
B_TW_TWCYS	SCNASVIKQA	CPKVSFEPIP	IHYCTPAGFA	ILKCNNKTFN	GTGTCTNVST
B_US_BC_L0	SCNTSVTTQA	CPKISFEPIP	IHYCAPAGFA	ILKCKDKKFN	GTGSCKNVST
B_US_DH123	SCNTSTLTQA	CPKVSFEPIP	IHYCAPAGFA	ILKCKDKKFN	GTGPCTNVST
B_US_JRCSF	SCNTSVITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNNKTFN	GKGQCKNVST
B_US_MNCG	SCNTSVITQA	CPKISFEPIP	IHYCAPAGFA	ILKCNDKKFS	GKGSCKNVST
B_US_P896	SCNTSVITQA	CPKVSFQPIP	IHYCVPAGFA	MLKCNNKTFN	GSGPCTNVST
B_US_RF_M1	HCNSSVITQA	CPKVSFEPIP	IHYCTPAGFA	ILKCNDKKFN	GTGPCKNVST
B_US_SF2_K	HCNRSVITQA	CPKVSFEPIP	IHYCTPAGFA	ILKCNNKTFN	GKGPCCTNVST
B_US_WEAU1	NCKSSTITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNDKKFN	GKGPCKNVST
B_US_WR27	HCNTSTITQA	CPKVSFEPIP	IHYCTPAGFA	ILKCNDKKFN	GTGQCKNVST
B_US_YU2_M	SCNTSVITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNDKKFN	GTGPCTNVST
BF1_BR_93B	NCNTSTLTQA	CPKVSWDPIP	IHYCAPAGYA	ILKCNDKKFN	GTGPCRVNST
C_BR_92BR0	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
C_BW_96BW0	NCNTSTITQA	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GSGPCNNVST
C_BW_96BW1	NCNTSTITQS	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
C_BW_96BW1	NCNTSTITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCTNVST
C_BW_96BW1	NCNTSAITQA	CPKVSFDPIP	IHYCAPAGYA	ILKCKNNTFN	GTGPCQNVST
C_ET_ETH22	NCNTSTITQA	CPKVSLDPIP	IHYCAPAGYA	ILKCRDKTFT	GTGPCHNVST
C_IN_93IN1	NCNTSAITQA	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCNNVST
C_IN_93IN9	NCNTSAITQA	CPKVTFDPIP	IHYCTPAGYA	ILKCKDKTFN	GTGPCHNVST
C_IN_93IN9	NCNTSAITQA	CPKVTFDPIP	IHYCTPAGYA	ILKCNEKTFN	GTGLCQNVST
C_IN_94IN1	SCNTSVITQA	CPKVSFDPIP	IHYCTPAGYA	ILKCNDKTFN	GTGPCRVNST
C_IN_95IN2	NCNTSALTQA	CPKVTFDPIP	IHYCAPAGYA	ILKCNNKTFN	GTGPCHNVST
CRF01_AE_C	HCNTSVIKQA	CPKVTFDPIP	IHYCTPAGYA	ILKCNEKNFN	GTGPCKNVSS
CRF01_AE_C	HCNTSVIKQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNDKNFN	GTGPCKNVSS
CRF01_AE_C	HCNTSVIKQA	CPKVSFDPIP	IHYCAPAGYA	ILKCNDKNFN	GTGPCKKVSS
CRF01_AE_T	NCSTSVIKQA	CPKISFDPIP	IHYCTPAGYV	ILKCNDKNFN	GTGPCKNVSS
CRF01_AE_T	NCNTSVIKQA	CPKISFDPIP	IHYCTPAGYA	ILKCNDKNFN	GTGPCKNVSS
CRF01_AE_T	NCNTSVIKQA	CPKVSFDPIP	IHYCTPAGYA	ILKCNDKNFN	GTGPCKNVSS
CRF01_AE_T	NCNDSVIKQA	CPKISFDPIP	IHYCTPAGYA	MLKCNDKNFN	GTGPCKNVSS
CRF01_AE_T	NCNTSVIKQA	CPKISFDPIP	IHYCTPAGYA	ILKCNDKNFN	GTGPCKNVSS

CRF01_AE_T	NCNTSVIKQA	CPKVSFDPIP	IHYCTPAGYA	ILKCNDKNFN	GTGPCKNVSS
CRF02_AG_F	NCHTSAITKA	CPRVTLEPIP	IHYCAPAGLA	ILKCNDKKFN	GTGLCKNVST
CRF02_AG_F	NCNTSAIKQA	CPKVTFEPIP	IHYCAPAGFA	ILKCNDKKFN	GTGPCKNVST
CRF02_AG_G	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCKDRNFN	GTGPCQNVST
CRF02_AG_N	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCKDKGFN	GTGPCKNVST
CRF02_AG_S	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNDKFN	GTGPCKNVST
CRF02_AG_S	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNEKDFN	GTGSCKNVST
CRF03_AB_R	SCNTSVVTQA	CPKISFEPIP	IHYCAPAGFA	ILKCNDKKFN	GTGPCTNVST
CRF03_AB_R	SCNTSVVTQA	CPKISFEPIP	IHYCAPAGFA	ILKCNDKKFN	GTGPCTNVST
CRF04_cpx_	NCNASTIKQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNEKNFT	GLGPCTNVSS
CRF04_cpx_	SCNTSNHKQA	CPKVTLEPNS	HTLLCPGWFC	DLKCNDKNSP	GLGSCTNVSP
CRF04_cpx_	NCNVSSITQA	CPKIKFEPIP	IHYCAPAGFA	ILQCNEKRFN	GSGPCKNVSS
CRF05_DF_B	NCNTSTIKQA	CPKVSWDPIP	IHYCAPAGYA	ILKCNEKKFS	GTGPCKNVST
CRF05_DF_B	NCNTSTIKQA	CPKVCWDPIP	IHYCAPAGYA	ILKCKEKRFN	GTGPCKNVST
CRF06_cpx_	NCNVSTIKQA	CPKVSFEPIP	IHYCAPAGFA	ILKCRDKDFN	GTGPCKNVST
CRF06_cpx_	NCNASTIKQA	CPKVTFEPIP	IHYCAPAGFA	ILKCRDKNFN	GTGPCKNVST
CRF06_cpx_	NCNTSTIKQA	CPKVSFEPIP	IHYCAPAGFA	ILKCRDKNFN	GTGPCKNVST
CRF06_cpx_	NCNASTIKQA	CPKVSFEPIP	IHYCAPAGFA	ILKCRDKNFN	GTGACKNVST
CRF11_cpx_	NCNVSTIKQA	CPKVTFEPIP	IHYCAPAGFA	ILKCRDKEFN	GTGPCKNVST
CRF11_cpx_	NCNVSAVKQA	CSKVTFEPIP	IHYCAPAGFA	ILKCRDKEFN	GTGPCKNVST
D_CD_84ZR0	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNDKRFN	GTGPCKNVSS
D_CD_ELI_K	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCRDKKFN	GTGPCTNVST
D_CD_NDK_M	NCDTSTITQA	CPKISFEPIP	IHFCAPAGFA	ILKCRDKKFN	GTGPCS NVST
D_UG_94UG1	NCNTSAITQA	CPKMTFEPIP	IHYCAPAGFA	ILKCNEKKFN	GTGPCKNVST
F1_BE_VI85	NCNTSTITQA	CPKVSWDPIP	IHYCAPAGYA	ILKCNDKRFN	GTGPCKNVST
F1_BR_93BR	NCDASTITQA	CPKVSWDPIP	IHYCAPAGYA	ILKCNEKNFT	GTGSCKNVST
F1_FI_FIN9	TCNTSTITQA	CPKVSWDPIP	IHYCAPAGYA	ILKCKDKRFN	GTGPCRNVST
F1_FR_MP41	NCNTSTIKQA	CPKVSWDPIP	IHYCAPAGYA	ILKCRDPRFN	GTGPCKNVST
F2_CM_MP25	NCNTSTLTQA	CPKVSFDPIP	IHYCAPAGFA	ILKCNNKTFN	GTGLCRNVST
F2KU_BE_VI	NCNTSAITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNEKEFN	GTGPCKNVST
G_BE_DRCBL	NCNVSTIKQA	CPKVTFEPIP	IHYCAPAGFA	ILKCVDKKFN	GTGTCNNVST
G_NG_92NG0	HCNVSTIKQA	CPKVNFDPIP	IHYCAPAGFA	ILKCRDKEYN	GTGPCKNVST
G_SE_SE616	HCNVSTIKQA	CPKVTFDPIP	IHYCAPTGFA	ILKCRDKEFN	GTGPCKNVST
H_BE_VI991	NCNTSVIKQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNGKKFN	GTGPCTNVST
H_BE_VI997	NCNTSVITQA	CPKVSFDPIP	IHYCAPAGFA	ILKCNNKTFN	GTGPCTNVST
H_CF_90CF0	NCNTSVITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNNKTFN	GTGLCTNVST
J_SE_SE702	NCNTSVIKQA	CPKVSFQPIP	IHYCAPAGFA	ILKCNNKTFN	GTGPCKNVST
J_SE_SE788	NCNTSVIKQA	CPKVSFQPIP	IHYCAPAGFA	ILKCNDKNFN	GTGSCKNVST
K_CD_EQTB1	NCNTSTVTQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNNNTCN	GTGPCTNVST
K_CM_MP535	NCNTSTITQT	CPKVTFEPIP	IHYCAPAGFA	ILKCKDKRFN	GTGPCKNVST
N_CM_YBF30	NCNTTAVTQA	CPKTSFEPIP	IHYCAPPGFA	IMKCNEGNFS	GNGSCTNVST
O_CM_ANT70	NCNSTTITQA	CPKVSFEPIP	IHYCAPAGYA	IFKCNSTEFN	GTGTCRNITV
O_CM_MVP51	NCNSTTIKQA	CPKVSFEPIP	IHYCAPTGFA	IFKCNSTDFN	GTGLCHNISV
O_SN_99SE_	NCNSTTIKQA	CPKVSFEPIP	IHYCAPAGYA	IFKCNSTEFN	GTGPCNNITA
O_SN_99SE_	NCNSTTIKQT	CPKVSFEAIP	IHYCAPAGYA	IFKCNNTGFN	GTGPCTNVTV
U_CD_83C	NCNTSTITQA	CPKVSFEPIP	IHYCAPAGFA	ILKCNDKKFN	GTGPCKNVST

	301			350
00BW0762_1	VQCTHGIKPV	VSTQLLLLNGS	LSEE.GIVIR	SENITNNAKT IIVHLNESVE
00BW0768_2	VQCTHGIKPV	VPTQLLLLNGS	LAEE.EIIIR	SEALTNNAKT I IARLNKSVE
00BW0874_2	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENLAKNKDT IIVHLNESVE
00BW1471_2	VLCTHGIKPV	VSTQLLLLNGS	LSEG.EIMIR	SENLTDNAKT IIVQLNKPVE
00BW1616_2	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIVIR	SENITNNAKI IIVHLNETVD
00BW1686_8	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIIR	SENMTDNVKT IIVHLNESVE
00BW1759_3	VQCTHEIKPV	VSTQLLLLNGS	LAEE.QIIIR	SENMTNNAKT IIVHLKKPVQ
00BW1773_2	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIIR	SENLTNNAKI IIVHLKDPVN
00BW1783_5	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENLTNNVKT IIVHLNESIQ
00BW1795_6	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIIR	SKNLTDNART IIVHLNESVQ
00BW1811_3	VQCTHGIKPV	VSTQLLLLNGS	LAEE.DIIIR	SENLTDNVKT IIVHLNESVE
00BW1859_5	VQCPHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENIEDNVKT IIVHLNESIE
00BW1880_2	VQCTHGIKPV	VATQLLLLNGS	LAEE.EIVIR	FKNITNNAKI IIVQLNTSVG
00BW1921_1	VQCTHGIKPV	VSTQLLLLNGS	LAEE.GIIIR	SENLTDNAKT ITVQLDQAVE
00BW2036_1	VQCTHGIRPV	VSTQLLLLNGS	LAEG.EIIIR	SENLAADNAKT IIVHFNESVE
00BW2063_6	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIIR	SKNITDNVKT IIVHLNEAVE
00BW2087_2	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENLTNNAKT IIVHLNDSVE
00BW2127_2	VQCTHGIKPL	VSTQLLLLNGS	LAQQ.QIMIT	SENLTNNAKI IIVHLKEAIN
00BW2128_3	VQCTHGIKPV	VSTQLLLLNGS	LAEE.IIIR.	SENLTNNAKT IIVQLKEPVK
00BW2276_7	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENLTNNVKT IIVHLNKSVE
00BW3819_3	VQCTHGIKPV	VSTQLLLLNGG	LAEK.EIIIR	SENLTNNVKT IIVHLNESVE
00BW3842_8	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIIR	SKNLSDNAKI IIVHLNESVG
00BW3871_3	VQCTHGIKPV	VSTQLLLLNGS	IAEG.EIIIR	SENLTNNAKT IIVHLNESVK
00BW3876_9	IQCTHGIKPV	VSTQLLLLNGS	LAEK.DIIIR	SENLTNNIKT IIVHLNDSVQ
00BW3886_8	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIIR	SENLTNNAKT IIVHLNESIE
00BW3891_6	VQCTHGIKPV	VSTQLLSNGS	LSEK.EIIIR	SENLTNNVKT IIVHLNESVE
00BW3970_2	IQCTHGIKPV	VSTQLLLLNGS	RAEK.EIIIR	SENMTNNAKT IIVHLNESIE
00BW5031_1	VQCTHGIKPV	VSTQLLLLNGS	LA.E.EIVIR	SENISDNAKT IIVHLNESVE
96BW01B21	VQCTHGIKPV	VSTQLLLLNGS	LAEE.DIIIR	SENLTDNVKT IIVHLNESIE
96BW0407	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENLTDNAKI IIVHLNESVR
96BW0502	VQCAHGIKPV	VSTQLLLLNGS	VAKG.EIIIR	SENLTNNAKI IIVQLNKPVK
96BW06_J4	VQCTHGIKPV	VSTQLLLLNGS	LAEE.KIIIR	SGNVTNNAKT IIVHLNESVE
96BW11_06	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIIR	FKNITNNAKT IIVHLNESVE
96BW1210	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENITDNVKT IIVHLNKSVE
96BW15B03	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIIR	SKNLSNNAYT IIVHLNDSVE
96BW16_26	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENLTNNAKT IIVHLNESVE
96BW17A09	VQCTHGIKPV	VSTRLLLLNGS	LAEE.EIIIR	SENLTNNAKI IMVQLNESIE
96BWMO1_5	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIIR	SKNITDNVKT IIVHLNESVE
96BWMO3_2	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENLTDNAKT IIVHFNESVQ
98BWMC12_2	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SENLTNNVKT IIVHLNESVE
98BWMC13_4	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIVIR	SENLTDNVKT IIVHLNQSVQ
98BWMC14_a	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EVVIR	SENLTNNAKT IIVQLKEPVK
98BWMO14_1	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIIR	SENITNNIKT IIVHLNEPVE
98BWMO18_d	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIIIR	SKNITNNANT IIVQLKDPVD
98BWMO36_a	IQCTHGIKPV	VSTQLLLLNGS	IAEE.EIIIR	SENLTDNVKT IIVHLKDPID
98BWMO37_d	VQCTHGIKPV	VSTQLLLLNGS	IAEE.EIIIR	SENLTNHFRT IIVQLNQSIE
99BW3932_1	VQCTHGIKPV	VSTQLLLLNGS	LAEE.EIVIR	SENLTNNVKT IIVHLNESIE
99BW4642_4	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIVR	SENLTDNVKT IIVHLDKPVG
99BW4745_8	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIVR	SENLTNNAKI IIVQLKQSVG
99BW4754_7	VQCTHGIKPV	VSTQLLLLNGS	RAEK.EVVIR	AESLTDNAKV IIVHLKDSVQ
99BWMC16_8	VQCTHGIKPV	VSTQLLLLNGS	LAEE.DIIIR	SKNLTDNAKT IIVHLNESVN
A2_CD_97CD	VQCAHGIRPV	ASTQLLLLNGS	LAEG.KVMIR	SENITDNAKN IIVQFNKPVP
A2_CY_94CY	VQCTHGIKPV	ASTQLLLLNGS	LAEGGKIMIR	SENITNNAKN IIVQFTKPVL
A2D_97KR	VQCTHGIRPV	ASTQLLLLNGS	LAEGNKTIIR	SANITDNTKN IIVQFTKPVQ
A2G_CD_97C	VQCTHGIRPV	VSTQLLLLNGS	LAEE.EVMIR	SENITDNAKN IIVQFDKPVE
A_BY_97BL0	VQCTHXIKPV	VSTQLLLLNGS	LAEKX.VMIR	SENITDNVKI IIVQLTEPVN
A_KE_Q23_A	VQCTHGIKPV	VSTQLLLLNGS	LAEN.ITIR	SENITNNAKI IIVQLVQPV
A_SE_SE659	VQCTHGIKPV	VSTQLLLLNGS	LAKG.IRIR	SENITNNVKT IIVQLDKPVN
A_SE_SE725	VQCTHGIKPV	VSTQLLLLNGS	LAEEK.IMIR	SENISDNAKT IIVQLTEPVT
A_SE_SE753	VQCTHGIKPV	VSTQLLLLNGS	LATK..IMIR	SENITNNAKT IIVQLVEPVE

A_SE_SE853	VQCTHGIKPV	VSTQLLLLNGS	LAREK.VMIR	SENITNNVKN	IIVQLKEPVE
A_SE_SE889	VQCTHGIRPV	VSTQLLLLNGS	LAETE.VMIR	AENITNNIKN	IIVQFNKSVE
A_SE_UGSE8	VQCTHGIRPV	VSTQLLLLNGS	LAKEE.VRIR	SENISDNAKT	IIVQFTKPVE
A_UG_92UG0	VQCTHGIRPV	VSTQLLLLNGS	LAEGK.VMIR	SENITNNVKN	IIVQLNESVT
A_UG_U455_	VQCTHGIKPV	VSTQLLLLNGS	LAERE.IRIR	SENFTNNAKT	IIVQLVNPVK
AC_IN_2130	VQCTHGIKPV	VSTQLLLLNGS	LAEKE.VIIR	SENITNNVKN	IIVQLAEPVR
AC_RW_92RW	VQCTHGIRPV	VSTQLLLLNGS	LAEKE.IIIR	SENITNNAKT	IIVQLNETVQ
AC_SE_SE94	VQCTHGIKPV	ISTQLLLLNGS	LSETG.VKIR	SENITNNAKT	IIVQLDEAVE
ACD_SE_SE8	VQCTHGIKPV	VSTQLLLLNGS	LAEKE.IIVR	SENLTNNAKI	IIIQLNETVK
ACG_BE_VI1	VQCTHGIKPV	VSTQLLLLNGS	LAEK..VVIR	SENITNNAKT	IIVQFDSVPK
AD_SE_SE69	VQCTHGIKPV	VSTQLLLLNGS	LA.EGKVRIR	SENITDNTKN	IIVQFTEPVT
AD_SE_SE71	VQCTHGIKPV	VSTQLLLLNGS	LAKEE.VIIR	SENITNNAKN	IIVQFVKPVT
ADHK_NO_97	VQCTHGIKPV	VSTQLLLLNGS	LA..EKVIIR	SKNITDNTKN	IIVHFNESVQ
ADK_CD_MAL	VQCTHGIKPV	VSTQLLLLNGS	LA.EEEIMIR	SENLTNNTKN	IIVQLNETVT
AG_BE_VI11	VQCTHGIKPV	VSTQLLLLNGS	LA.EEEIIVR	SENFTNNAKV	IIVQLKEPIE
AG_NG_92NG	VQCTHGIKPV	VSTQLLLLNGS	LA.EGEIVIR	SENLTNNAKV	IIVQLNKTIG
AGHU_GA_VI	VQCTHGIRPV	VTTQLLLLNGS	LA.EGEIIR	SENITENTKN	IIVQLNETVE
AGU_CD_Z32	VQCTHGIKPV	VSTQLLLLNGS	LAEKE.VRIR	SENFSDNAKI	IIVQLAKPVN
AJ_BW_BW21	VQCTHGIKPV	VSTQLLLLNGS	IA.EEEIIR	SENITNNAKT	IIVQLNNTVE
B_AU_VH_AF	VQCTHGIRPV	VSTQLLLLNGS	LA.EKEIVIR	SDNFTDNAKS	IIVQLNESVE
B_CN_RL42_	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVIK	FSNFTDNARV	IIVQLNESVE
B_DE_D31_U	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVIR	SDNFTDNAKT	IIVQLKESVE
B_DE_HAN_U	VQCTHGIRPV	VSTQLLLLNGS	LA.EKEVVIR	SDNFTDNTKT	IIVHLNESVE
B_FR_HXB2_	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVIR	SVNFTDNAKT	IIVQLNTSVE
B_GA_OYI_	VQCTHGIKPV	VSTQLLLLNGS	LA.EEEVIIR	SSNFTNNAKI	IIVQLNKSVE
B_GB_CAM1_	VQCTHGIRPV	VSTQLLLLNGS	LA.EKEVVIR	SENFTNNAKT	IIVQLKEPVE
B_GB_GB8_C	VQCTHGIRPV	VSTQLLLLNGS	LA.EEKVVIR	SDNFTDNVKT	IIVQLKEAVE
B_GB_MANC_	IQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVLIR	SDNFTDNAKT	IIVHLNESVE
B_KR_WK_AF	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEIVLR	SENFTNNAKT	IIVQLNASVE
B_NL_3202A	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVIR	SANFSNNAKT	IIVQLNESVA
B_TW_TWCYS	VQCTHGIRPV	VSTQLLLLNGS	IA.EEEILIK	SENITNNAKT	IIIQLNKSVK
B_US_BC_L0	VQCTHGIKPV	VSTQLLLLNGS	LA.EEEVVIR	SANFSDNAKT	IIVQLKEAVE
B_US_DH123	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVIR	SSNFTDNAKI	IIVQLNETVE
B_US_JRCSF	VQCTHGIRPV	VSTQLLLLNGS	LA.EEKVVIR	SDNFTDNAKT	IIVQLNESVK
B_US_MNCG_	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVIR	SENFTDNAKT	IIVHLNESVQ
B_US_P896_	VQCTHGIRPV	VSTQLLLLNGS	LA.EEDIVIR	SENFTDNAKT	IIVQLNESVV
B_US_RF_M1	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVIR	SENFTDNVKT	IIVQLNASVQ
B_US_SF2_K	VQCTHGIRPI	VSTQLLLLNGS	LA.EEEVVIR	SDNFTNNAKT	IIVQLNESVA
B_US_WEAU1	VQCTHGIRPV	VSTQLLLLNGS	LA.EEDIVIR	SENFTDNAKN	IIVQLNVSIE
B_US_WR27_	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEVVIR	SANFTNNAKT	IIVQLKESVE
B_US_YU2_M	VQCTHGIRPV	VSTQLLLLNGS	LA.EEEIVIR	SENFTNNAKT	IIVQLNESVV
BF1_BR_93B	VQCTHGIKPV	VSTQLLLLNGS	LA.EKDIIIR	SONISDNAKT	IIVQLNVSVP
C_BR_92BR0	IQCTHGTGPV	VSTQLLLLNGS	LAEK.EIIR	SKNLTNDVKT	IIVHLNESVE
C_BW_96BW0	VQCTHGIKPV	VSTQLLLLSGS	LAEK.EIVIR	SENLTNNAKI	IIVHLNKTVR
C_BW_96BW1	VQCTHGIKPV	VSTPLLLNGS	LAEK.EIIR	FKIITNNAKT	IIVHLNESVE
C_BW_96BW1	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIR	SENITDNVKT	IIVHLNKSVE
C_BW_96BW1	VQCTHGIKPV	VSTQLLLLNGS	LAEGGEIIR	SKNLSNNAYT	IIVHLNDSVE
C_ET_ETH22	VQCTHGIKPV	VSTQLLLLNGS	IAEG.ETIIR	FENLTNNAKI	IIVQLNESVE
C_IN_93IN1	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIR	SENLTNNVKT	IIVHLNQSVQ
C_IN_93IN9	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIR	SENLTNNAKT	IIVHLNQSVQ
C_IN_93IN9	VQCTHGIKPV	VSTQLLLLNGS	LAEG.EIIR	SENLTNDVKT	IIVHLNQSVQ
C_IN_94IN1	VQCTHGIKPV	VSTQLLLLNGS	LSEG.EIIR	SENLTNNVKT	IIVHLNKSVE
C_IN_95IN2	VQCTHGIKPV	VSTQLLLLNGS	LAEG.GIIR	SENLTNNVKT	IIVHLNQPVE
CRF01_AE_C	VQCTHGIKPV	VSTQLLLLNGS	LAEK.DIIR	SENLTNNAKT	IIVHLNKSVE
CRF01_AE_C	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIR	SEDLTDNAKT	IIVHLNKSIE
CRF01_AE_C	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIR	SENLTNNAKT	IIVHLNKSVE
CRF01_AE_T	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIR	SENLTNNAKT	IIVHLNKSVE
CRF01_AE_T	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIR	FENLTNNAKT	IIVHLNKSVE
CRF01_AE_T	VQCTHGIKPV	VSTQLLLLNGS	LAEK.KIIR	SENLTNNAKT	IIVHLHESVE
CRF01_AE_T	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIR	SENLTNNAKT	IIVHLNKSVE
CRF01_AE_T	VQCTHGIKPV	VSTQLLLLNGS	LAEK.EIIR	SEDLTNNAKT	IIVHLNKSVE

CRF01_AE_T	VQCTHGIKPV	VSTQLLLNGS	LAKE.EIIIR	SENLTNNAKT	IIVHLNKSVK
CRF02_AG_F	VQCTHGIKPV	VSTQLLLNGS	LAEEE.VVIR	SENITNNAKN	IIVQLVAPVR
CRF02_AG_F	VQCTHGIKPV	VSTQLLLNGS	LAEEE.VVIR	SENITNNAKN	IIVQLVTPVR
CRF02_AG_G	VQCTHGIKPV	VSTQLLLNGS	LAEEE.IVIR	SENITNNVKN	IIVQLAKPVR
CRF02_AG_N	VQCTHGIKPV	VSTQLLLNGS	LAEGE.VVIR	SENITNNAKT	IIVQLANPVK
CRF02_AG_S	VQCTHGIKPV	VSTQLLLNGS	LAEEE.IVIR	SENFTNNAKI	IIVQLHESVK
CRF02_AG_S	VQCTHGIKPV	VSTQLLLNGS	LAEGD.IVIR	SENISNNAKT	IIVQLNKPVW
CRF03_AB_R	VQCTHGIKPV	VSTQLLLNGS	LA.EEEVVIR	SVNFTDNTKT	IIVQLKEPVE
CRF03_AB_R	VQCTHGIKPV	VSTQLLLNGS	LA.EEEVVIR	SVNFTDNTKT	IIVQLKEPVE
CRF04_cpx_	VRCTHGIKPV	VSTQLLLNGS	LA.TEEVVIR	SKNITDNTKN	IIVQLAKAVK
CRF04_cpx_	VQCTHGIKPV	VSTQLLLNGS	LA.TGGVVIR	SKNFTDNPKN	IIVQLDKAVK
CRF04_cpx_	VQCTHGIKPV	VSTQLLLNGS	LS.TEGVLR	SKNFTDNTKN	IIVQLAEAVK
CRF05_DF_B	VQCTHGIKPV	VSTQLLLNGS	LA.KEGIIIR	SONISDNAKN	IIVHLNESVH
CRF05_DF_B	VQCTHGIKPV	VSTQLLLNGS	LA.EESIIIR	SONILDNTKT	IIVHLNESVQ
CRF06_cpx_	VQCTHGIKPV	VSTQLLLNGS	LA.EGNITIK	TENITDNTKN	IIVQLNQPVE
CRF06_cpx_	VQCTHGIKPV	VSTQLLLNGS	LA.EEEIIIK	SKNLTDNTKI	IIVQLNKSVE
CRF06_cpx_	VQCTHGIKPV	VSTQLLLNGS	LA.EDEIIIK	SENHTNNAKI	IIVQLNKTVO
CRF06_cpx_	VQCTHGIKPV	VSTQLLLNGS	LA.EEEIIIK	TENLTDNSKN	IIVQLNKSIE
CRF11_cpx_	VQCTHGIKPV	VSTQLLLNGS	LA.EEKVKIR	SENFTNNAKT	IIVQFNNTVR
CRF11_cpx_	VQCTHGIKPV	VSTQLLLNGS	LA.EGEVRIR	SENLTNNAKT	IIVQLNSTVR
D_CD_84ZR0	VQCTHGIRPV	VSTQLLLNGS	LA.EEEIVIR	SENLTNNAKI	IIVHLNQSVE
D_CD_ELI_K	VQCTHGIRPV	VSTQLLLNGS	LA.EEEVIIR	SENLTNNAKN	IIAHLNESVK
D_CD_NDK_M	VQCTHGIRPV	VSTQLLLNGS	LA.EEEIIIR	SENLTNNVKT	IIVQLNASIV
D_UG_94UG1	VQCTHGIKPV	VSTQLLLNGS	LA.EEEIIIR	SENLTNNAKI	IIVQLNESVP
F1_BE_VI85	VQCTHGIKPV	VSTQLLLNGS	LA.EEGIVIR	SONISNNAKT	IIVHLNESVQ
F1_BR_93BR	VQCTHGIKPV	VSTQLLLNGS	LA.EGEIVIR	SONISDNAKT	IIVHLNESVQ
F1_FI_FIN9	VQCTHGIRPV	VSTQLLLNGS	LS.EGGIIIR	SONLSDNAKT	IIVHLNESVQ
F1_FR_MP41	VQCTHGIRPV	VSTQLLLNGS	LA.EEDIIIR	SONISDNAKT	IIVHLNESVQ
F2_CM_MP25	VQCTHGIKPV	VSTQLLLNGS	LA.EEKMIIR	SENISDNTKT	IIVQFKNPVK
F2KU_BE_VI	VQCTHGIRPV	ISTQLLLNGS	LA.EKEIIIR	SGNISDNTKN	IIVQLNETVE
G_BE_DRCBL	VQCTHGIKPV	VSTQLLLNGS	LA.EKDIIIS	SENISDNAKV	IIVHLNRSVE
G_NG_92NG0	VQCTHGIKPV	VSTQLLLNGS	LA.EEDIRIR	SENFTDNTKV	IIVQLNNSIE
G_SE_SE616	VQCTHGIKPV	VSTQLLLNGS	LA.EGKIKVR	SENFTDNTKV	IIVQLNKTVE
H_BE_VI991	VQCTHGIRPV	VSTQLLLNGS	LA.EVEEVIIR	SKNITDNTKN	IIVQLNEPVQ
H_BE_VI997	VQCTHGIKPV	VSTQLLLNGS	LA.EGQVIIR	SKNISDNTKN	IIVQLDSPIE
H_CF_90CF0	VQCTHGIRPV	VSTQLLLNGS	LA.EEQIIIR	TKNISDNTKN	IIVQLKTPVN
J_SE_SE702	VQCTHGIKPV	VSTQLLLNGS	VA.EGDIIIR	SENISDNAKN	IIVQLNDTVE
J_SE_SE788	VQCTHGIKPV	VSTQLLLNGS	IA.EGDIIIR	SENISDNAKN	IIVQLNKTVE
K_CD_EQTB1	VQCTHGIKPV	VSTQLLLNGS	LA.EEEIIIR	SEDITKNNTKN	IIVQLNEAVE
K_CM_MP535	VQCTHGIKPV	VSTQLLLNGS	LA.EEEIIIR	SENITDNTKN	IIVQLNETVQ
N_CM_YBF30	VQCTHGIKPV	ISTQLILNGS	LNTDGIVIR.	NDSHSN	LLVQWNETVP
O_CM_ANT70	VTCTHGIRPT	VSTQLILNGT	LS.KGKIRMM	AKDILEGGKN	IIVTLNSTLN
O_CM_MVP51	VTCTHGIKPT	VSTQLILNGT	LS.REKIRIM	GKNITESAKN	IIVTLNTPIN
O_SN_99SE_	VTCTHGIKPT	VSTQLILNGT	LS.EGNIRIM	GKNISDNMKN	IIVTLNSTIN
O_SN_99SE_	VTCTHGIRPT	VSTQLILNGT	IS.EGKIRIM	GKNISDTGKN	IIVTINSTIN
U_CD_83C	VQCTHGIRPV	VSTQLLLNGS	LSEE.EVIIR	SENITNNAKT	IIVQLNETVK

	351		400
00BW0762_1	IVCTRPNNN.	.TRRSVRIG.	.PGQTFYATG DIIGN IRQAHCNISK
00BW0768_2	IECTRPNNN.	.TRKSIRIG.	.PGQTFYATG DIIGD IREAHCNINK
00BW0874_2	IVCTRPNNN.	.TRKSIRIG.	.PGQTFYATG DVIGD IRQAHCNTSE
00BW1471_2	IVCTRPNNN.	.TRKSVRIG.	.PGQTFYATG DIIGD IRQAHCNISN
00BW1616_2	IVCTRPNNN.	.TRKSMRIG.	.PGQTFYATG DIIGN IREAHCNISK
00BW1686_8	INCTRPNNN.	.TRKSIRIG.	.PGQTFYATG AIIGD IRQAYCTVNA
00BW1759_3	IVCTRPNNN.	.TRRSVRIG.	.PGQTFYARG DIIGN IRQAHCNISR
00BW1773_2	ITCTRPNNN.	.TRKSIRIG.	.PGQTFYATG DIIGD IRQAHCNISA
00BW1783_5	INCTRPNNN.	.TRKSVRIG.	.PGQTFYATG DIIGD IRQAHCNISE
00BW1795_6	IVCTRPNNN.	.TRKSVRLG.	.PGQAFYATG DIIGD IRKAYCTINE
00BW1811_3	IVCTRPNNN.	.TRKSIRIG.	.PGQAFFATG EIIGD IRQAHCNISA
00BW1859_5	IVCIRPNNN.	.TRKSIRIG.	.PGQTFYATG DIIGN IREAHCNITR
00BW1880_2	IVCTRPNNN.	.TKKSMRIG.	.PGQTFYATG DIIGD IRQAHCNISE
00BW1921_1	IECTRPNNN.	.TRKSIRIG.	.PGQTFYATG DIIGN TRQAHCNVSA
00BW2036_1	IECIRPNNN.	.TRKSIRIG.	.PGQVFFYATG DIIGD IREAHCNITE
00BW2063_6	IVCTRPNNN.	.TRKSVRIG.	.PGQTFYATG EVIGD IREAHCNISE
00BW2087_2	IVCTRPNNN.	.TRKSIRIG.	.PGQAFYATD AIIGD IRQAHCNISR
00BW2127_2	IVCTRPNNN.	.TRTSIRIG.	.PGHSFFATN GIIGD IRQAHCSISK
00BW2128_3	INCTRPNNN.	.TRKSIRIG.	.PGQAFYATG DIIGD IRQAHCNISK
00BW2276_7	IVCVRPNNN.	.TRKSVRIG.	.PGQTFEAT NIIGD IREAHCNISE
00BW3819_3	IKCTRPNNN.	.TRRSVRIG.	.PGQAFYTN DIIGD IRLAHCNISK
00BW3842_8	IVCTRPNNN.	.TRKSIRIG.	.PGQTFYAAG DIIGN IRQAHCNISE
00BW3871_3	ITCTRPNNN.	.TRESIRIG.	.PGQTFYATG DIIGD IRKAYCNISI
00BW3876_9	IVCTRPNNN.	.IRKSVRIG.	.PGQAFYATG DIIGD IREAYCNING
00BW3886_8	IVCVRPNNN.	.TRKSIRIG.	.PGQTFYATG EIIGN IRQAYCSISG
00BW3891_6	IECTRPNNN.	.TRRSIRIG.	.PGQTFYATG EIIGD IRQAYCTINE
00BW3970_2	IECIRPNNN.	.TRKSIRIG.	.PGQTFYATN GMIGD IRQAHCNISG
00BW5031_1	IECRRPNNN.	.TGKSVRIG.	.PGQTFEATG GIIGE IRRAHCDING
96BW01B21	INCTRPNNN.	.TRKSIRIG.	.PGQTFYAAG EIIGK IRLAYCNISE
96BW0407	IECTGPNNN.	.TRKSMRIG.	.PGQTFYATG EIVGD IRQAHCNISE
96BW0502	IVCVRPNNN.	.TRKSVRIG.	.PGQTFYATG EIIGD IRQAYCIINK
96BW06_J4	IVCTRPNNN.	.TRKSIRIG.	.PGQTFYAT DIIGD IRQAYCNVSK
96BW11_06	IVCIRPNNN.	.TRKSVRIG.	.PGQTFYATE AIIGN IREAHCNISE
96BW1210	IVCTRPNNN.	.TRKSIRIG.	.PGQTFYATG DIIGD IRQAHCNISK
96BW15B03	IVCTRPNNN.	.TRKGIRIG.	.PGQTFYATE NIIGD IRQAHCNISA
96BW16_26	IVCIRPNNN.	.TRKSIRIG.	.PGQTFEATG DIIGD IRQAHCIING
96BW17A09	IVCTRPNNN.	.TRKSTRIG.	.RGQTFYAMG RIIGD IRQAHCNISG
96BWM01_5	IECTRPNNN.	.TRRSVRIG.	.PGQAFYATG DIIGD IRAAHCNISE
96BWM03_2	INCTRPNNN.	.TRKSVRIG.	.PGQAFYATG DIIGD IRKAYCNISK
98BWMC12_2	IVCTRPNNN.	.TRKSMRIG.	.PGQIFYATG DIIGD IREAHCNISK
98BWMC13_4	IECTRPNNN.	.TRKSMRIG.	.PGQAFYATG EIIGN IRQAYCNINE
98BWMC14_a	IVCTRPNNN.	.TRKSIRIG.	.PGQTFYATG DIIGD IRQAHCNISE
98BWM014_1	IVCTRPNNN.	.TRTSIRIG.	.PGQTFYATG DIIGD IRQAHCNISE
98BWM018_d	ILCVRPSNN.	.TRKSVRIG.	.PGQTFYATG DIIGD IRQAHCNISA
98BWM036_a	IVCTRPNNN.	.TRKSVRIG.	.PGQTFEATG DIIGD IRQAHCNISK
98BWM037_d	INCTRPNNN.	.TRKSIRIG.	.PGQAFYATN DIIGD IRQAHCNISE
99BW3932_1	IVCIRPNNN.	.TRKSIRIG.	.PGQTFYATG AIIGN IREAYCNISG
99BW4642_4	IVCIRPNNN.	.TRKSIRIG.	.PGQTFYATG DIIGN IKEAYCNIKE
99BW4745_8	IECIRPNNN.	.TRKSIRIG.	.PGQTFYATG EIIGD IRKAHCTINK
99BW4754_7	INCTRPNNN.	.TRKSMRIG.	.PGQTFYATG EIIGD IRQAHCNISR
99BWMC16_8	ITCTRPNNN.	.TRKSIRIG.	.PGQTFYATG DIIGD IRQAHCSINK
A2_CD_97CD	INCTRPNNN.	.TRKSIRFG.	.PGQAFYTN NIIGD IRQAHCNISI
A2_CY_94CY	ITCIRPNNN.	.TRKSIRFG.	.PGQAFYTN EIIGD IRQAHCNINK
A2D_97KR	INCTRPDVG.	.QRRSVRIG.	.PGRAFYTRO TYTR.QAKGD IRQAQCNISS
A2G_CD_97C	ITCIRPNNN.	.TRKSIRFG.	.PGQAFYTN SIIGD IRQAYCNISK
A_BY_97BL0	ITCIRPGNN.	.TRTSIRIG.	.PGQTFYATG DVIXD IRKAYCNVSR
A_KE_Q23_A	IKCIRPNNN.	.TRKSIRIG.	.PGQAFYATG DIIGD IRQAHCNVTR
A_SE_SE659	ITCIRPYHN.	.TRTRIHIH.	.PG.RSFYTG DIKGS IRQAHCTVNR
A_SE_SE725	INCTRPNNN.	.TRTSIRIG.	.PGQAFYATG DITGD IRQAHCNVSR
A_SE_SE753	INCTRPNNN.	.TRTSVPIG.	.PGKVFFYATG EIIGD IRQAHCNVSK

A_SE_SE853	INCTRPNN.	.TRKSIRIG.	.PGQAFYATG	EVIGD	IRQAHCNVSR
A_SE_SE889	IICIRPNN.	.TRKSIRIG.	.PGQAFYATG	DIIGD	IRQAYCDVNR
A_SE_UGSE8	IICTRPNN.	.TRKSIRIG.	.PGQAFYGMG	DIIGD	IRKAHCNVSR
A_UG_92UG0	INCTRPNN.	.TRRSVRIG.	.PGQTFYATG	DIIGD	IRQAHCNVSG
A_UG_U455	INCSRPNNT	KNIRRYSIG.	.SGQAFYVTG	KIIGD	IRQAHCNVSR
AC_IN_2130	INCTRPNN.	.TRTSIRIG.	.PGQTFYTS.	NIIGD	IRQAHCNVSR
AC_RW_92RW	INCSRPNNN.	.TRKSVHIG.	.PGQAFYATG	DVIGD	IRQAYCTVNG
AC_SE_SE94	INCTRPNN.	.TRRSVHIG.	.PGQAFYATG	DITGD	IRKAHCIVNG
ACD_SE_SE8	INCTRPNN.	.TRNSIRIG.	.PGQAFYATG	AITGD	IRQAHCNVSR
ACG_BE_VI1	INCTRPNN.	.TRKSVRIG.	.PGQTFYATG	DIIGD	IRQAHCNISG
AD_SE_SE69	INCTRPNNNT	.RK.SVRIG.	.PGQALYVTG	GII..G...D	IRQAFCEVNR
AD_SE_SE71	INCTRPNN.	.TRKSVHMG.	.PGKVIFYATG	DIIGD	IRQAHCNVSK
ADHK_NO_97	INCTRIANNT	RKS..IHIG.	.PGQAFYAAE	PVI..G...D	IRQAHCNISE
ADK_CD_MAL	INCTRPNNNT	.RR.GIHFG.	.PGQALYTTG	IVG.....D	IRRAYCTINE
AG_BE_VI11	INCTRPNNNT	RKSIGLPG.	...QAFYATG	DII..GD...	IRQAHCNVSG
AG_NG_92NG	INCTRPNNNT	RKSIRIGPG.	...QAFYATG	EII..G....	
AGHU_GA_VI	INCTRPNNNT	RKG..IRIG.	.PGRVIYATS	AIT..G...D	IRQAHCNISK
AGU_CD_Z32	ITCMRPNNY.	.TRKSIHIG.	.PGRALYPEG	DIIGD	IRQAHCNVSR
AJ_BW_BW21	IKCVRPANNT	RKGIHTGPG.	...QVLYATG	AVV..GD...	IRQAHCNVSR
B_AU_VH_AF	IHCMRPNNNT	.RK.GIYVG.	.PGRHIYATE	KIV..G...D	IRQAHCNISR
B_CN_RL42	IKCIRPNNNT	.RK.SIHLG.	.PGKAWYTTG	QII..G...D	IRQAHCNLSS
B_DE_D31_U	INCTRPNNYT	.SK.RIRIG.	.ARRAFYTKG	KII..G...D	IRQAHCNISG
B_DE_HAN_U	INCTRPNNNT	.RK.GIHIG.	.PGRAVYTTG	RIV..G...D	IRLAHCNISR
B_FR_HXB2	INCTRPNNNT	.RK.RIRIQR	GPGRAFVTIG	KIG.....N	MRQAHCNISR
B_GA_OYI	INCTRPNNNT	.RN.RISIG.	.PGRAFHTTK	QII..G...D	IRQAHCNLSR
B_GB_CAM1	INCTRLNNNT	.RK.SIAIG.	.PGRTVYATD	RII..G...D	IRQAHCNLSS
B_GB_GB8_C	INCTRPNNNT	.RK.GIYMG.	.PGRRFYTTG	RII..G...D	IRQAHCNISK
B_GB_MANC	INCTRPNNNS	.RK.SIYIG.	.PGRRFHVTR	AVT..G...D	IRQAHCNISK
B_KR_WK_AF	INCTRLNNNT	.RK.SIRIG.	.PGSTFYATG	AII..G...D	IRQAHCNISR
B_NL_3202A	INCTRPNNNT	.RK.GIHIG.	.PGKAFYATG	QII..G...D	IRQAHCNLSR
B_TW_TWCYS	INCTRPNNIS	KRR.SMHIG.	.TGRVFYTTG	..I..G...N	IRQAHCNLSK
B_US_BC_L0	INCTRPNNKT	.RK.RITTG.	.PGRVYTTG	EIV..G...D	IRQAHCNLSR
B_US_DH123	INCTRPNNNT	.RK.GITLG.	.PGRVFYTTG	EIV..G...D	IRKAHCNISK
B_US_JRCSF	INCTRPNNNT	.RK.SIHIG.	.PGRAFYTTG	EII..G...D	IRQAHCNISR
B_US_MNCG	INCTRPNNYK	.RK.RIHIG.	.PGRAFYTTG	NII..G...T	IRQAHCNISR
B_US_P896	INCTRPNNNT	.RR.RLSIG.	.PGRAFYARR	NII..G...D	IRQAHCNISR
B_US_RF_M1	INCTRPNNNT	.RK.SITKG.	.PGRVIYATG	QII..G...D	IRKAHCNLSR
B_US_SF2_K	INCTRPNNNT	.RK.SIYIG.	.PGRAFYTTG	RII..G...D	IRKAHCNISR
B_US_WEAU1	INCTRPNNNT	.RK.KITLG.	.PGRVLYTTG	EII..G...D	IRRAHCNLSR
B_US_WR27	INCTRPNNKI	.RR.RIHIG.	.PGRAFYTDR	..V..G...D	IRQAYCNISG
B_US_YU2_M	INCTRPNNNT	.RK.SINIG.	.PGRALYTTG	EII..G...D	IRQAHCNLSK
BF1_BR_93B	INCTRPNNNT	RKS..IPIG.	.PGRAFYTTG	EII..G...D	IRKAHCNVSG
C_BR_92BR0	INCTRPNN.	.TRKSIRIG.	.PGQAFYATG	EIIGD	IRQAHCNISR
C_BW_96BW0	IVCTRPNN.	.TRRSVRIG.	.PGQTFYATG	EIIGD	IR.AHCNISE
C_BW_96BW1	IVCIRPNN.	.TRKSVRIG.	.PGQTFYATE	AIIGN	ISEAHCNISE
C_BW_96BW1	IVCTRPNN.	.TRKSIRIG.	.PGQTFYATG	DIIGD	IRQAHCNISK
C_BW_96BW1	IVCTRPNN.	.TRKGIRIG.	.PGQTFYATE	NIIGD	IRQAHCNISA
C_ET_ETH22	ITCTRPNN.	.TRESIRIG.	.PGQTFYATG	DIIGD	IRQAHCNISE
C_IN_93IN1	IVCTRPNN.	.TRKSIRIG.	.PGQTFYATG	DIIGD	IRQAHCNISR
C_IN_93IN9	IVCTRPNN.	.TRKSIRIG.	.PGQTFYATG	EIIGD	IRQAHCNISK
C_IN_93IN9	IECVRPNN.	.TRESIRIG.	.PGQTFYATG	EIIGD	IRQAHCNISA
C_IN_94IN1	IVCTRPNN.	.TRKSIRIG.	.PGQTFYATG	EIVGN	IRQAHCNISK
C_IN_95IN2	IMCTRPNN.	.TRKSIRIG.	.PGQTFYATG	DIIGD	IRQAHCNISE
CRF01_AE_C	INCTRPNN.	.MRTSARIG.	.PGQVFYKTG	SITGD	IRKAYCEING
CRF01_AE_C	INCTRPNN.	.VRISARIG.	.PGRVFHTTG	NINGD	IRKAYCEINK
CRF01_AE_C	INCTRPNN.	.MRTSVRIG.	.PGRVFYKTG	SITGD	IRKAYCEING
CRF01_AE_T	INCTRPNN.	.MRTSMRIG.	.PGQVFYRTG	SITGD	IRKAYCEING
CRF01_AE_T	INCTRPNN.	.TRTSITMG.	.PGQVFYRTG	DIIGD	IRKAYCEING
CRF01_AE_T	INCTRPNN.	.KRTRTSIG.	.QGRVLYRTG	DITGN	IGKPYCEING
CRF01_AE_T	INCTRPNN.	.EEKRMTIG.	.PGKVFYSTG	KITGD	IRKAYCVING
CRF01_AE_T	INCTRPNN.	.TRTSITIG.	.PGRVFYRTG	DIIGN	IRKAYCEING

CRF01_AE_T	INCTRP.TI.	.YKKKTTMG.	.PARVYHRTG	DVIGD	IRKAYCQING
CRF02_AG_F	INCPRPNNN.	.TRKSVRIG.	.PGQTFYATG	DIIGD	IRQAHCNVSR
CRF02_AG_F	INCTRPNNN.	.TRKSVRIG.	.PGQTFYATG	DIIGD	IRKAHCNVSR
CRF02_AG_G	INCTRPNNN.	.TRKSVRIG.	.PGQTFYATG	GIIGD	IRQAHCNVSR
CRF02_AG_N	INCTRPNNN.	.TRKGVHIG.	.PGQAFYATG	DIIGD	IRQAHCNVSK
CRF02_AG_S	INCTRPGNN.	.TRKSVRIG.	.PGQTFYATG	DIIGD	IRQAHCNVSW
CRF02_AG_S	INCTRPSNN.	.TRKSVRIG.	.PGQTFYTTG	AVIGD	IRQAHCNVSR
CRF03_AB_R	INCTRPNNNT	.RK.GIHIG.	.PGRAFYATG	DIT..G...D	IRQAHCNISI
CRF03_AB_R	INCTRPNNNT	.RK.GIHIG.	.PGRAFYATG	DII..G...D	IRQAYCNISR
CRF04_cpx_	INCTRPGNNT	RKS..VHIG.	.PGLTWYATG	EII..G...D	IRQAHCNISG
CRF04_cpx_	INCTGLNNNT	GGSERIGIG.	.PGHTWYATG	NIV..G...D	IRQAHCNISG
CRF04_cpx_	INCTRPNNNT	RKG..VHIG.	.PGKTWFATG	EVI..G...D	IRKAHCNISE
CRF05_DF_B	INCTRPNNNT	RKS..IHLG.	.PGQAFYATG	DII..G...D	IRKAHCNVSR
CRF05_DF_B	INCTRPNNNT	RKS..IPLG.	.PGQAFYTTG	DII..G...D	IRKAHCNVSG
CRF06_cpx_	IRCTRPGNNT	RKSISFGPG.	...QAFIATG	DII..GD...D	IRQAHCNVSR
CRF06_cpx_	ISCSRPNNT	RKSIHIGPG.	...QAFYATG	EII..GN...D	IRKAHCNVSR
CRF06_cpx_	IRCTRPSNNT	RKSIPLGPG.	...QAFYATG	DII..GD...D	IRQAHCNVSR
CRF06_cpx_	IKCTRPNNNT	RKSISFAPG.	...QAFYATG	DII..GD...D	IRQAHCNVSR
CRF11_cpx_	INCTRPGNNT	RKSIHLGPG.	...HAFYATG	AII..GD...D	IRQAHCNVSK
CRF11_cpx_	INCTRPNNNT	RKGIHIGPG.	...QAFYATG	DII..GD...D	IKQAHCNVSR
D_CD_84ZR0	INCTRPYKKE	.RQ.RTPIG.	.QGQALYTTR	YT.....TRI	IGQAYCNISG
D_CD_ELI_K	ITCARPYQNT	.RQ.RTPIG.	.LGQSLYTTR	SR.....SI	IGQAHCNISR
D_CD_NDK_M	INCTRPYKYT	.RQ.RTSIG.	.LRQSLYTIT	GKK..KKTGY	IGQAHCNISR
D_UG_94UG1	INCIRPYNNT	.RQ.STRIG.	.PGQALFTTK	VIG.....D	IRQAHCNISG
F1_BE_VI85	INCTRPNNNT	RKG..IHLG.	.PGQTFYATG	AII..G...D	IRKAHCNISG
F1_BR_93BR	INCTRPNNNT	RKR..ISLG.	.PGRVFYTTG	EII..G...D	IRKAHCNVSG
F1_FI_FIN9	INCTRPNNNT	RKS..IRIG.	.PGQSFYATG	EII..G...D	IRKAHCNISG
F1_FR_MP41	INCTRPNNNT	RKS..IHLG.	.PGQAFYATG	DII..G...D	IKKAYCEING
F2_CM_MP25	INCTRPNNNT	RRS..IHIG.	.PGRAFYATG	EII..G...D	TRKAHCNISE
F2KU_BE_VI	IVCIRPGNNT	RKS..IRIG.	.PGQTFYATG	DII..G...D	IRQAHCNITG
G_BE_DRCBL	INCTRPNNNT	RRSVAIGPG.	...QAFYTTG	EVI..GD...D	IRKAHCNVSW
G_NG_92NG0	INCIRPNNNT	RKSIPIGPG.	...QAFYATG	DII..GD...D	IRQAHCNVSR
G_SE_SE616	INCTRPNNNT	MKRIRMGIGP	..GQTFYATG	AII..GD...D	IRQAHCNVTK
H_BE_VI991	INCTRTGNNT	RKS..IRIG.	.PGQAFYATG	DII..G...D	IRRAYCNISG
H_BE_VI997	ITCTRPNNNT	RKG..IHFG.	.PGQAFYATG	DII..G...N	IRQAHCNVSE
H_CF_90CF0	ITCTRPNNNT	RTS..IHLG.	.PGRAFYATG	DII..G...D	IRQAHCNISR
J_SE_SE702	IVCTRPNNNT	RKGIHMGPG.	...QVLYATG	EII..GD...D	IRKAYCNISR
J_SE_SE788	IVCYRPNNNT	RKGIHMGPG.	...QVLYATG	EII..GN...D	IRETHCNISE
K_CD_EQTB1	INCTRPSNNT	RKS..IHIG.	.PGRAFYATG	DII..G...D	IRQAHCNISG
K_CM_MP535	INCTRPNNNT	RKS..IHMG.	.PGKAFYTTG	DII..G...D	IRQAHCNISG
N_CM_YBF30	INCTRPGNN.	.TGGQVQIG.	.PAMTFYNIE	K.....IVGD	IRQAYCNVSK
O_CM_ANT70	MTCERP.QI.	.DIQEMRIG.	..PMAWYSMG	IGG..TAGNS	SRAAYCKYNA
O_CM_MVP51	MTCIREGIA.	.EVQDIYTG.	..PMRWRSMT	LKRSNNTSPR	SRVAYCTYNK
O_SN_99SE_	MTCVRQGNQ.	.SVQEIQIG.	..PMAWYSMS	LAQE.GKPNN	SRIAYCKYNI
O_SN_99SE_	MTCERPGNQ.	.TVQKILTG.	..PVAWYSMG	LKN...NLTN	SRAASCKYNS
U_CD_83C	INCTRPGSDK	KIRQSIRIG.	.PGKVIFYAKG	GI	TGQAHCNITD

	401		450
00BW0762_1	TKWNKTLRSRI	GEKLKEHFNP	...KTIRFN. .SSAGGDLEI TTHSFNCRGE
00BW0768_2	EKWNSTLQGV	REKLEKHFPN	...KNITFE. .PSSGGDLEI TTHSFNCRGE
00BW0874_2	QKWNKTLQV	GKKLAEHFPN	...KTIIFN. .SSSGGDLEI TTHSFNCGGE
00BW1471_2	SNWNSTLQV	ARKLEKYFPN	...KTIKFO. .PSSGGDLEI TTHSFNCRGE
00BW1616_2	TKWENTLHMV	SEKLKENFPN	...KTIVFN. .SSAGGDLEI TTHSFNCRGE
00BW1686_8	SKWNETLQKV	KKKLGEHFPD	...KNITFE. .SPSGGDLEI TTHSFNCRGE
00BW1759_3	GEWNETLMEV	SKELRKYFPN	...KNITFA. .PSSGGDLEI TTHSFNCRGE
00BW1773_2	AQWNKTLQEV	GAKLEEHPN	...KTIKFN. .QSSGGDLEI TTHSFICRGE
00BW1783_5	KAWNKTLLHRV	SEXLKEHFNP	...KTIKFT. .SPSGGDLEI IQHTFNCGGE
00BW1795_6	SKWITTLHRV	SEKLKEHFNP	...KAINFT. .QPKGGDLEI TTHSFNCRGE
00BW1811_3	TNWNKTLQMV	SEKLQQHFPN	...KTIKFD. .KHSGGDLEI TTHSFNCRGE
00BW1859_5	EEWNKTLQGV	EEKLEHFNP	...KKITFK. .PSSGGDLEI TTHSFNCGGE
00BW1880_2	DVWRKTLFNV	SNKLKEYFPK	...RNITFN. .SSTGGDLEI TTHSFNCRGE
00BW1921_1	GAWNKTLLQKV	GEELRKHFNP	...KTIQFN. .SSSGGDLEI TTHSFNCRGE
00BW2036_1	SAWNRTLHRV	SKKLREHFNP	...TTIKFO. .PASGGDLEI TTHSFNCRGE
00BW2063_6	EKWNKTLYRV	SEKLKEYFPN	...KTIKFD. .QPTGGDLEI TKHSFNCRGE
00BW2087_2	DKWNKTLQV	GEKLAEHFPN	...KTIKFA. .PSSGGDLGI TTHSFNCRGE
00BW2127_2	DAWNETLQV	GKKLEEFNP	...KTIKFA. .NSSGGDLEI TTHSFNCGGE
00BW2128_3	EEWNKTLREV	KGKLGKHFNP	...KTIMFA. .PSSGGDLEI TAHSFNCRGE
00BW2276_7	NQWNETLQV	GKKLKEHFNP	...KTIKFE. .QSSGGDLEI TMHSFNCRGE
00BW3819_3	KAWNKTLLQV	VKKLKEHFNP	...RTIKFT. .PPPGGNLEI TTHSFNCRGE
00BW3842_8	GNWTKTLQV	SRKLRLQIFNK	...SNIEFK. .PHSGGDPEI TTHSFNCRGE
00BW3871_3	HEWNKTLQV	KKKLGEHFNP	...KTIKFO. .PSSGGDLEI ATHTFNCRGE
00BW3876_9	SDWNRTLQV	KNKLGEHFNP	...TVIKFE. .PSSGGDLEI TTHSFNCRGE
00BW3886_8	GNWTETLRRV	KNKLGEHFNP	...KTITFG. .PSSGGDLEI TTHSFNCRGE
00BW3891_6	STWNRTLQEV	SRKLVERFPN	...KTIRFO. .PPSGGDLEI TTHSFNCRGE
00BW3970_2	ADWNRTLQGV	GRKLAGYFPN	...KTISFO. .PSSGGDLEI TTHSFNCGGE
00BW5031_1	TKWTETLQKI	SEKLRGYFK.	...KTIIFA. .PSSGGDPEI TTHSFNCRGE
96BW01B21	EVWKKTLQV	GRKLKEHFNP	...KTIQFO. .PPSGGDLEI TTHSFNCRGE
96BW0407	KDWNKTLHRV	RKKLAEHFS.	...KNITFK. .PSSGGDLEI TTHSFNCGGE
96BW0502	TEWNSTLQGV	SKKLEEHFSK	...KAIKCE. .PSSGGDLEI TTHSFNCRGE
96BW06_J4	TNWNKTLKGV	KGELRKHFNP	...KNITFO. .PASGGDLEI TTHSFNCRGE
96BW11_06	SQWNKTLHRV	IEKLKEHFNP	...KTIGFS. .QAAGGDLEI TTHSFNCGGE
96BW1210	GAWNRTLQV	GKKLKEHFNP	...KTIRFK. .ESSGGDLEI TTHSFNCGGE
96BW15B03	GEWNKAVQV	SAKLREHFNP	...KTIEFO. .PSSGGDLEI TTHSFNCRGE
96BW16_26	SEWKRTLQV	SEKLGKHFNP	...KTIKFA. .PHSGGDLEI TTHVLNCRGE
96BW17A09	.WNNTLQV	VKKLRKHFPN	...KTITFA. .PPSGGDLEI TTHSFNCGGE
96BWM01_5	SKWNKILYRV	SEKLKEHFNP	...KTIQFG. .QPIGGDLEI TTHSFNCRGE
96BWM03_2	GEWAKVMQV	TGKLKEHFP.	...KNITFO. .PPSGGDLEI TTHSFNCRGE
98BWM012_2	QKWNKTLQV	GKKLAEFNP	...RTIAFN. .SSARGDLEI ATHSFNCRGE
98BWM013_4	SLWNETLYKV	SEKLKEYFN.	...TTIEFO. .QPAGGDLEI TTHSFNCRGE
98BWM014_a	EEWNNSLQV	AKKLREHFNP	...KTIAFN. .SSSGGDLEI TTHSFNCRGE
98BWM014_1	GNWTKTLHRV	GEKLKEHFNP	...KTIKFA. .PPSGGDLEI IMHSFNCGGE
98BWM018_d	DDWKSTLQV	SEKLRRHFNP	...KTIVFN. .SPSGGDLEI ITHSFNCEGE
98BWM036_a	ERWNKTLQEV	GEKLERHFNP	...KTIEFK. .PSSGGDLEI TTHSFNCRGE
98BWM037_d	KKWYKALHGV	REKLKVLFPN	...KNISFO. .PAAGGDLEI TTHSFNCKGE
99BW3932_1	DAWNKTLQV	GRKLKEYFPG	...STIRFA. .PHSGGDLEI TTHSFNCRGE
99BW4642_4	SEWNRTLQV	GEKLGKYFPS	...KTIKFN. .SSSGGDLEI TTHSFNCRGE
99BW4745_8	KAWNRTLQEV	GRKLAEHFPN	...KTIKFO. .PHSGGDPEI TMHSFTCGGE
99BW4754_7	SQWNKTLQEV	SEKLREKFPN	...KTIQFN. .SSTGGDLEI TTHSFNCGGE
99BWM016_8	EAWNKTLLRV	SKKLREHFNP	...KTIIFD. .KSSGGDLEI TTHSFNCGGE
A2_CD_97CD	TEWNATLKKV	VEQLREHFNP	...KTIIFN. .SSSGGDLEI TTHSFNCGGE
A2_CY_94CY	TLWNETLQV	AEQLREKFPK	...KTIIFT. .NSSGGDPEI TTLSFNCRGE
A2D_97KR	RQWNETLQV	AEQLRKYFSN	...KTIIFT. .NSSGGDIEI TTHSVNCGGE
A2G_CD_97C	AGWNETLQV	AEQLGKHFSG	...KNITFA. .NSSGGDLEI TTHSFNCGGE
A_BY_97BL0	AAXNSTLQKI	STQLRKYFNN	...KTIIFK. .SSTGXDLEI TTHSFNCGGE
A_KE_Q23_A	SRWNKTLQEV	AEKLRTYFGN	...KTIIFA. .NSSGGDLEI TTHSFNCGGE
A_SE_SE659	SEWNNTLQV	AKQLRTYFGN	...KTIIFT. .NSSGGDLEI TTHSFNCKGE
A_SE_SE725	SSWNKTLQDI	VTQLRVYWN.	...RTIIFN. .SSSGGDLEI TTHSFNCGGE
A_SE_SE753	SKWNATLQV	AIKLREYFDD	...KTIIFT. .KPSGGDLEI TTHSFNCGGE

A_SE_SE853	AKWNKTLHEV	AKQLRTYFNN	...KTIIIFT.	.NSSGGDLEI	THTVNCGGE
A_SE_SE889	TEWNEALQKV	VNQLKTHFKN	...KTIIIFN.	.SSSGGDLEI	TTHSFNCGGE
A_SE_UGSE8	SKWNETLKKV	AIQLRKYWN.	...TTIIIFT.	.NSSGGDLEI	TTHSFNCGGE
A_UG_92UG0	SQWNKTLHQV	VEQLRKYWNN	...NTIIIFN.	.SSSGGDLEI	TTHSFNCAGE
A_UG_U455	RDWNRTIQOV	AEQLKKKFNN	...KTIIIFA.	.SSSGGDIEI	TTHSFNCGGE
AC_IN_2130	AEWNKALNKI	GKQLRKYFVN	...KTIKFA.	.NSSGGDLEI	TTHSFNCEGE
AC_RW_92RW	TKWNRTLQKV	AEKLSHYFEN	..ITTIIFK.	.NSSGGDLEI	TTHSFNCGGE
AC_SE_SE94	TKWNKTLHKV	VTQLRKYFVN	...KPIIFT.	.PSSGGDVEV	TTHSFNCRGE
ACD_SE_SE8	SEWNKTLQOV	AKKLGDPLNK	...TEIIIFK.	.PPSGGDLEI	TTHSFNCGGE
ACG_BE_VI1	KEWNKTLQAV	GKKLAEYYPN	...KTINFT.	.QASGGDLEI	VTHSFNCGGE
AD_SE_SE69	TKWDKTLREV	AIQLKHYYG.	..NKTIVIFAN	.SS.GGDIEI	TTHSFNCRGE
AD_SE_SE71	SAWNNTLQOV	VIQLRRYFNN	...KTIIIFT.	.NSSGGDLEI	TTHSFNCGGE
ADHK_NO_97	GSWMKTLHKV	ATQLXQHFS.	..NKTIIIFNA	.SA.GGDIEI	TTHSFNCAGE
ADK_CD_MAL	TEWDKTLQOV	AVKLGSLLN.	..KTKIIIFNS	.SS.GGDPEI	TTHSFNCRGE
AG_BE_VI11	KDWGKMLQEV	SRQLKKFFNN	...KTIFFNS	.SA.GGDLEI	TTHSFNCRGE
AG_NG_92NG	QEWQEMLOKV	QAQLEQVFN.	...KSITFNS	.SA.GGDLEI	TTHSFNCRGE
AGHU_GA_VI	EQWNRTLERV	KEKLGRHFK.	..NKTITFKP	.AS.GGDPEV	TMHIFNCRGE
AGU_CD_Z32	KEWSETLSKV	AAQLRKHFVN	T.RTDIIIFA.	.NSSGGDVEI	TTHSFNCGGE
AJ_BW_BW21	KNWTDTLHKV	TAKLKEYFN.	...TTIEFQP	.AS.AGDLEI	MTHTFNCGGE
B_AU_VH_AF	TNWTSVLRQI	AVKLRERFK.	..NKTIVFNH	.SS.GGDPEI	VRHSFNCGGE
B_CN_RL42	TKWNNTLKQI	TKKLREQFG.	..NKTIVFNQ	.SS.GGDPEI	VMHSFNCGGE
B_DE_D31_U	AKWDSTLRQI	VKKLRERFG.	..NKTIVFNQ	.SS.GGDPEI	VTHSFNCGGE
B_DE_HAN_U	ARWNKTLNQI	FRKLREIRQF	.ENKTIVFNR	.SS.GGDPEI	VMHSFNCGGE
B_FR_HXB2	AKWNNTLKQI	ASKLREQFG.	..NNKTIIIFKQ	.SS.GGDPEI	VTHSFNCGGE
B_GA_OYI	ATWEKTLEQI	ATKLARKQFR.	.N.KTIAFDR	.SS.GGDPEI	VMHSFNCGGE
B_GB_CAM1	TKWNNTLKQI	VTKLKEQFG.	..NKTIIIFNQ	.SS.GGDPEI	VMHSFNCGGE
B_GB_GB8_C	EKWNNTLHQI	VIELRKQFR.	..NKTIVFNQ	.SS.GGDPEI	VMHSFNCGGE
B_GB_MANC	AKWEKTLKQI	VEKLREKFG.	..NKTIIIFNQ	.SS.GGDPEI	VTHSFNCGGE
B_KR_WK_AF	EKWNNTLKQI	VIKLGEQFG.	..NSNIIVFKQ	.SS.GGDPEI	VMHSFICGGE
B_NL_3202A	AKWNNTLKQI	VSKLRKQFG.	..NKTIVFSQ	.PL.GGDPEI	VMHSFNCGGE
B_TW_TWCYS	AEWNNTLPQI	VKKFREQFG.	..NKTIVFNQ	.SS.GGDLEI	VMHSFNCGGE
B_US_BC_L0	AKWNNTLRQI	VIKLREQF.	.ENKTIVFNQ	.SS.GGDPEI	VMHSFNCGGE
B_US_DH123	VKWHNTLKR	VEKLREKFE.	..NKTIVFNK	.SS.GGDPEI	VMHSFNCGGE
B_US_JRCSF	AQWNNTLKQI	VEKLREQF.	..NNKTIVFTH	.SS.GGDPEI	VMHSFNCGGE
B_US_MNCG	AKWNNTLRQI	VSKLKEQFK.	..NKTIVFNQ	.SS.GGDPEI	VMHSFNCGGE
B_US_P896	AKWNNTLQOI	VIKLREKF.	.RNKTIAFNQ	.SS.GGDPEI	VMHSFNCGGE
B_US_RF_M1	AQWNNTLKQV	VTKLREQF.	.DNKTIVFTH	.SS.GGDPEI	VLHSFNCGGE
B_US_SF2_K	AQWNNTLEQI	VKKLREQFG.	..NNKTIVFNQ	.SS.GGDPEI	VMHSFNCRGE
B_US_WEAU1	TSWNNTLKQI	VEKLREIKQF	.KNKTIVFKQ	.SS.GGDPEI	VMHSFNCGGE
B_US_WR27	TKWKNTLEKI	VAKIREIKQF	.KNKTIVFNH	.SS.GGDPEI	VMHSFNCGGE
B_US_YU2_M	TQWENTLEQI	AIKLKEQFG.	..NNKTIIIFNP	.SS.GGDPEI	VTHSFNCGGE
BF1_BR_93B	TKWNETLEKV	RAKLKPHFPN	...ATIKFNS	.SS.GGDLEI	TMHSFNCRGE
C_BR_92BR0	TAWNNTLQEV	GKKLAEHFPN	...KAIFKA.	.KHSGGDLEI	TTHSFNCRGE
C_BW_96BW0	RDWNNTLNRV	SKKLAEHFPN	...KTIEFK.	.PSSGGDLEI	TTHSFNCRGE
C_BW_96BW1	SQWNNTLQRV	SEKLKEHFPN	...KTIFKN.	.QPAGGDLEI	TTHSFNCGGE
C_BW_96BW1	GAWNNTLQV	GKKLKEHFPN	...KTIRFK.	.ESSGGDLEI	TTHSFNCGGE
C_BW_96BW1	GEWNKAVQV	SAKLREHFPN	...KTIEFQ.	.PSSGGDLEI	TTHSFNCRGE
C_ET_ETH22	EKWNNTLQKV	KEKLQKHFPN	...KTIEFK.	.PSSGGDLEI	TTHSFNCGGE
C_IN_93IN1	DKWNETLQV	GKKLAEHFHN	...KTIKFA.	.SSSGGDLEI	TTHSFNCRGE
C_IN_93IN9	ENWTDTLQV	SKKLAEHFPN	...KTIKFD.	.SPSGGDLEI	TTHSFNCRGE
C_IN_93IN9	DRWNNTLQV	GEKLAEHFPN	...KTIKFA.	.PSSGGDLEI	TTHSFNCRGE
C_IN_94IN1	RDWNNTLQV	SEKLAKHFPN	...KTIKFA.	.PSSGGDLEI	TTHSFNCRGE
C_IN_95IN2	DKWNETLQV	SKKLAEHFPN	...KTIIIFN.	.SSSGGDLEI	TTHSFNCRGE
CRF01_AE_C	TKWNETLKQV	TKKLREHFPN	...KTIIIFQ.	.PSSGGDPEI	TMHHFNCRGE
CRF01_AE_C	TKWKETLKQV	TRKLREHLNG	...TMTISFR	.PSSGGDPEI	TMHHFNCRGE
CRF01_AE_C	TKWNETLQOI	IRKLEEHFNN	...KTIIQFKP	.PSSGGDLEI	TMHHFNCRGE
CRF01_AE_T	TKWNKVLKQV	TEKLKEHFPN	...KTIIIFQ.	.PPSGGDLEI	TMHHFNCRGE
CRF01_AE_T	TKWNEVLKQV	AGKLKEHFPN	...KTIIIFK.	.PPSGGDLEI	TMHHFNCRGE
CRF01_AE_T	TKWNKVLNQV	TEKLKEHFPN	...RNISFQ.	.PPSGGDLEI	TMHHFICRGE
CRF01_AE_T	TKWNETLKQV	AGKLREHFPN	...KTIIIFQ.	.PPSGGDLEI	TMHHFNCRGE
CRF01_AE_T	TKWNKVLKQV	TEKLKEHFN.	...KTIIIFQ.	.PPSGGDLEI	TMHHFNCRGE

CRF01_AE_T	TKWNKVLKQV	TEKLKEHFN.	...KTIIFQ.	.PPSGGDLEI	TMHHFNCRGG
CRF02_AG_F	SEWNRTLQQV	ATQLRKHFN.	...KTIIFA.	.NSSGGDIEI	TTHSFNCGGE
CRF02_AG_F	SKWNNTLQQV	AIQLRKHFN.	...TTIIFA.	.NPSGGDIEI	TTHSFNCGGE
CRF02_AG_G	TDWNTTLQQV	ATQLGKYFRD	T..TRIKFD.	.NPSGGDLEI	MTHSFNCGGE
CRF02_AG_N	TEWNKTLHQV	VTQLKTYFKN	...TTIIFA.	.NPLGGDVEI	TTHSFNCGGE
CRF02_AG_S	QQWNKTLHDV	ATKLREYFNN	...TTIIFD.	.EPSGGDLEI	TTHSFNCGGE
CRF02_AG_S	EKWNSTLQKV	VTKLKGKHFNS	...SKIIFT.	.NSSGGDLEI	TTHSFNCGGE
CRF03_AB_R	TKWNNTLKQI	VIKLRKQFG.	..NKTIVFNQ	.SS.GGDPEI	VMHSFNCGGE
CRF03_AB_R	TKWNNTLEQI	VSKLRKQFR.	..NKTIVFNQ	.SS.GGDPEI	VMHSFNCGGE
CRF04_cpx_	NDWNTTLKVI	SEELKRLFP.	..NKTIKFAP	.PV.GGDLEI	TTHSFNCKGE
CRF04_cpx_	SDWNEALQKV	VVKLREHFP.	..NKTIIFNQ	.SS.GGDLEI	TTHSFNCGGE
CRF04_cpx_	KDWNNTLQKI	VDELKRFHP.	..NKNITFAP	.SA.GGDVEI	TTHSFRLGGE
CRF05_DF_B	EQWNKTLIQV	AKELQSHFP.	..NKTIKFNS	.SS.GGDLEI	TMHSFNCRGE
CRF05_DF_B	AQWNKTLQV	KEELRAHIKD	IGNKTIVFNS	.SA.GGDLEI	TSHIFNCRGE
CRF06_cpx_	ANWTDILGEV	KVKLEEVFNN	...THITFKS	.SA.GGDLEI	TTHSFNCGGE
CRF06_cpx_	KAWNSMLQNV	TAKLKELFNN	...KNITFNS	.SA.GGDLEV	TTHSFNCGGE
CRF06_cpx_	TAWKETLQNV	TEKLKQLLN.	...TNITFNP	.SA.GGDLEI	TTHSFNCRGE
CRF06_cpx_	TDWNNMLKNV	TTKLIEVFK.	...KNITFNS	.SA.GGDLEI	TTHSFNCGGE
CRF11_cpx_	AEWLNTLQQV	ATQLRGKFN.	...KTIIFDN	.PSPGGDIEI	TSHSFNCRGE
CRF11_cpx_	ADWNNTLQQV	AEQLHNNFN.	...KTIVFNE	.HS.GGDLEV	TTHSFNCGGE
D_CD_84ZR0	VKWNNTLRQV	ARKLGNLLN.	..QTKIIFKP	.SS.GGDPEI	TTHSFNCGGE
D_CD_ELI_K	AQWSKTLQQV	ARKLGTLLN.	..KTIIFKP	.SS.GGDPEI	TTHSFNCGGE
D_CD_NDK_M	AEWNKALQV	ATKLGNLLN.	..KTTITFKP	.SS.GGDPEI	TSHMLNCGGD
D_UG_94UG1	AGWNKTLQV	AEKLGNLLN.	..QTTIIFKP	.SS.GGDPEI	TTHSFNCGGE
F1_BE_VI85	TQWNNTLEYV	KAELKSHFPN	N..TAIKFNQ	.SS.GGDLEI	TMHSFNCRGE
F1_BR_93BR	TQWRNTLAKV	KAKLGSYFPN	...ATIKFNS	.SS.GGDLEI	TRHNFNCMGE
F1_FI_FIN9	EQWNKTLDRV	KAELKLHFNK	...TIQFNS	.SS.GGDLEI	TMHSFNCRGE
F1_FR_MP41	TQWSKTKTQV	QEKLRALFNK	...TIKFNQ	.SS.GGDLEI	TMHSFNCRGE
F2_CM_MP25	KQWYDTLIKI	ATEFKDQYN.	...KTVGFQP	.SA.GGDLEI	TTHSFNCRGE
F2KU_BE_VI	ENWNKTLEGV	KAKLHGFFTN	...KTIIFKP	.HS.GGDPEV	VMHTFNCGGE
G_BE_DRCBL	TKWNETLRDV	OAKLQEYFIN	...KSIEFNS	.SS.GGDLEI	TTHSFNCGGE
G_NG_92NG0	IKWREMLKNV	TAQLRKIYN.	..NKNITFNS	.SA.GGDLEI	TTHSFNCRGE
G_SE_SE616	RKWKEALQNV	AAELGKIFNK	S.SENITFNS	.SA.GGDLEI	TTHSFICRGE
H_BE_VI991	KQWNETLHKV	ITKLGSYFD.	..NKTIILOP	.PA.GGDIEI	ITHSFNCGGE
H_BE_VI997	EKWNKTLQOI	ATQLSKYFV.	..NRTLIFKP	.HS.GGDLEV	TTHSFNCRGE
H_CF_90CF0	TDWNKTLHQV	VTQLGIHLN.	..NRTISFKP	.NS.GGDMEV	RTHSFNCRGE
J_SE_SE702	KDWNNTLRRV	AKKLREHFN.	...KTIDFTS	.PS.GGDIEI	TTHSFNCGGE
J_SE_SE788	RDWSNTLRRV	ATKLREHFN.	...KTINFTS	.PS.GGDIEI	VTHSFNCGGE
K_CD_EQTB1	GQWNKTVNQV	KKELGKHFN.	...KTIIFQP	.SS.GGDPQV	TRHIFNCRGE
K_CM_MP535	EKWNMTLSRV	KEKLKEHFN	...GTITFKP	.PNPGGDPEI	LTHMFNCAGE
N_CM_YBF30	ELWEPMWNRT	REEIKKILGK	...NNITFRA	RERNEGDLEV	THLMFNCRGE
O_CM_ANT70	TDWGKILKQT	AERYLELVNN	TGSINMTFN.	.HSSGGDLEV	THLHFNCHGE
O_CM_MVP51	TVWENALQQT	AIRYLNLVNQ	TENVTIIFS.	.RTSGGDAEV	SHLHFNCHGE
O_SN_99SE	SDWEKALKQT	AERYLDLRNN	TNTVNITFE.	.RSIGGDSEV	THLHFNCHGE
O_SN_99SE	SVWEEALKQT	AERYLELMNN	TNTVNITFN.	.HSTGGDPEV	THLHFNCHGE
U_CD_83C	GEWRNTLQV	AIALRRQFNN	...KSIIFN.	.SSSGGDIEI	TTHTFNCGGE

	451		500
00BW0762_1	FFYCNTTRLF	NGTYN.....	STGD TNS.....TN STITLQCRIK
00BW0768_2	FFYCDTSNLF	NKTRR.....	DN.. ..AN ETITLPCRIK
00BW0874_2	FFYCNTSRLF	NSTYN.....	PNST YIEGR...SN ATITLQCRIK
00BW1471_2	FFYCYTTKLF	NSTYN.....	STYT GSESN..... ITIPCRIC
00BW1616_2	FFYCNTSKLF	NGTYN.....	SNNN TA..... DITLQCRIK
00BW1686_8	FFYCNTSNLS	NETYL.....	ANLT SNVTK....N ATITLPCRIK
00BW1759_3	FFYCNTSNLF	NNTYR.....	ADNN ITNDNSN... ITLQCRIK
00BW1773_2	FFYCNTSALF	NSTYN.....	STNT SGHN....DT RIITLPCRIK
00BW1783_5	FFYCNTSKLF	NGTYN.....	GTS. ...ISS...N SSITLQCRIK
00BW1795_6	FFYCNTSELF	NGTYN.....	STG. ...DSN...S NLITLQCRIK
00BW1811_3	FFYCNTSQLF	NGTYM.....	PNTY MS....SSDN RNITIPCRIC
00BW1859_5	FFYCNTTHLF	NGNG.....	ESD INITLPCRIK
00BW1880_2	FFYCDTTKLF	NGTYN.....	STEQ TN..... STITLQCRIK
00BW1921_1	FFYCNTSQLF	NGTYN.....	DT.Y ESNSG....N STITLPCRIK
00BW2036_1	FFYCDTSKLF	NSSYN.....	DTL YSYNS....T ANITLPCRLK
00BW2063_6	FFYCNTSQLF	NSSYS.....	RHN. ...NTS...N STITLPCNIK
00BW2087_2	FFYCNTSGLF	N.....	GTF NGT...HSTN TNITLPCRIK
00BW2127_2	FFYCNTTILF	NSTYY.....	P... NTK...SDTT ETITLPCRIK
00BW2128_3	FFYCNTSLLF	DETQL.....	SKE. N NTINIQCRIK
00BW2276_7	FFYCNTSKLF	NGTYM.....	PNYN TSN...SSNN SNITLPCRIK
00BW3819_3	FFYCNTSGLF	NGTYN.....	G... TND...NDTD SDITLPCRIK
00BW3842_8	FFYCNTSLLF	NSSYN.....	GNSS YNDTGS...N STITLQCRIK
00BW3871_3	FFYCNTSILF	NDTYW.....	FNGT ANDTG....S NNITIPCRIC
00BW3876_9	FFYCNTSGLF	NNNLI.....	NNG. AE DTIRLPCRIK
00BW3886_8	FFYCNTSKLF	NSTNN.....	NTE. ..SES....N ATITLPCRIK
00BW3891_6	FFYCNISRLF	NRPNM.....	TKNM TSDIKNN... STITLPCRIK
00BW3970_2	FFYCNTSSLF	NNTYR.....	PTYW PGTE....SN STITLQCRIK
00BW5031_1	FFYCNTSQLF	NSTYR.....	ANTS NS..... NITLPCRIK
96BW01B21	FFYCDTSELF	NSTYM.....	SNGG NISS.....S TIIMLPCRIE
96BW0407	FFYCNTSRLF	NESYN.....	FDES YWN.N...TN KTIMLPCRIK
96BW0502	FFYCDTSQLF	NSTYS.....	PSNG TENK....LN GTITITCRIC
96BW06_J4	FFYCNTSRLF	DETYL.....	S... GTDED....N GTITLPCRIK
96BW11_06	FFYCNTSKLF	NSTYI.....	QLN. .STETP...N STITLPCRIK
96BW1210	FFYCNTSQLF	NSTYN.....	Y MPS...NNTG TNITLQCRIK
96BW15B03	FFYCNSSKLL	NSSYN.....	GTSY RGTESN...S SIITLPCRIK
96BW16_26	FFYCNTSKLF	NSTYN.....	STDR SNN.....T DNITIQCRIK
96BW17A09	FFYCNTSILF	NSTYN.....	STYT GSDSNS.... TITIPCRIC
96BWM01_5	FIYCNTSKLF	NGTYN.....	STG. TS...N STITLSCRIC
96BWM03_2	FFYCNTSELF	NGTYN.....	GTD. ..NNS....N KTITLLCRIC
98BWMC12_2	FFYCNTSGLF	NSTYN.....	PNST YTESK...AN SNITLHCRIC
98BWMC13_4	FFYCNTTKLF	NGTYS.....	QPN. .STGTP...H SNITLPCRIK
98BWMC14_a	FFYCNTSQLF	NSTYN.....	G... RNSTT....N ATITLPCRIK
98BWM014_1	FFYCNTSKLF	NSTYN.....	ATY NST...DTSN STITIPCRIC
98BWM018_d	FFYCNTSGLF	NS.....	AFNDN...SG GTITLQCRIE
98BWM036_a	FFYCNTSGLF	NSTYY.....	SNKT SSN...MTN EIITIPCRIC
98BWM037_d	FFYCNTSKLF	NTSWL.....	DSYI SNTG....NN SIITLPCRIK
99BW3932_1	FFYCNTSRLF	NSTYN.....	P... NTK...SNTG SWIILPCRIK
99BW4642_4	FFYCNTSKLF	TYQSN.....	TY.. VAN STITLPCRIK
99BW4745_8	FFYCNTSELF	NSTYN.....	ANTY NTATGNNS.. TTIILPCRIK
99BW4754_7	FFYCNTSKLF	NSTFN.....	SNHG DST....GN DPLTIPCRIC
99BWMC16_8	FFYCNTSNLF	NNTYY.....	PNMT NTDTK...SN LTITLPCRIK
A2_CD_97CD	FFYCNTTGLF	NSTWEN....	GTNK QNYTE...SN DTITLQCRIK
A2_CY_94CY	FFYCNTTGLF	NGTWNN....	GTWN GPYTPNN.TN GSIILPCRIK
A2D_97KR	FFYCDTSGLF	NSTWPAN...	ASRE NEEKD...R. .NVTLPCRIK
A2G_CD_97C	FFYCNTTNLF	NSTFNTT...	SLFN STGRNGTNDN TTITIPCRIC
A_BY_97BL0	FFYCNTTDLF	NSTX.....	DGTVT NSTKAN.... GTITLPCRIK
A_KE_Q23_A	FFYCNTSGLF	NSTWY.....	VNSTW NDTDSTQESN DTITLPCRIK
A_SE_SE659	FFYCNTSSLF	NSTWS.....	NDNNT QGSNSTET.K GTITLPCRIK
A_SE_SE725	FFYCNTSGLF	NSTWS.....	Q.NDT GVSNSTES.N DTIILPCRIK
A_SE_SE753	FFYCNTSGLF	NSTIL.....	NSTKM NDNASRESYD DTITLQCRIK

A_SE_SE853	FFYCNTSGLF	NSTWS.....	SNASE	PMSNSTES.N	DTITLQCRIR
A_SE_SE889	FFYCNTSGLF	NSTWN.....	GTDSM	QKLNST....	GNITLPCRIK
A_SE_UGSE8	FFYCNTSGLF	NSSWN.....	END.T	KVNYNTES.N	DTITLQCRIR
A_UG_92UG0	FFYCNTSGLF	NSTWV.....	NGTTS	STSN.....	GTITLPCRIK
A_UG_U455_	FFYCNTSGLF	NSIWN.....	GSMSN	DMGP.....N	GTITLQCRIR
AC_IN_2130	FFYCNTSGLF	NGTWNASMQ.	ES	NSTESN....	ETIILPCRIK
AC_RW_92RW	FFYCNTSGLF	NSTWS.....	KR	NGTWQSNSTE	LNITLPCRIK
AC_SE_SE94	FFYCDTSGLF	NSTWPFNS..	T	NSTGPN....	GTITLQCRIR
ACD_SE_SE8	FFYCNTSGLF	NSTWV.....	NGSRE	SNSTDN....	DTITLPCRIK
ACG_BE_VI1	FFYCNTSGLF	NSTYN.....	PSYN	STESVN...E	TTIILPCKIK
AD_SE_SE69	FFYCNTTGLF	NSTWNDTAT.	EQKP	N..	DTIRLQCRIR
AD_SE_SE71	FFYCNTSGLF	NSTWN.....	NTDSM	QESHSTET.N	DTITLPCRIK
ADHK_NO_97	FFYCNTSQLF	NSTWNHTST.	YNST	EN.....	GTITLPCRIK
ADK_CD_MAL	FFYCNTSKLF	NSTWQNNGA.	RLSN	S..TE.ST..	GSITLPCRIK
AG_BE_VI11	FFYCNTSALF	NFSSETNST.	FP.N.....		TTTLTPCRIK
AG_NG_92NG	FFYCNTSGLF	NESGGNDT..			.TITLPCRIK
AGHU_GA_VI	FFYCNTTKLF	NDTENKN...		..NDAEN...	KTITLPCRIK
AGU_CD_Z32	FFYCNTSGLF	NSTWK.....	NSTSI	NDTVSN....	GTITLPCRIK
AJ_BW_BW21	FFYCNTSGLF	NKSLLNETS.	NETT	DGAN.....	NTITLTCRIK
B_AU_VH_AF	FFYCNSTQLF	NSTWFNSTG.	NDTE	RATNN..T..	ENITLPCRIK
B_CN_RL42_	FFYCNTSQLF	NSTWNDTG..	T	WNDTTGNS..	.TITLPCRIK
B_DE_D31_U	FFYCNSAQLF	NSTWNDTK..	ES	NNTNG.....	.TITLPCRIK
B_DE_HAN_U	FFYCNSTKLF	NSTWNNTST.	WN..	DNGND.....	.TITLPCRIK
B_FR_HXB2_	FFYCNSTQLF	NSTWFNSTW.	STEG	SNNTGSD..	.TITLPCRIK
B_GA_OYI_	FFYCNTSQLF	NSTWNDTTR.	AN..	.STEV.....	.TITLPCRIK
B_GB_CAM1_	FFYCNTTQLF	NTTWLFNGT.	WNDT	EGLNNTER..	.NITLPCRIK
B_GB_GB8_C	FFYCKTAQLF	NSTWNSTGN.	GTIK	SNTTE.....	.IITLPCRIK
B_GB_MANC_	FFYCNSTQLF	NSTWNTGND.	TRES	NDTNN..T..	GNITLPCRIK
B_KR_WK_AF	FFYCNTTQLS	NSTWQRSDG.	TWN	TGGLNETK..	ENITLPCRIK
B_NL_3202A	FFYCNSTQLF	NSTWNDTGN.	VTER	SNNNE.....	.NITLPCRIK
B_TW_TWCYS	FFYCNATPLF	NSTWNATST.	LNAT	NEENE.....	.NITLLCRIK
B_US_BC_L0	FFYCKSTQLF	NSTWAGNNT.	WNSS	AERSDDTG..	GNITLPCRIK
B_US_DH123	FFYCNTKKLF	NSTWNGTEG.	SYNI	EGND.....	.TITLPCRIK
B_US_JRCSE	FFYCNSTQLF	NSTWNDTEK.	SSG.	TEGND.....	.TITLPCRIK
B_US_MNCG_	FFYCNTSPLF	NSTWNGNNT.	WNNT	TGSNN.....	.NITLQCKIK
B_US_P896_	FFYCNTAQLF	NSTWNVTTG.	TNG.	TEGND.....	.IITLQCRIR
B_US_RF_M1	FFYCNTTQLF	NSTWNSTEG.	SNNT	GGND.....	.TITLPCRIK
B_US_SF2_K	FFYCNTTQLF	NNTWRLNHT.	EG..	TKGND.....	.TITLPCRIK
B_US_WEAU1	FFYCNSTQLF	NSTWHANGT.	WKNT	EGADN.....	.NITLPCRIK
B_US_WR27_	FFYCNSTQLF	NSTWNSTEG.	NS..	TWSDK.....	.IIRLPCRIK
B_US_YU2_M	FFYCNSTQLF	..TWNDTRK.	LN..	.NTGR.....	.NITLPCRIK
BF1_BR_93B	FFYCNTSGLF	NDTVDN....			GTITLPCRIK
C_BR_92BR0	FFYCNTSSLF	NSTYT.....	PNST	ENITGT..EN	SIITIPCRIK
C_BW_96BW0	FFYCNTSRLF	NESYS.....	FNES	HWSND...TN	ATITLPCRIK
C_BW_96BW1	FFYCNTSKLF	NGTYI.....	QPNS	.TEDTP...N	STITLPCRIK
C_BW_96BW1	FFYCNTSQLF	NSTYN.....	S.TY	MPS...NNTG	TNITLQCRIR
C_BW_96BW1	FFYCNSSKLL	NSSYN.....	GTSY	RGTESN...S	SIITLPCRIK
C_ET_ETH22	FFYCNTSNLF	NSTKL.....	E...	LFNSS...TN	LNITLQCRIR
C_IN_93IN1	FFYCNTSGLF	NGTYM.....	PTYM	PNGTESN.SN	STITIPCRIK
C_IN_93IN9	FFYCNTSGLF	NGTYN.....	TSSD	GNS.....S	STITIPCRIK
C_IN_93IN9	FFYCNTSSLF	DSLFN.....	PNGT	RNDT.....N	LTITIPCRIK
C_IN_94IN1	FFYCNTSGLF	NSTYM.....	SGTY	MNSSADM.NS	SYITIPCRIK
C_IN_95IN2	FFYCNTSGLF	NRTYM.....	PNDT	KSNSSSN.PN	ANITIPCRIK
CRF01_AE_C	FFYCNTTKLF	NSTWT.....	TNE	IMEEFKGTNS	STITLPCRIK
CRF01_AE_C	FFYCNTTALF	NSTWI.....	N.G	TMQEVNGTNS	GNITLPCRIK
CRF01_AE_C	FFYCNTTRLF	N.....		..ISTNGTTN	GTITLPCRIK
CRF01_AE_T	FFYCNTTQLF	NNTCI.....	GNE	TMK...GCNG	.TITLPCRIK
CRF01_AE_T	FFYCNTTQLF	NSTWT.....	GNE	TME...GSNG	.TITLPCRIK
CRF01_AE_T	FFYCNTTRLF	NNTCI.....	GNK	TMK...ECND	.TITLPCRIK
CRF01_AE_T	FFYCNTTKLF	NSTWI.....	GNE	TIG...SSG	.NIILPCRIK
CRF01_AE_T	FFYCNTTKLF	NNTCL.....	GNE	TMA...GCND	.TITLPCRIK

CRF01_AE_T	FFYCNTTKLF	NSTWR.....	GNE	TIESREGYNK	.TIILPCKIK
CRF02_AG_F	FFYCNTSELF	NSTW.....	..NSTWDNSS	NHIESNHT.E	GNITLQCRIK
CRF02_AG_F	FFYCNTSELF	N.....	...STWDNSL	NHTESNHT.E	DNITLQCRIK
CRF02_AG_G	FFYCNTSGLF	NSTWYKN...	..STWYSNST	ASSNHTEL.N	STITLQCKIK
CRF02_AG_N	FFYCNTSKLF	N.....	...STWDNSN	STANHTGS.N	DTITLQCRIK
CRF02_AG_S	FFYCNTSNLF	NRTWNHNGTW	NAPGPFNDTE	DKTINGTE.D	KTITLQCRIK
CRF02_AG_S	FFYCNTAELF	NSTWASN...	.TNGIWASNI	NASNKNKA.N	DTITLKCKIK
CRF03_AB_R	FFYCNTTKLF	NSTWNGTEE.	LN..	.NTEG.....	DIVTLPCRIK
CRF03_AB_R	FFYCNTTKLF	NSTWNNTTEE.	SN..	.NTKG.....	DIVTLPCRIK
CRF04_cpx_	FFYCNTTPLF	NSTHMQNGT.	NIT.	S.TDSTN...	STITLQCRLK
CRF04_cpx_	FFYCNTSGLF	NSTYMFNST.	NRTN	T.TNGTN...	STITLPCRIK
CRF04_cpx_	FFYCNTSDLF	NRTYMVNKN.	ETNS	T.NTTDE...	KIIRLPCRIK
CRF05_DF_B	FFYCDTSKLF	NATVFNDTV.	FNAT	MFNND...SD	KNIILPCKIK
CRF05_DF_B	FFYCNTSGLF	NVTVP.....		..NNE.....	.TITLPCRIK
CRF06_cpx_	FFYCNTSNLF	NTSDLFNTS.	R..G	NDTN.....	TTITLPCRIK
CRF06_cpx_	FFYCNTSQLF	NNNITDSNE.		...T.....	TNFTLPCRIK
CRF06_cpx_	FFYCNTSQLF	NSSIPESENE.		...T.....	DIITLPCRIK
CRF06_cpx_	FFYCNTSQLF	NSSNLNNNS.		SDNN.....	GTITLPCRIN
CRF11_cpx_	FFYCNTSGLF	NNTWLFNST.	WNSS	QELNGT...E	PNITLPCRIK
CRF11_cpx_	FFYCNTSGLF	NSTWYANDN.	TSTQ	NDMQSN...D	.TITLPCRIK
D_CD_84ZR0	FFYCNTSGLF	NSAWNISGH.	STGL	N.....D..	TIITIPCRIC
D_CD_ELI_K	FFYCNTSGLF	NSTWNISAW.	NNIT	ESNNS.TN..	TNITLQCRIK
D_CD_NDK_M	FFYCNTSRLF	NSTWNQTN.	TGFN	N..	GTVTLPCRIK
D_UG_94UG1	FFYCNTTRLF	NSTWKRNNS.	EWRS	D..NT.PD..	ETITLQCRIK
F1_BE_VI85	FFYCDTSGLF	NDTGSN....		N	GTITLPCRIK
F1_BR_93BR	FFYCNTDELF	NDTKFND...		...TG...FN	GTITLPCRIK
F1_FI_FIN9	FFYCNTSLLF	NNTVPN....		N	GTITLPCRIK
F1_FR_MP41	FFYCDTSGLF	NESEKY....		N	GTIILPCKIK
F2_CM_MP25	FFYCNTTILF	NHTRVNDIL.	SNNH	TR.....EN	DTITLPCRIK
F2KU_BE_VI	FFYCNTTRLF	NDTLNHT...		ID	QNITLPCRIK
G_BE_DRCBL	FFYCNTSGLF	NNSILKSNI.		SENN.....	DTITLNCKIK
G_NG_92NG0	FFYCNTSGLF	NNNISNIN..		...N.....	ETITLPCRIK
G_SE_SE616	FFYCNTSGLF	NSSLLRSNS.		SE.N.....	GTITLPCRIK
H_BE_VI991	FFYCNTTKLF	NSTWTNSSY.	TNDT	YNSNSTEDIT	GNITLQCKIK
H_BE_VI997	FFYCNTSGLF	NSSWTGDNI.	NMPN	DTG.....	KNITLPCRIK
H_CF_90CF0	FFYCNTSGLF	NSSWEMHTN.	YTSN	DTKG...N..	ENITLPCRIK
J_SE_SE702	FFYCNTSTLF	NSSWDENNI.	KDTN	STNDN.....	TTITIPCKIK
J_SE_SE788	FLYCNTSKLF	NSSWDKNSI.	EATN	DTSX.....	ATITIPCKIK
K_CD_EQTB1	FSYCDTDTV	DDTEEE....		ED	TTITIPCRIC
K_CM_MP535	FFYCNTTKLF	NETGE.....		N	GTITLPCRIK
N_CM_YBF30	FFYCNTSKLF	NEELLN....	ETG.		EPITLPCRIR
O_CM_ANT70	FFYCNTAKMF	NYTFS.....	CNGTTC	SVSNVSQ.G.	NNGTLPCKLR
O_CM_MVP51	FFYCNTSGMF	NYTFIN....	CTKSGC	QEIKGSNETN	KNGTIPCKLR
O_SN_99SE_	FFYCNTSKMF	NYTFS.....	CIGTNC	TSNQNSSNS.	NDTRIYCRIK
O_SN_99SE_	FFYCNTSQMF	NYTFS.....	CTRNC	IRQSNSS...	INGTISCRIK
U_CD_83C	FFYCNTSELF	TGIWNG....	TWDK	NCTSTESNCT	GNITLPCRIK

DEMANDE OU BREVET VOLUMINEUX

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

CECI EST LE TOME 1 DE 3
CONTENANT LES PAGES 1 À 208

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

JUMBO APPLICATIONS/PATENTS

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

THIS IS VOLUME 1 OF 3
CONTAINING PAGES 1 TO 208

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

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

WHAT IS CLAIMED IS:

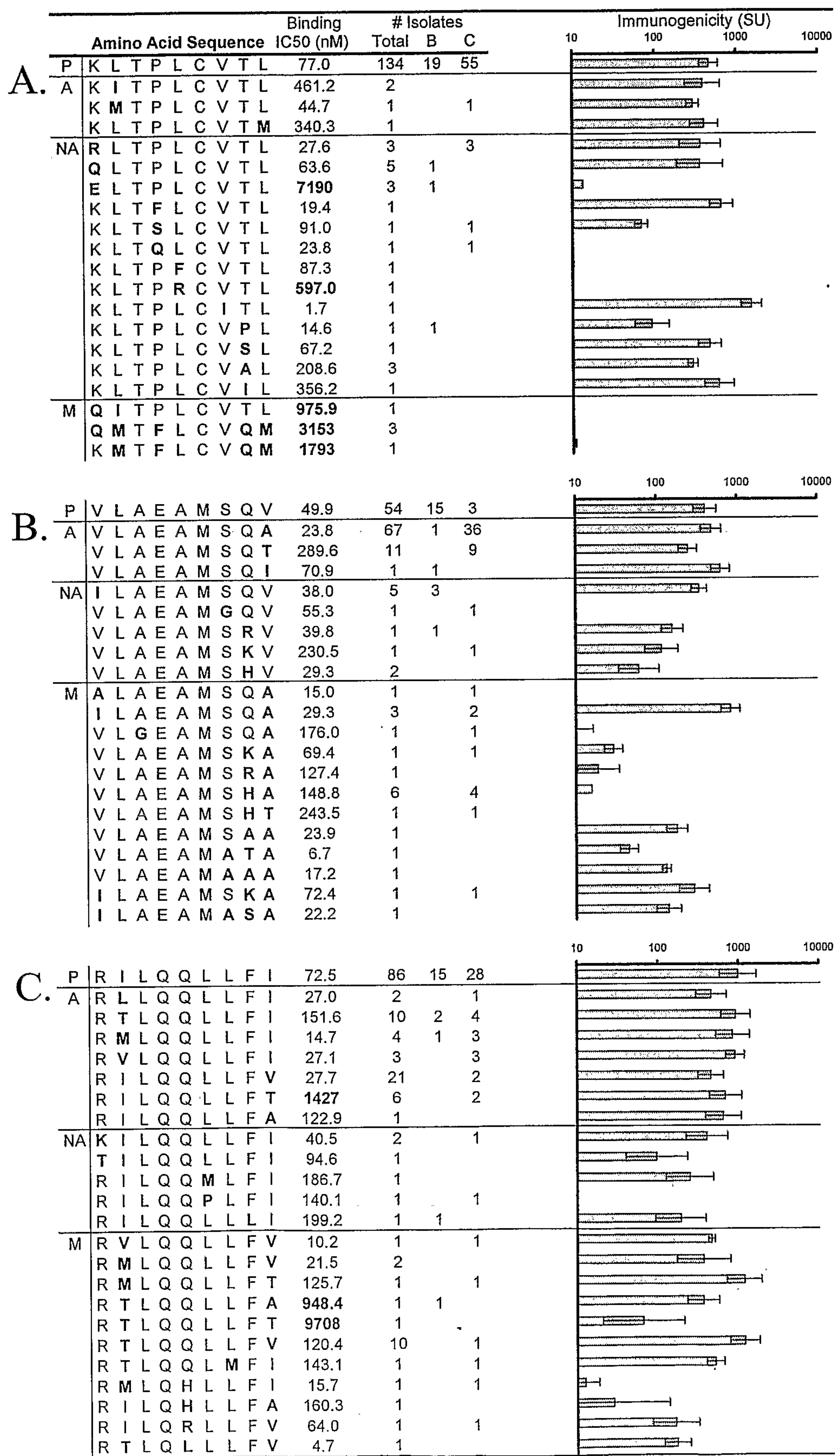
1. A method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising
 - a) identifying, from a particular antigen of an infectious agent, variants of a peptide epitope 8-11 amino acids in length, each variant comprising primary anchor residues of the same HLA class I binding motif; and
 - b) determining whether one of said variants comprises only conserved non-anchor residues in comparison to at least one remaining variant, thereby identifying a candidate peptide epitope.
2. A method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising
 - a) identifying, from a particular antigen of an infectious agent, variants of a peptide epitope 8-11 amino acids in length, each variant comprising primary anchor residues of the same HLA class I binding motif;
 - b) determining whether each of said variants comprises conserved, semi-conserved or non-conserved non-anchor residues in comparison to each of the remaining variants; and
 - c) identifying a variant which comprises only conserved non-anchor residues in comparison to at least one remaining variant.
3. A method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising
 - a) identifying, from a particular antigen of an infectious agent, a population of variants of a peptide epitope 8-11 amino acids in length, each peptide epitope comprising primary anchor residues of the same HLA class I binding motif;
 - b) choosing a variant selected from the group consisting of:
 - i) a variant which comprises preferred primary anchor residues of said motif; and
 - ii) a variant which occurs with high frequency within the population of variants; and

- c) determining whether the variant of (b) comprises only conserved non-anchor residues in comparison to at least one remaining variant, thereby identifying a candidate peptide epitope.
4. A method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising
- a) identifying, from a particular antigen of an infectious agent, a population of variants of a peptide epitope 8-11 amino acids in length, each peptide epitope comprising primary anchor residues of the same HLA class I binding motif;
 - b) choosing a variant selected from the group consisting of:
 - i) a variant which comprises preferred primary anchor residues of said motif; and
 - ii) a variant which occurs with high frequency within the population of variants; and
 - c) determining whether the variant of (b) comprises conserved, semi-conserved or non-conserved non-anchor residues in comparison to each of the remaining variants; and
 - d) identifying a variant which comprises only conserved non-anchor residues in comparison to at least one remaining variant.
5. The method of claim 1, wherein (b) comprises identifying a variant which comprises only conserved non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.
6. The method of claim 2 or 3, wherein (c) comprises identifying a variant which comprises only conservative non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.
7. The method of claim 4, wherein (d) comprises identifying a variant which comprises only conservative non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

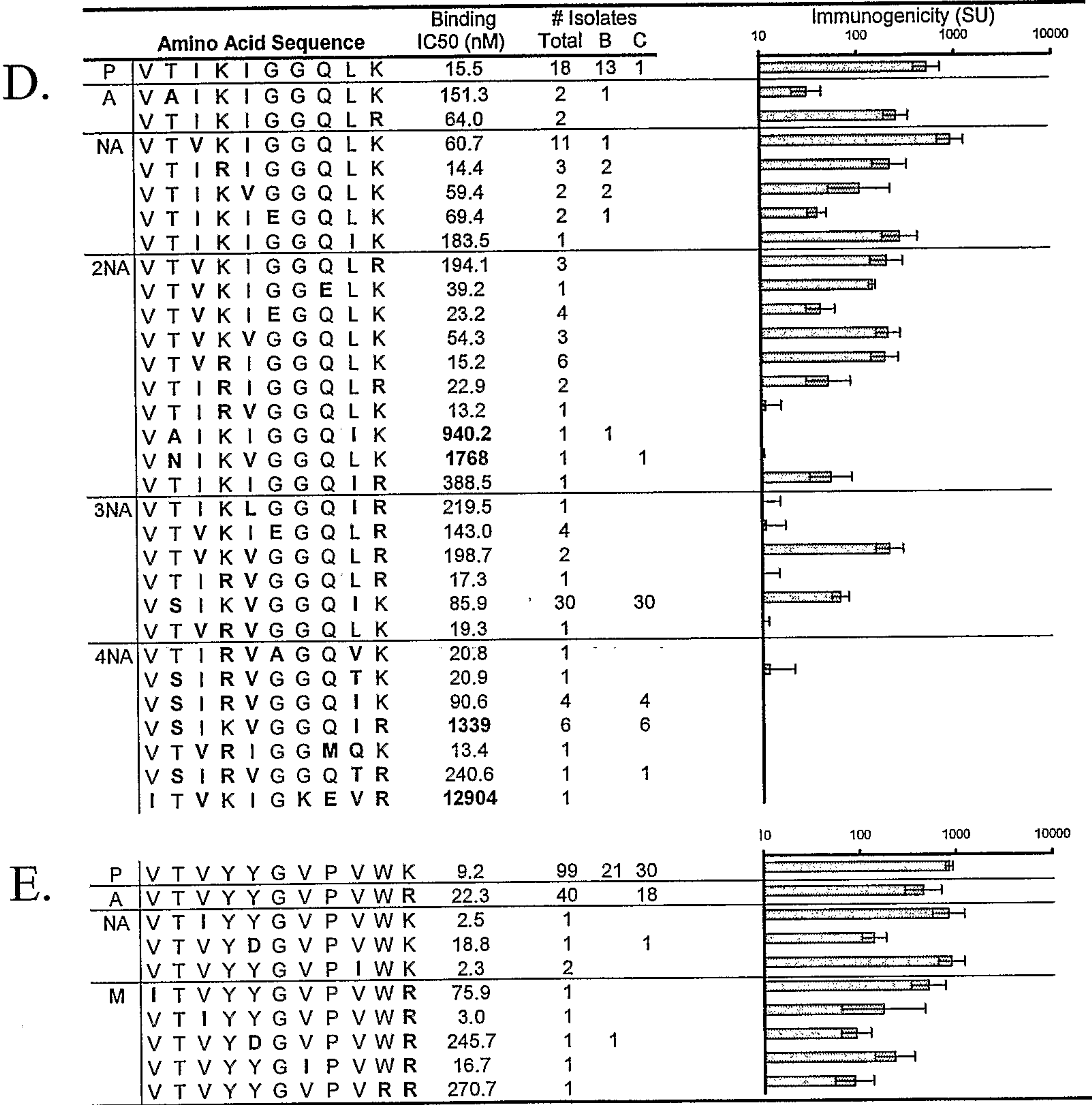
8. The method of any of claims 1-4, wherein (a) comprises aligning the sequences of said antigens.
9. The method of claim 3 or 4, wherein (b) comprises choosing a variant which comprises preferred primary anchor residues of said motif.
10. The method of claim 3 or 4, wherein (b) comprises choosing a variant which occurs with high frequency within said population.
11. The method of claim 10, wherein (b) comprises ranking said variants by frequency of occurrence within said population.
12. The method of claim 3 or 4 wherein (b) comprises choosing a variant which comprises preferred primary anchor residues of said motif and which occurs with high frequency within said population.
13. The method of claim 12, wherein (b) comprises ranking said variants by frequency of occurrence within said population.
14. The method of any of claims 1-13, wherein the identified variant comprises the fewest conserved anchor residues in comparison to each of the remaining variants.
15. The method of any of claims 1-4, wherein the remaining variants comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 27, 28, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 220, 240, 260, 280, or 300 variants.
16. The method of any of claims 1-15, wherein the infectious agent is selected from the group consisting of: HIV, HBV, HCV, HPV, *Plasmodium falciparum*, Influenza virus, and Dengue virus, Epstein-Barr virus, *Mycobacterium tuberculosis*, *Chlamydia*, *Candida albicans*, *Cryptococcus neoformans*, *Coccidioides spp.*, *Histoplasma spp.*, *Aspergillus fumigatis*, *Plasmodium spp.*, *Trypanosoma spp.*, *Schistosoma spp.*, and *Leishmania spp.*
17. The method of claim 16, wherein the infectious agent is selected from the group consisting of: HIV, HBV, HCV, HPV, *Plasmodium falciparum*, Influenza virus, and Dengue virus.
18. The method of claim 16, wherein the infectious agent is HIV and the antigen is selected from the group consisting of: Gag, Env, Pol, Nef, Rev, Tat, Vif, Vpr, and Vpu.

19. The method of claim 16, wherein the infectious agent is HBV and the antigen is selected from the group consisting of: Pol, Env, Core, and NS1/Env2.
20. The method of claim 16, wherein the infectious agent is HCV and the antigen is selected from the group consisting of: Core, E1, E2, NS1, NS2, NS3, NS4, and NS5.
21. The method of claim 16, wherein the infectious agent is HPV and the antigen is selected from the group consisting of: E1, E2, E3, E4, E5, E6, E7, L1, and L2.
22. The method of claim 16, wherein the infectious agent is *Plasmodium falciparum* and the antigen is selected from the group consisting of: CSP, SSP2, EXP1, LSA1.
23. The method of any claims 1-4, wherein the selected variant and the at least one remaining variant comprise different primary anchor residues of the same motif or supermotif.
24. The method of any of claims 1-4, wherein the motif or supermotif is selected from the group consisting of those in Tables 1-2.
25. The method of any of claims 1-4, wherein the conserved non-anchor residues are at any of positions 3-7 of said variant.
26. The method of any of claims 1-4, wherein the variant comprises only 1-3 conserved non-anchor residues compared to at least one remaining variant.
27. The method of any of claims 26, wherein the variant comprises only 1-2 conserved non-anchor residues compared to at least one remaining variant.
28. The method of any of claims 27, wherein the variant comprises only 1 conserved non-anchor residue compared to at least one remaining variant.
29. The method of claim 16, wherein the infectious agent is HPV, and further wherein, the HPV infectious agent is selected from the group consisting of HPV strains 16, 18, 31, 33, 45, 52, 56, and 58.
30. The method of any of claims 1-29, wherein the variants are a population of naturally occurring variants.

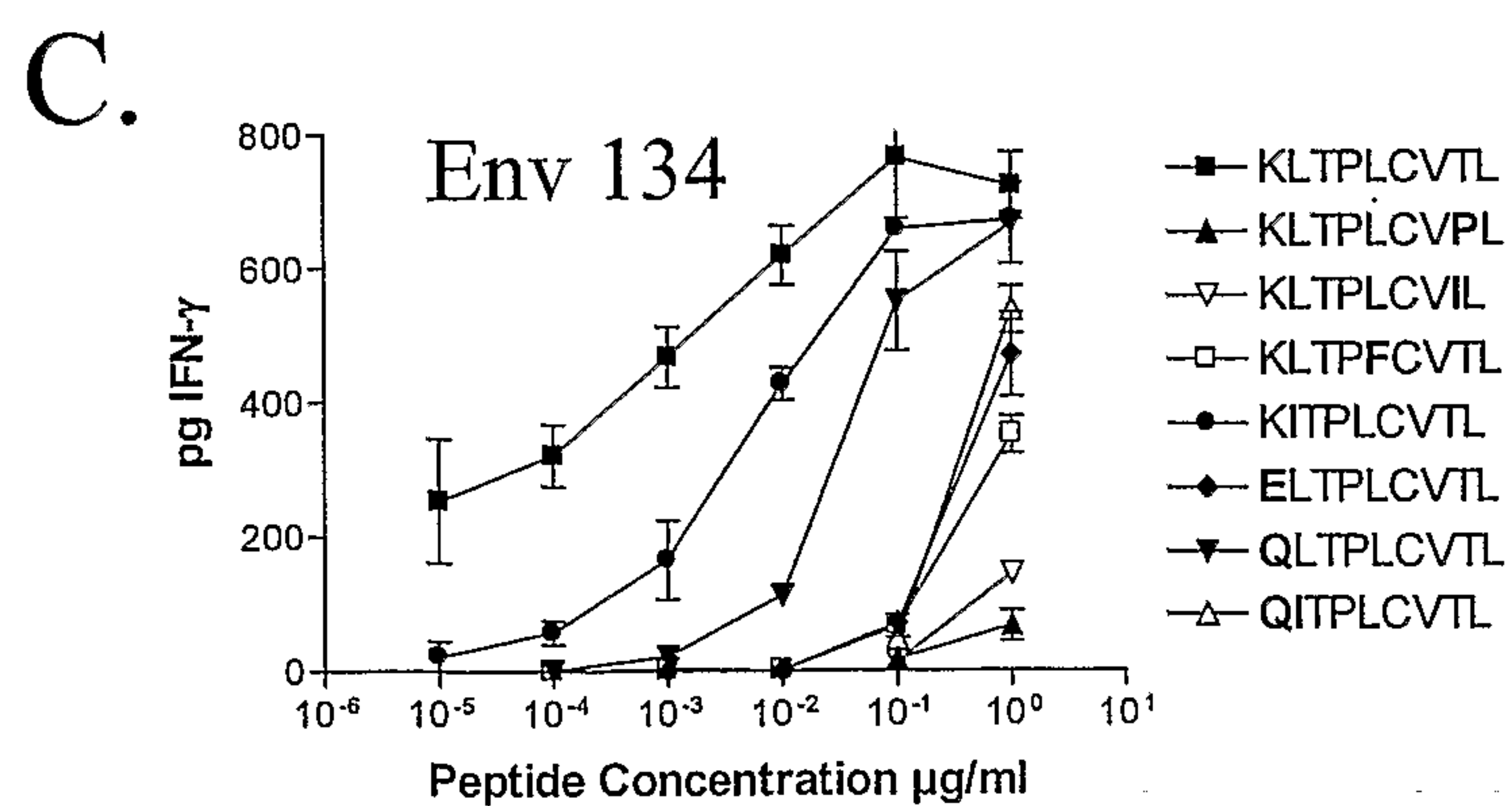
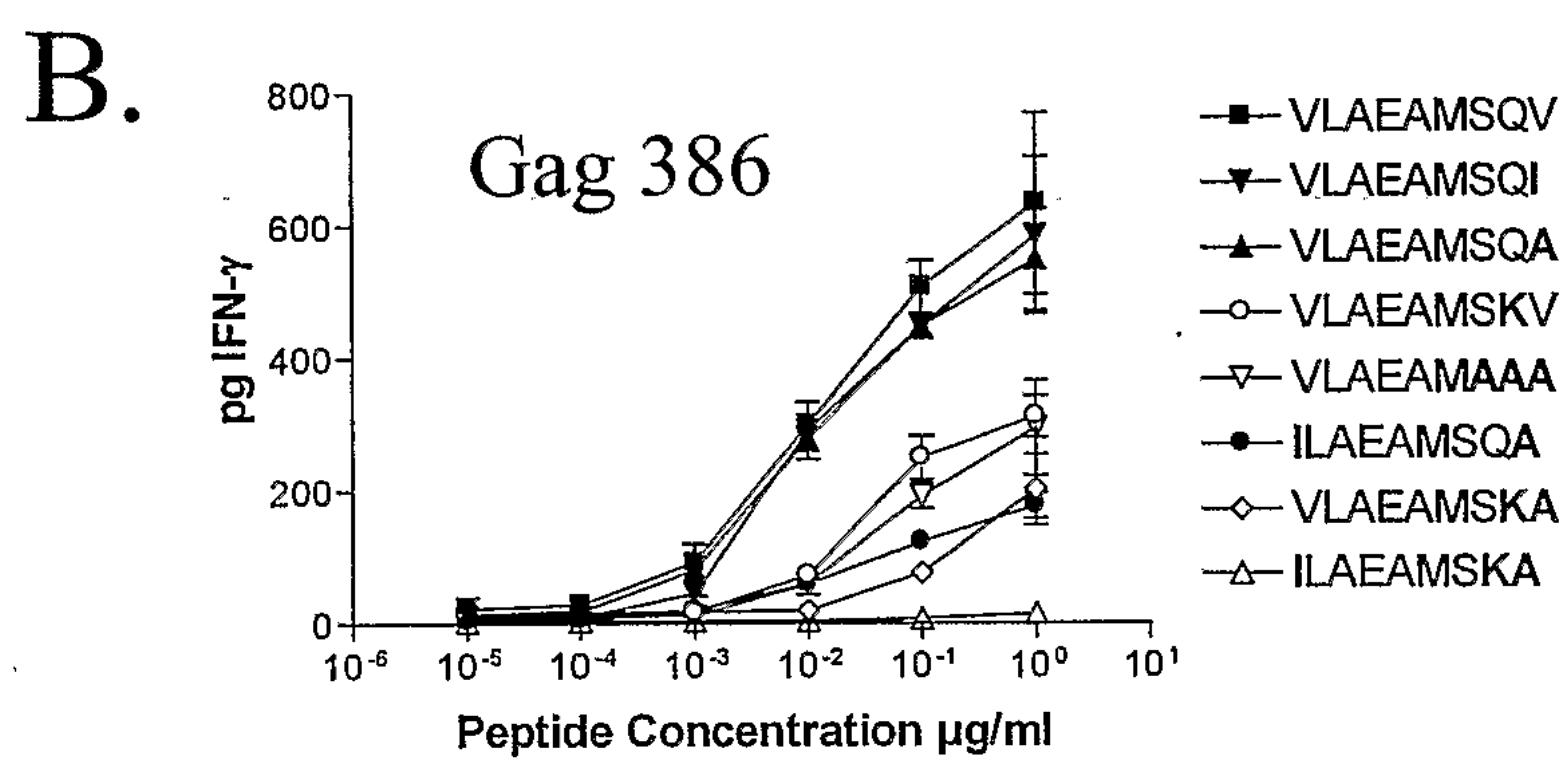
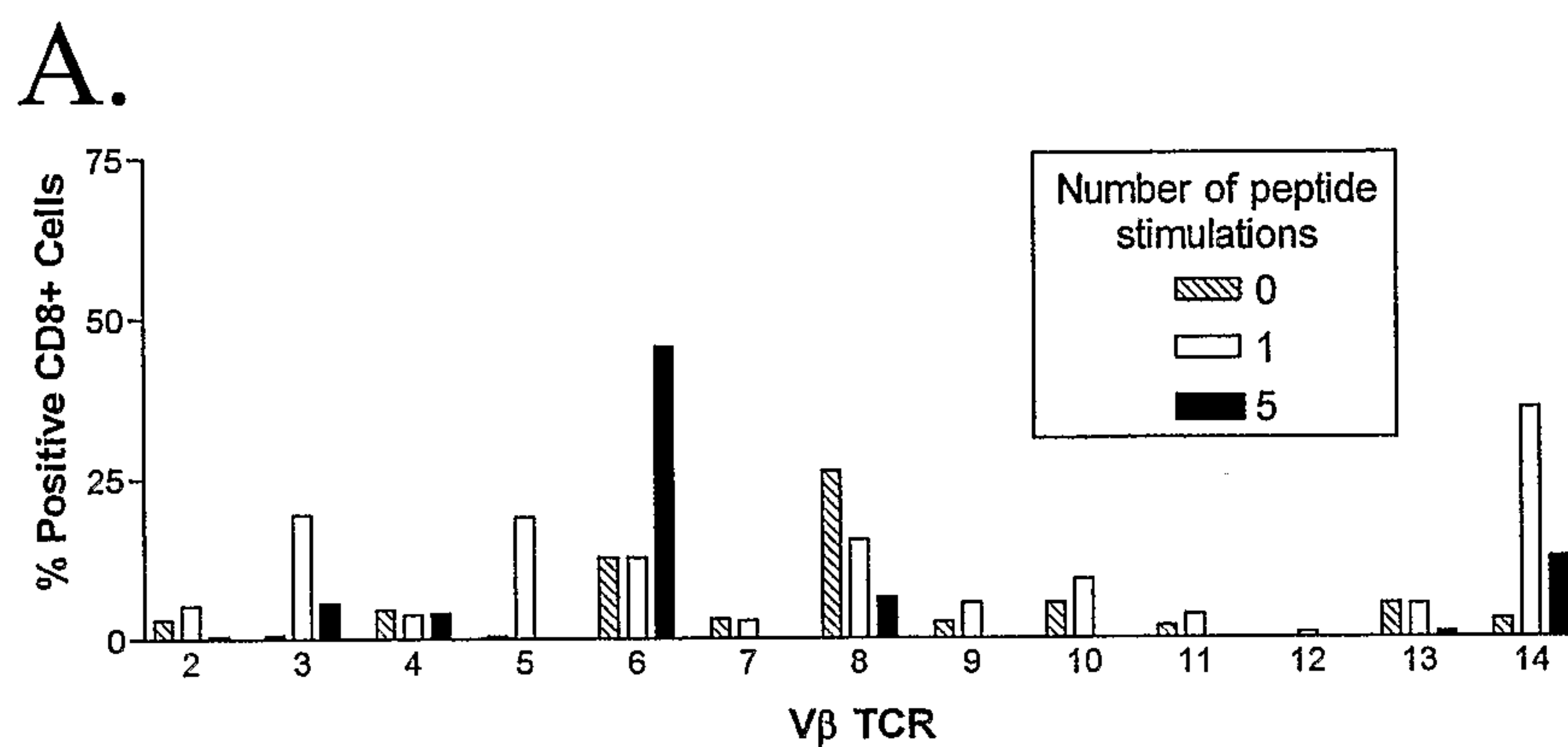
FIGS. 1A-1C



FIGS. 1D-1E



FIGS. 2A-2C



FIGS. 3A-3B

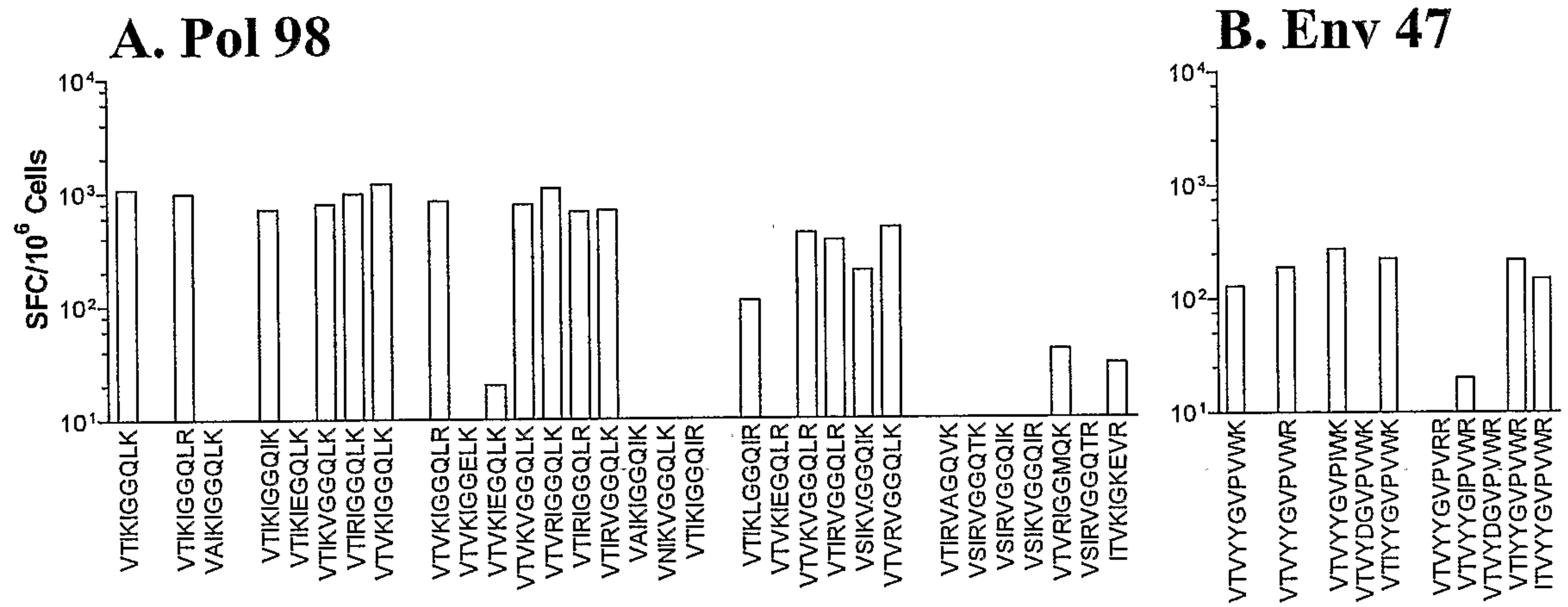


FIG. 4

