



(51) International Patent Classification:

G01N 33/50 (2006.01) *G01N 33/53* (2006.01)
G01N 33/52 (2006.01) *G01N 33/493* (2006.01)

(21) International Application Number:

PCT/US2014/053676

(22) International Filing Date:

2 September 2014 (02.09.2014)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/881,767 24 September 2013 (24.09.2013) US

(71) Applicant: DREXEL UNIVERSITY [US/US]; 3141 Chestnut Street, Philadelphia, Pennsylvania 19104 (US).

(72) Inventors: SCHLECHT, Hans Peter; 901 N. Penn Street, R902, Philadelphia, Pennsylvania 19123 (US). VOSSELLER, Keith; 2104 Green Street, Apartment #3, Philadelphia, Pennsylvania 19130 (US).

(74) Agents: DOYLE, Kathryn et al.; 1500 Market Street, Suite 3800, Philadelphia, Pennsylvania 19102 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: NOVEL NON-INVASIVE METHODS OF MONITORING HIV VIRAL LOADS

(57) Abstract: The present invention provides a method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient, comprising analyzing a sample comprising urine from the patient for the presence and/or concentration of at least one protein.



WO 2015/047669 A1

TITLE OF THE INVENTION

Novel Non-Invasive Methods of Monitoring HIV Viral Loads

5

CROSS-REFERENCE TO RELATED APPLICATIONS

The present application claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Patent Application No. 61/881,767, filed September 24, 2013, which application is incorporated herein by reference in its entirety.

10

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR
DEVELOPMENT

This invention was made with government support under 1R03AI083149-01A1 and 5R03AI083149-02 awarded by National Institute of Allergy and Infectious Diseases (National Institutes of Health). The government has certain rights in the invention.

15

BACKGROUND OF THE INVENTION

Viral load testing (*i.e.*, measuring the number of copies of HIV in the blood) is the only way to accurately assess the level of viral replication in HIV-infected patients. Routine monitoring of viral load helps reinforce a patient's adherence to anti-retroviral therapy (ART), thereby ensuring viral suppression and preventing treatment failure before it occurs. Routine testing also ensures that health care workers can diagnose treatment failure early on when drug resistance occurs, and appropriately switch patients from first-line ART to more effective second-line treatment regimens. With large numbers of patients throughout the world already on treatment for several years, ensuring patients can be tested for viral load is a global priority. Furthermore, viral load monitoring is a critical component of programs that aim to reduce transmission rates.

20

25

For patients on ART, the World Health Organization (WHO) recommends viral load testing twice yearly. Unfortunately, viral load testing remains largely unavailable in resource-limited settings, in which the majority of HIV-infected patients reside. Viral load testing is rarely available or convenient in poor countries, resulting in avoidable morbidity and mortality and increasing the risk of transmission of drug-resistant forms of the virus.

30

It is thus critical that access to viral load testing in resource-limited settings be prioritized as part of the fight against HIV/AIDS. Current viral load tests are fairly complex,

requiring specialized laboratory facilities. Unfortunately, the majority of HIV-infected patients rely on points of service without reliable power supply or highly trained staff. In such cases, transport of samples to central reference laboratories is unfeasible and/or cost-prohibitive. Further, a lack of market competition for viral load testing kits results in high testing costs. Simple tests that can be performed at a community-based clinics, and/or a point-of-care test that can be performed at a point of service, are now urgently needed throughout the world.

There is a need in the art for novel convenient and effective methods of identifying and/or monitoring patients with (un)controlled HIV infection. Such methods may be used to determine whether the patient is responding to anti-retroviral therapy. The present invention fulfills this need.

BRIEF SUMMARY OF THE INVENTION

The invention includes a method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient. The invention also includes a kit for assessing or monitoring systemic HIV viral load in an HIV-infected human patient.

In certain embodiments, the method includes analyzing a test sample comprising urine from the patient for the presence or concentration of at least one protein, whereby a test data set is obtained.

In certain embodiments, the methods includes comparing the test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample.

In certain embodiments, the methods allows for assessing and/or monitoring the HIV viral load in the patient.

In certain embodiments, the patient has received or is receiving a first anti-HIV medication. In other embodiments, the patient is a new-born human or an infant younger than about 18 months of age.

In certain embodiments, the test sample is prepared by a method comprising subjecting urine from the patient to at least one procedure selected from the group consisting of protein isolation and protein digestion. In other embodiments, the test sample is analyzed using mass spectrometry, a quantum dot assay or a chromophore assay. In yet other embodiments, the test sample is analyzed using a method comprising contacting the test sample with an antibody or aptamer. In yet other embodiments, the antibody is at least one selected from the group consisting of a polyclonal antibody, monoclonal antibody, Fv, Fab,

F(ab)₂, single chain antibody, human antibody, humanized antibody, and fragments and derivatives thereof. In yet other embodiments, the antibody or aptamer is used in an immunoassay. In yet other embodiments, the immunoassay comprises at least one selected from the group consisting of immunoturbidimetry, immunonephelometry, ELISA assay, radioimmunoassay, chemiluminescence immunoassay, immunofluorescence, immunoprecipitation, immunoelectrophoresis, and flow cytometry-based immunoassay.

In certain embodiments, the control sample comprises an urine sample from an untreated HIV-infected control human. In other embodiments, the untreated HIV-infected control human is the human patient before receiving anti-HIV medication. In yet other embodiments, the control sample comprises an urine sample from an HIV-uninfected control human. In yet other embodiments, the control sample comprises an urine sample from an HIV-infected control human with controlled infection.

In certain embodiments, the concentration of the protein in the patient's urine is higher by at least a multiplicity factor than the concentration of the protein in the urine from an HIV-uninfected control human or from an HIV-infected control human with controlled infection, wherein the patient is identified as having uncontrolled HIV infection, whereby the patient is prescribed a second anti-HIV medication that is distinct from the first anti-HIV medication. In other embodiments, the multiplicity factor is selected from the group consisting of about 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 30, 40, 50, 75, 100, 125, 250, 500 and 1,000.

In certain embodiments, the concentration of the protein in the patient's urine is lower by at least a multiplicity factor than the concentration of the protein in the urine from an HIV-uninfected control human or from an HIV-infected control human with controlled infection, wherein the patient is identified as having controlled HIV infection, whereby the patient continues to be prescribed the first anti-HIV medication. In other embodiments, the multiplicity factor is selected from the group consisting of about 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 30, 40, 50, 75, 100, 125, 250, 500 and 1,000.

In certain embodiments, the concentration of the protein in the patient's urine is equal to or greater than a multiplicity factor of the concentration of the protein in the urine from an untreated HIV-positive control human, wherein the patient is identified as having uncontrolled HIV infection, whereby the patient is prescribed a second anti-HIV medication which is distinct from the first anti-HIV medication. In other embodiments, the multiplicity factor is selected from the group consisting of about 1, 0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3,

0.2, 0.1, 0.05, 0.025, 0.01, 0.005, 0.0025, 0.001, 0.0005, 0.00025, 0.0001, 0.00005 and 0.00001.

In certain embodiments, the concentration of the protein in the patient's sample is lower than a multiplicity factor of the concentration of the protein in the urine from an untreated HIV-positive control human, wherein the patient is identified as having controlled HIV infection, whereby the patient continues to be prescribed the first anti-HIV medication. In other embodiments, the multiplicity factor is selected from the group consisting of about 0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.025, 0.01, 0.005, 0.0025, 0.001, 0.0005, 0.00025, 0.0001, 0.00005 and 0.00001.

In certain embodiments, the at least one protein has an accession number selected from the group consisting of Q8TD57 (SEQ ID NO:1), Q18PE1 (SEQ ID NO:2), Q8NFH5 (SEQ ID NO:3), Q8WYL5 (SEQ ID NO:4), Q8IYD8 (SEQ ID NO:5), O14654 (SEQ ID NO:6), Q96AP4 (SEQ ID NO:7), Q9UQ35 (SEQ ID NO:8), Q8N6W0 (SEQ ID NO:9), Q9H792 (SEQ ID NO:10), Q9H497 (SEQ ID NO:11), Q9UE35 (SEQ ID NO:12), O00743 (SEQ ID NO:13), Q8WXF8 (SEQ ID NO:14), P81274 (SEQ ID NO:15), Q8NG08 (SEQ ID NO:16), Q96AE7 (SEQ ID NO:17), Q9BZM4 (SEQ ID NO:18), Q5T2D3 (SEQ ID NO:19), Q8IXT5 (SEQ ID NO:20), Q9P225 (SEQ ID NO:21), and Q9Y2I9 (SEQ ID NO:22).

In certain embodiments, the at least one protein has an accession number selected from the group consisting of P41222 (PTGDS) (SEQ ID NO:23), P14151 (SELL) (SEQ ID NO:24), Q06418 (TYRO3) (SEQ ID NO:25), P52306 (RAP1GDS1) (SEQ ID NO:26), and Q9Y5Y7 (LYVE1) (SEQ ID NO:27).

In certain embodiments, the kit includes an antibody or aptamer that binds to at least one protein with an accession number selected from the group consisting of Q8TD57 (SEQ ID NO:1), Q18PE1 (SEQ ID NO:2), Q8NFH5 (SEQ ID NO:3), Q8WYL5 (SEQ ID NO:4), Q8IYD8 (SEQ ID NO:5), O14654 (SEQ ID NO:6), Q96AP4 (SEQ ID NO:7), Q9UQ35 (SEQ ID NO:8), Q8N6W0 (SEQ ID NO:9), Q9H792 (SEQ ID NO:10), Q9H497 (SEQ ID NO:11), Q9UE35 (SEQ ID NO:12), O00743 (SEQ ID NO:13), Q8WXF8 (SEQ ID NO:14), P81274 (SEQ ID NO:15), Q8NG08 (SEQ ID NO:16), Q96AE7 (SEQ ID NO:17), Q9BZM4 (SEQ ID NO:18), Q5T2D3 (SEQ ID NO:19), Q8IXT5 (SEQ ID NO:20), Q9P225 (SEQ ID NO:21), and Q9Y2I9 (SEQ ID NO:22).

In certain embodiments, the kit includes an antibody or aptamer that binds to at least one protein with an accession number selected from the group consisting of P41222

(PTGDS) (SEQ ID NO:23), P14151 (SELL) (SEQ ID NO:24), Q06418 (TYRO3) (SEQ ID NO:25), P52306 (RAP1GDS1) (SEQ ID NO:26), and Q9Y5Y7 (LYVE1) (SEQ ID NO:27).

In certain embodiments, the kit includes an applicator. In other embodiments, the kit includes an instructional material for the use of the kit. In yet other embodiments, the instruction material comprises instructions for analyzing a test sample comprising urine from the patient for the presence or concentration of the at least one protein.

In certain embodiments, the kit further comprises a test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample. In other embodiments, the control sample comprises an urine sample from an untreated HIV-infected control human. In yet other embodiments, the control sample comprises an urine sample from an HIV-uninfected control human. In yet other embodiments, the control sample comprises an urine sample from an HIV-infected control human with controlled infection.

BRIEF DESCRIPTION OF THE DRAWINGS

The following detailed description of specific embodiments of the invention will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating the invention, there are shown in the drawings specific embodiments. It should be understood, however, that the invention is not limited to the precise arrangements and instrumentalities of the embodiments shown in the drawings.

Fig. 1 is a table illustrating characteristics of the study population in Example 1.

Fig. 2, comprising Figs. 2A-2B, is a table illustrating a selected list of proteins identified in the urine of HIV-infected patients in Example 1.

Fig. 3, comprising Figs. 3A-3F, is a table illustrating a selected list of proteins identified in the urine of HIV-infected patients in Example 2. Highlighted are proteins that are unique to HIV urine proteomes compared to non-HIV urine, as well as proteins that display greatly increased abundance in HIV urine proteomes compared to non-HIV urine. Relative abundance is reflected in the columns displaying spectral counts for each peptide/protein identified.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to the unexpected discovery of a novel, non-invasive method for monitoring or assessing HIV viral load in a human. The method

comprises analyzing an urine sample from the human for the presence and/or concentration of one or more protein markers that are associated with active systemic HIV replication. The method allows for the monitoring or assessment of systemic HIV replication and/or infection in a human and the identification of a human with uncontrolled HIV infection.

5 In certain embodiments, change in the urinary proteome, as compared to the urinary proteome of an untreated HIV-infected control human or a HIV-uninfected control human, correlates with systemic HIV replication. In other embodiments, change in the urinary proteome, as compared to the urinary proteome of an untreated HIV-infected control human or an HIV-uninfected control human, acts as a surrogate for serum HIV viral load. In
10 yet other embodiments, the urine proteome of an HIV-infected human with high serum viral loads (such as, but not limited to, equal to or greater than about 1,000 copies/mL) can be distinguished from the urine proteome of an HIV-infected human with low serum viral loads (such as, but not limited to, equal to or less than about 200 copies/mL, or equal to or less than 400 copies/mL).

15 In one aspect, the method of the invention allows for HIV treatment monitoring using a rapid point-of-care urine test. In certain embodiments, the human has been or is being administered highly active antiretroviral therapy (HAART). In other embodiments, the human has uncontrolled HIV infection. In yet other embodiments, the human has controlled HIV infection.

20 As disclosed herein, the urinary proteome in subjects with uncontrolled HIV infection was analyzed using mass spectrometry. In certain embodiments, analysis of the urine samples identified thousands of peptides corresponding to human-unique proteins. Although no HIV proteins were detected, several host proteins were found exclusively in the urine of patients infected with HIV as compared to published surveys of the non-HIV-
25 infected human urinary proteome. In certain embodiments, these HIV-specific proteomic signatures provide insights into the human physiological response to HIV infection and serve as novel HIV biomarkers in urine.

Definitions

30 Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described.

As used herein, each of the following terms has the meaning associated with it in this section.

The articles “a” and “an” are used herein to refer to one or to more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

The term “about” as used herein, when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass variations of $\pm 20\%$ or $\pm 10\%$, more preferably $\pm 5\%$, even more preferably $\pm 1\%$, and still more preferably $\pm 0.1\%$ from the specified value, as such variations are appropriate to perform the disclosed methods.

As used herein, the term “acceptable carrier” means an acceptable material, composition or carrier, such as a liquid or solid filler, stabilizer, dispersing agent, suspending agent, diluent, excipient, thickening agent, solvent or encapsulating material, involved in carrying or transporting a compound useful in the methods of the invention such that it may perform its intended function. Each carrier must be “acceptable” in the sense of being compatible with the other compounds useful in the methods of the invention, and not interfering with the method of the invention. Some examples of materials that may serve as acceptable carriers include: sugars, such as lactose, glucose and sucrose; starches, such as corn starch and potato starch; cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; powdered tragacanth; malt; gelatin; talc; excipients, such as cocoa butter and suppository waxes; oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; glycols, such as propylene glycol; polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; esters, such as ethyl oleate and ethyl laurate; agar; buffering agents, such as magnesium hydroxide and aluminum hydroxide; surface active agents; alginic acid; pyrogen-free water; isotonic saline; Ringer’s solution; ethyl alcohol; phosphate buffer solutions; and other compatible substances.

As used herein, “acceptable carrier” also includes any and all coatings, antibacterial and antifungal agents, and absorption delaying agents, and the like that are compatible with the activity of the compound useful in the methods of the invention. Supplementary active compounds may also be incorporated into the compositions. Other additional ingredients that may be included in the compositions used in the practice of the invention are known in the art and described, for example in Remington's Pharmaceutical Sciences (Genaro, Ed., Mack Publishing Co., 1985, Easton, PA), which is incorporated herein by reference.

The term “antibody” as used herein refers to an immunoglobulin molecule that specifically binds with an antigen. An antibody of the invention includes intracellularly expressed antibody, or intrabody. Antibodies can be intact immunoglobulins derived from natural sources or from recombinant sources and can be immunoreactive portions of intact immunoglobulins. Antibodies are typically tetramers of immunoglobulin molecules. The antibodies in the present invention may exist in a variety of forms including, for example, polyclonal antibodies, monoclonal antibodies, Fv, Fab and F(ab)₂, as well as single chain antibodies, human antibodies, and humanized antibodies (Harlow, *et al.*, 1999, In: Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY; Harlow, *et al.*, 1989, In: Antibodies: A Laboratory Manual, Cold Spring Harbor, New York; Houston, *et al.*, 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; Bird *et al.*, 1988, Science 242:423-426).

The term “antibody fragment” refers to a portion of an intact antibody and refers to the antigenic determining variable regions of an intact antibody. Examples of antibody fragments include, but are not limited to, Fab, Fab', F(ab')₂, and Fv fragments, linear antibodies, scFv antibodies, and multispecific antibodies formed from antibody fragments.

An “antibody heavy chain” as used herein refers to the larger of the two types of polypeptide chains present in all antibody molecules in their naturally occurring conformations.

An “antibody light chain” as used herein refers to the smaller of the two types of polypeptide chains present in all antibody molecules in their naturally occurring conformations. κ and λ light chains refer to the two major antibody light chain isotypes.

The term “antigen” or “Ag” as used herein is defined as a molecule that provokes an immune response. This immune response may involve either antibody production, or the activation of specific immunologically-competent cells, or both. The skilled artisan will understand that any macromolecule, including virtually all proteins or peptides, can serve as an antigen. Furthermore, antigens can be derived from recombinant or genomic DNA. A skilled artisan will understand that any DNA, which comprises a nucleotide sequences or a partial nucleotide sequence encoding a protein that elicits an immune response therefore encodes an “antigen” as that term is used herein. Furthermore, one skilled in the art will understand that an antigen need not be encoded solely by a full length nucleotide sequence of a gene. It is readily apparent that the present invention includes, but is not limited to, the use of partial nucleotide sequences of more than one gene and that these nucleotide sequences are arranged in various combinations to elicit the desired

immune response. Moreover, a skilled artisan will understand that an antigen need not be encoded by a “gene” at all. It is readily apparent that an antigen can be generated synthesized or can be derived from a biological sample. Such a biological sample can include, but is not limited to a tissue sample, a tumor sample, a cell or a biological fluid.

5 “Antisense” refers particularly to the nucleic acid sequence of the non-coding strand of a double stranded DNA molecule encoding a polypeptide, or to a sequence which is substantially homologous to the non-coding strand. As defined herein, an antisense sequence is complementary to the sequence of a double stranded DNA molecule encoding a polypeptide. It is not necessary that the antisense sequence be complementary solely to the coding portion of the coding strand of the DNA molecule. The antisense sequence may be complementary to regulatory sequences specified on the coding strand of a DNA molecule encoding a polypeptide, which regulatory sequences control expression of the coding sequences.

As used herein, the term “applicator” refers to any device including, but not limited to, a hypodermic syringe, a pipette, an automatic sample probe and the like, for administering the compounds and compositions of the invention.

A “constitutive” promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a cell under most or all physiological conditions of the cell.

20 The term “container” includes any receptacle for holding a composition useful within the methods of the invention. For example, in one embodiment, the container is the packaging that contains the composition. In other embodiments, the container is not the packaging that contains the composition, *i.e.*, the container is a receptacle, such as a box or vial that contains the packaged composition or unpackaged composition and the instructions for use of the composition. Moreover, packaging techniques are well known in the art. It should be understood that the instructions for use of the composition may be contained on the packaging containing the composition, and as such the instructions form an increased functional relationship to the packaged product. However, it should be understood that the instructions may contain information pertaining to a procedure that allows for implementation of a method of the invention.

As used herein, the term “controlled HIV infection” in a human refers to an HIV-infected human who is receiving HIV treatment and has low serum viral loads (such as, but not limited to, equal to or less than about 200 copies/mL, or equal to or less than about 400 copies/mL).

The term “derivative” includes any purposefully generated peptide that in its entirety, or in part, comprises an amino acid sequence substantially similar to a variable domain amino acid sequence of an antibody that binds one of the proteins contemplated in the invention. Derivatives of the antibodies of the present invention may be characterized by single or multiple amino acid substitutions, deletions, additions, or replacements. These derivatives may include: (a) derivatives in which one or more amino acid residues are substituted with conservative or non-conservative amino acids; (b) derivatives in which one or more amino acids are added; (c) derivatives in which one or more of the amino acids of the amino acid sequence used in the practice of the invention includes a substituent group; (d) derivatives in which amino acid sequences used in the practice of the invention or a portion thereof is fused to another peptide (*e.g.*, serum albumin or protein transduction domain); (e) derivatives in which one or more nonstandard amino acid residues (*e.g.*, those other than the 20 standard L-amino acids found in naturally occurring proteins) are incorporated or substituted into the amino acid sequences used in the practice of the invention; (f) derivatives in which one or more non-amino acid linking groups are incorporated into or replace a portion of the amino acids used in the practice of the invention; and (g) derivatives in which one or more amino acid is modified by glycosylation.

The term “encoding” refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (*i.e.*, rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of mRNA corresponding to that gene produces the protein in a cell or other biological system. Both the coding strand, the nucleotide sequence of which is identical to the mRNA sequence and is usually provided in sequence listings, and the non-coding strand, used as the template for transcription of a gene or cDNA, can be referred to as encoding the protein or other product of that gene or cDNA.

As used herein, the term “endogenous” refers to any material from or produced inside an organism, cell, tissue or system.

As used herein, the term “fragment,” as applied to a protein or peptide, refers to a subsequence of a larger protein or peptide. A “fragment” of a protein or peptide may be at least about 10 amino acids in length; for example, at least about 50 amino acids in length; more preferably, at least about 100 amino acids in length; even more preferably, at least

about 200 amino acids in length; particularly preferably, at least about 300 amino acids in length; and most preferably, at least about 400 amino acids in length.

The term “heterologous” as used herein is defined as DNA or RNA sequences or proteins that are derived from the different species.

5 The term “homologous” refers to the sequence similarity or sequence identity between two polypeptides or between two nucleic acid molecules. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit, *e.g.*, if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous at that position. The percent of homology between two sequences
10 is a function of the number of matching or homologous positions shared by the two sequences divided by the number of positions compared X 100. For example, if 6 of 10 of the positions in two sequences are matched or homologous then the two sequences are 60% homologous. By way of example, the DNA sequences ATTGCC and TATGGC share 50% homology. Generally, a comparison is made when two sequences are aligned to give
15 maximum homology.

 The term “immunoglobulin” or “Ig” as used herein is defined as a class of proteins that function as antibodies. Antibodies expressed by B cells are sometimes referred to as the BCR (B cell receptor) or antigen receptor. The five members included in this class of proteins are IgA, IgG, IgM, IgD, and IgE. IgA is the primary antibody that is present in
20 body secretions, such as saliva, tears, breast milk, gastrointestinal secretions and mucus secretions of the respiratory and genitourinary tracts. IgG is the most common circulating antibody. IgM is the main immunoglobulin produced in the primary immune response in most subjects. It is the most efficient immunoglobulin in agglutination, complement fixation, and other antibody responses, and is important in defense against bacteria and viruses. IgD is
25 the immunoglobulin that has no known antibody function, but may serve as an antigen receptor. IgE is the immunoglobulin that mediates immediate hypersensitivity by causing release of mediators from mast cells and basophils upon exposure to allergen.

 An “inducible” promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene
30 product to be produced in a cell substantially only when an inducer which corresponds to the promoter is present in the cell.

As used herein, the term “instructional material” includes a publication, a recording, a diagram, or any other medium of expression which can be used to communicate the usefulness of a compound, composition or delivery system of the invention in the kit for

detecting or monitoring the conditions, diseases or disorders recited herein. Optionally, or alternately, the instructional material can describe one or more methods of detecting or monitoring the conditions, diseases or disorders in a cell or a tissue of a mammal. The instructional material of the kit of the invention can, for example, be affixed to a container that contains the identified compound, composition or delivery system of the invention or be shipped together with a container that contains the identified compound, composition or delivery system. Alternatively, the instructional material can be shipped separately from the container with the intention that the instructional material and the compound be used cooperatively by the recipient.

The term “isolated” means altered or removed from the natural state. For example, a nucleic acid or a peptide naturally present in a living animal is not “isolated,” but the same nucleic acid or peptide partially or completely separated from the coexisting materials of its natural state is “isolated.” An isolated nucleic acid or protein can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

The term “isolated nucleic acid” refers to a nucleic acid segment or fragment which has been separated from sequences which flank it in a naturally occurring state, *i.e.*, a DNA fragment that has been removed from the sequences which are normally adjacent to the fragment, *i.e.*, the sequences adjacent to the fragment in a genome in which it naturally occurs. The term also applies to nucleic acids that have been substantially purified from other components which naturally accompany the nucleic acid, *i.e.*, RNA or DNA or proteins, that naturally accompany it in the cell. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or that exists as a separate molecule (*i.e.*, as a cDNA or a genomic or cDNA fragment produced by PCR or restriction enzyme digestion) independent of other sequences. It also includes a recombinant DNA that is part of a hybrid gene encoding additional polypeptide sequence.

In the context of the present invention, the following abbreviations for the commonly occurring nucleic acid bases are used. “A” refers to adenosine, “C” refers to cytosine, “G” refers to guanosine, “T” refers to thymidine, and “U” refers to uridine.

As used herein, the term “monoclonal antibody” includes antibodies that display a single binding specificity and affinity for a particular epitope. These antibodies are mammalian-derived antibodies, including murine, human and humanized antibodies.

Unless otherwise specified, a “nucleotide sequence encoding an amino acid

sequence” includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. The phrase nucleotide sequence that encodes a protein or an RNA may also include introns to the extent that the nucleotide sequence encoding the protein may in some version contain an intron(s).

5 Unless otherwise specified, a “nucleotide sequence encoding an amino acid sequence” includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns.

10 The term “operably linked” refers to functional linkage between a regulatory sequence and a heterologous nucleic acid sequence resulting in expression of the latter. For example, a first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably
15 linked DNA sequences are contiguous and, where necessary to join two protein coding regions, in the same reading frame.

 As used herein, the terms “patient” and “subject” and “individual” refer interchangeably to a human or a non-human mammal. Non-human mammals include, for example, livestock and pets, such as ovine, bovine, porcine, canine, feline and murine
20 mammals. In certain embodiments, the patient or subject is human.

 As used herein, the terms “peptide,” “polypeptide,” and “protein” are used interchangeably, and refer to a compound comprised of amino acid residues covalently linked by peptide bonds. A protein or peptide must contain at least two amino acids, and no limitation is placed on the maximum number of amino acids that can comprise a protein’s or
25 peptide’s sequence. Polypeptides include any peptide or protein comprising two or more amino acids joined to each other by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types. “Polypeptides” include, for example, biologically
30 active fragments, substantially homologous polypeptides, oligopeptides, homodimers, heterodimers, variants of polypeptides, modified polypeptides, derivatives, analogs, fusion proteins, among others. The polypeptides include natural peptides, recombinant peptides, synthetic peptides, or a combination thereof.

The term “polynucleotide” as used herein is defined as a chain of nucleotides. Furthermore, nucleic acids are polymers of nucleotides. Thus, nucleic acids and polynucleotides as used herein are interchangeable. One skilled in the art has the general knowledge that nucleic acids are polynucleotides, which can be hydrolyzed into the monomeric “nucleotides.” The monomeric nucleotides can be hydrolyzed into nucleosides. As used herein polynucleotides include, but are not limited to, all nucleic acid sequences which are obtained by any means available in the art, including, without limitation, recombinant means, *i.e.*, the cloning of nucleic acid sequences from a recombinant library or a cell genome, using ordinary cloning technology and PCRTM, and the like, and by synthetic means.

The term “promoter” as used herein is defined as a DNA sequence recognized by the synthetic machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a polynucleotide sequence.

As used herein, the term “promoter/regulatory sequence” means a nucleic acid sequence which is required for expression of a gene product operably linked to the promoter/regulatory sequence. In some instances, this sequence may be the core promoter sequence and in other instances, this sequence may also include an enhancer sequence and other regulatory elements which are required for expression of the gene product. The promoter/regulatory sequence may, for example, be one which expresses the gene product in a tissue specific manner.

By the term “specifically binds,” as used herein with respect to an antibody, is meant an antibody that recognizes a specific antigen, but does not substantially recognize or bind other molecules in a sample. For example, an antibody that specifically binds to an antigen from one species may also bind to that antigen from one or more species. But, such cross-species reactivity does not itself alter the classification of an antibody as specific. In another example, an antibody that specifically binds to an antigen may also bind to different allelic forms of the antigen. However, such cross reactivity does not itself alter the classification of an antibody as specific. In some instances, the terms “specific binding” or “specifically binding,” can be used in reference to the interaction of an antibody, a protein, or a peptide with a second chemical species, to mean that the interaction is dependent upon the presence of a particular structure (*e.g.*, an antigenic determinant or epitope) on the chemical species; for example, an antibody recognizes and binds to a specific protein structure rather than to proteins generally. If an antibody is specific for epitope “A”, the presence of a

molecule containing epitope A (or free, unlabeled A), in a reaction containing labeled “A” and the antibody, will reduce the amount of labeled A bound to the antibody.

As used herein, the term “substantially the same” amino acid sequence is defined as a sequence with at least 70 %, preferably at least about 80 %, more preferably at least about 90 %, even more preferably at least about 95 %, and most preferably at least 99 %
5 homology to another amino acid sequence, as determined by the FASTA search method in accordance with Pearson & Lipman, Proc. Natl. Inst. Acad. Sci. USA 1988, 85:2444-2448.

By the term “synthetic antibody” as used herein is meant an antibody that is generated using recombinant DNA technology, such as, for example, an antibody expressed
10 by a bacteriophage as described herein. The term should also be construed to mean an antibody that has been generated by the synthesis of a DNA molecule encoding the antibody and which DNA molecule expresses an antibody protein, or an amino acid sequence specifying the antibody, wherein the DNA or amino acid sequence has been obtained using synthetic DNA or amino acid sequence technology which is available and well known in the
15 art.

A “tissue-specific” promoter is a nucleotide sequence that, when operably linked with a polynucleotide encodes or specified by a gene, causes the gene product to be produced in a cell substantially only if the cell is a cell of the tissue type corresponding to the promoter.

The term “transfected” or “transformed” or “transduced” as used herein refers
20 to a process by which exogenous nucleic acid is transferred or introduced into the host cell. A “transfected” or “transformed” or “transduced” cell is one that has been transfected, transformed or transduced with exogenous nucleic acid. The cell includes the primary subject cell and its progeny.

As used herein, the term “uncontrolled HIV infection” refers to an HIV-
25 infected human who is receiving HIV treatment and yet has high serum viral loads (such as, but not limited to, equal to or greater than 1,000 copies/mL).

The phrase “under transcriptional control” or “operatively linked” as used
30 herein means that the promoter is in the correct location and orientation in relation to a polynucleotide to control the initiation of transcription by RNA polymerase and expression of the polynucleotide.

A “vector” is a composition of matter comprising an isolated nucleic acid and used to deliver the isolated nucleic acid to the interior of a cell. Numerous vectors are known in the art including, but not limited to, linear polynucleotides, polynucleotides associated with

ionic or amphiphilic compounds, plasmids, and viruses. Thus, the term “vector” includes an autonomously replicating plasmid or a virus. The term should also be construed to include non-plasmid and non-viral compounds which facilitate transfer of nucleic acid into cells, such as, for example, polylysine compounds, liposomes, and the like. Examples of viral vectors include, but are not limited to, adenoviral vectors, adeno-associated virus vectors, retroviral vectors, and the like.

Ranges: throughout this disclosure, various aspects of the invention can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 2.7, 3, 4, 5, 5.3, and 6. This applies regardless of the breadth of the range.

Description

The present invention relates to the unexpected discovery of a novel, non-invasive method for monitoring and/or assessing HIV viral load in a human. The method comprises analyzing an urine sample from the human for the presence of one or more protein markers that are associated with active systemic HIV replication. The method allows for the monitoring of systemic HIV replication and/or infection in a human, and/or the identification of a human with uncontrolled HIV infection.

As disclosed herein, in one aspect, a survey of the urinary proteome in subjects with highly active HIV infection was performed, and the results were then compared with published studies of the HIV-uninfected human urinary proteome. A remarkable overlap of proteins identified in the present HIV urine as compared with HIV-uninfected urine was observed: 863 of the 885 proteins found in three or more of the 19 samples of HIV urine were proteins also identified in HIV-uninfected urine. This level of correspondence indicates that the methods used herein broadly surveyed HIV urine proteomes, and that comparison with reported HIV-uninfected human urine proteomes is a valid strategy to identify candidate novel HIV urine biomarkers. HIV-1-derived proteins were not observed in urine, but several host proteins in the urine of HIV-infected subjects were not observed in

multiple studies of the normal human urinary proteome. These proteins stem from a wide range of cellular processes.

In certain embodiments, the unique urine proteins found in the greatest number of samples (14 of 19) were docking protein 7 (DOK7) and dynein heavy-chain 3 (DNAH3). DOK7 is a key component for proper formation of neuromuscular synapses and has no known interaction with HIV-1. The dynein heavy-chain 2 (DNAH2) isoform was also identified as unique to HIV urine samples. The peptide identifications clearly distinguish between the two dynein heavy-chain isoforms. For example, the peptide SVLTAAGNLK identified in HIV urine samples is unique to DNAH3. Conversely, the DNAH2 peptide LLMRIGDKEVEYNTNFR, not found in isoform 3, was identified in the HIV urine samples. Thus, both of these proteins, with functionally related roles in force generation during microtubule-based movement, are independent HIV urine-specific candidate markers, despite having no known interaction with HIV-1.

This study is the first general survey of urinary proteomics in HIV-infected subjects with active systemic viral replication. While no HIV-1 specific proteins were observed, several host proteins were found exclusively in the urine of subjects infected with HIV as compared to published surveys of the non-HIV-infected human urinary proteome. These HIV specific proteomic signatures provide insights in to the human physiological response to HIV infection and potentially serve as novel HIV biomarkers in urine.

Methods

The invention includes a method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient. In certain embodiments, the patient has received or is receiving a first anti-HIV medication. In other embodiments, the patient is a new-born human. In yet other embodiments, the patient is an infant under about 18 months of age.

The method comprises obtaining a bodily sample from the human. In certain embodiments, the sample comprises urine. In other embodiments, the first anti-HIV medication comprises ART. In yet other embodiments, the patient has received or is receiving ART.

The method further comprises analyzing the test sample comprising urine from the patient for the presence and/or concentration of one or more proteins contemplated within the invention.

In certain embodiments, the test sample is processed, using methods such as but not limited to protein isolation and/or protein digestion. In other embodiments, the processed sample is analyzed by mass spectrometry, whereby the presence and/or concentration of specific peptides in the sample may be correlated with the presence and/or concentration of one or more proteins contemplated within the invention.

In certain embodiments, the sample is analyzed for the presence and/or concentration of a protein using a quantum dot assay and/or chromophore assay. Such analysis is known to those skilled in the art (Stepanenko, *et al.*, 2011, "Modern fluorescent proteins: from chromophore formation to novel intracellular applications," *Biotechniques* 51(5):313-8; Mehta, *et al.*, "Surface modified quantum dots as fluorescent probes for biomolecule recognition," 2014, *J. Nanosci. Nanotechnol.* 14(1):447-59; Geszke-Moritz & Moritz, 2013, "Quantum dots as versatile probes in medical sciences: synthesis, modification and properties," *Mater. Sci. Eng. C Mater. Biol. Appl.* 33(3):1008-21).

In certain embodiments, the sample is analyzed for the presence and/or concentration of a protein contemplated within the invention using an antibody or aptamer that binds to the protein. In other embodiments, the antibody is at least one selected from the group consisting of a polyclonal antibody, monoclonal antibody, Fv, Fab, F(ab)₂, single chain antibody, human antibody, humanized antibody, and fragments and derivatives thereof. In yet other embodiments, the analysis for the presence and/or concentration of the protein contemplated within the invention comprises an immunoassay. In yet other embodiments, the immunoassay comprises at least one selected from the group consisting of immunoturbidimetry, immunonephelometry, ELISA assay, radioimmunoassay, chemiluminescence immunoassay, immunofluorescence, immunoprecipitation, immunoelectrophoresis, and flow cytometry-based immunoassay.

The method further comprises comparing the presence and/or concentration of the protein in the test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample. In certain embodiments, the control sample comprises an urine sample from an untreated HIV-infected control human. In other embodiments, the untreated HIV-infected control human is the patient before receiving anti-HIV medication. In other embodiments, the control sample comprises an urine sample from an HIV-uninfected control human. In other embodiments, the control sample comprises an urine sample from an HIV-infected control human with controlled infection.

In certain embodiments, comparison of the results for the test data set and the control data set allows for the monitoring and/or assessment of the systemic HIV load in the patient.

5 In certain embodiments, the concentration of the protein in the patient's urine is higher by at least a multiplicity factor than the concentration of the protein in the urine sample from an HIV-uninfected control human or an HIV-infected control human with controlled infection, and the patient is identified as having uncontrolled HIV infection. In other embodiments, the concentration of the protein in the patient's urine is lower by at least a multiplicity factor than the concentration of the protein in the urine sample from an HIV-
10 uninfected control human or an HIV-infected control human with controlled infection, and the patient is identified as having a controlled HIV infection. In other embodiments, the multiplicity factor is selected from the group consisting of about 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 30, 40, 50, 75, 100, 125, 250, 500 and 1,000.

15 In certain embodiments, the concentration of the protein in the patient's urine is equal to or greater than a multiplicity factor of the concentration of the protein in the urine sample from an untreated HIV-positive control human, and the patient is identified as having uncontrolled HIV infection. In other embodiments, the concentration of the protein in the patient's urine is lower than the concentration of the protein in the urine sample from an
20 untreated HIV-positive control human, and the patient is identified as having a controlled HIV infection. In other embodiments, the multiplicity factor is selected from the group consisting of about 1, 0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.025, 0.01, 0.005, 0.0025, 0.001, 0.0005, 0.00025, 0.0001, 0.00005 and 0.00001.

25 In certain embodiments, the patient is identified as having controlled HIV infection, and the patient continues to be prescribed the first anti-HIV medication.

In certain embodiments, the patient is identified as having an uncontrolled HIV infection, and the patient is prescribed a second anti-HIV medication.

30 In certain embodiments, the patient is identified as having an uncontrolled HIV infection and has not received any anti-HIV medication (such as for example a newborn), and the patient is prescribed an anti-HIV medication.

In certain embodiments, the at least one protein has an accession number selected from the group consisting of Q8TD57 (SEQ ID NO:1), Q18PE1 (SEQ ID NO:2), Q8NFH5 (SEQ ID NO:3), Q8WYL5 (SEQ ID NO:4), Q8IYD8 (SEQ ID NO:5), O14654 (SEQ ID NO:6), Q96AP4 (SEQ ID NO:7), Q9UQ35 (SEQ ID NO:8), Q8N6W0 (SEQ ID

NO:9), Q9H792 (SEQ ID NO:10), Q9H497 (SEQ ID NO:11), Q9UE35 (SEQ ID NO:12), O00743 (SEQ ID NO:13), Q8WXF8 (SEQ ID NO:14), P81274 (SEQ ID NO:15), Q8NG08 (SEQ ID NO:16), Q96AE7 (SEQ ID NO:17), Q9BZM4 (SEQ ID NO:18), Q5T2D3 (SEQ ID NO:19), Q8IXT5 (SEQ ID NO:20), Q9P225 (SEQ ID NO:21), and Q9Y2I9 (SEQ ID NO:22).

In certain embodiments, the at least one protein has an accession number selected from the group consisting of P41222 (PTGDS) (SEQ ID NO:23), P14151 (SELL) (SEQ ID NO:24), Q06418 (TYRO3) (SEQ ID NO:25), P52306 (RAP1GDS1) (SEQ ID NO:26), and Q9Y5Y7 (LYVE1) (SEQ ID NO:27).

Antibodies

Using conventional techniques, the skilled artisan may use the nucleotide and amino acid sequences of the proteins contemplated within the invention to prepare an antigenic peptide for use in generating corresponding antibody. The sequence for the proteins contemplated within the invention are listed in Tables 1-2.

Alternatively, the skilled artisan may utilize a commercially available antibody against a protein contemplated within the invention. The skilled artisan may also obtain commercially available antibodies and modify them using conventional methods such as coupling to other antibodies, partial digestion, pegylation or covalent modification. Modified antibodies may then be used in the methods of the invention as described herein. Antibodies useful in the practice of the present invention may be polyclonal, monoclonal, synthetic or fragments of any of the above.

It will be appreciated that an antibody used in the invention may be monovalent, divalent or polyvalent in order to achieve antigen binding. Monovalent immunoglobulins are dimers (HL) formed of a hybrid heavy chain associated through disulfide bridges with a hybrid light chain. Divalent immunoglobulins are tetramers (H2L2) formed of two dimers associated through at least one disulfide bridge.

The invention also includes functional equivalents of the antibodies described herein. Functional equivalents have binding characteristics comparable to those of the antibodies, and include, for example, hybrid and single chain antibodies, as well as fragments thereof. Methods of producing such functional equivalents are disclosed for example in PCT Application Nos. WO 1993/21319 and WO 1989/09622. Functional equivalents include polypeptides with amino acid sequences substantially the same as the amino acid sequence of

the variable or hypervariable regions of the antibodies raised against proteins contemplated within the invention, according to the practice of the present invention.

Functional equivalents of the antibodies further include fragments of antibodies that have the same, or substantially the same, binding characteristics to those of the whole antibody. Such fragments may contain one or both Fab fragments or the F(ab')₂ fragment. Preferably the antibody fragments contain all six complement determining regions of the whole antibody, although fragments containing fewer than all of such regions, such as three, four or five complement determining regions, are also functional. The functional equivalents are members of the IgG immunoglobulin class and subclasses thereof, but may be or may combine any one of the following immunoglobulin classes: IgM, IgA, IgD, or IgE, and subclasses thereof. Heavy chains of various subclasses, such as the IgG subclasses, are responsible for different effector functions and thus, by choosing the desired heavy chain constant region, hybrid antibodies with desired effector function are produced. Preferred constant regions are gamma 1 (IgG1), gamma 2 (IgG2 and IgG), gamma 3 (IgG3) and gamma 4 (IgG4). The light chain constant region can be of the kappa or lambda type.

The monoclonal antibodies may be advantageously cleaved by proteolytic enzymes to generate fragments retaining the antigen binding site. For example, proteolytic treatment of IgG antibodies with papain at neutral pH generates two identical so-called "Fab" fragments, each containing one intact light chain disulfide-bonded to a fragment of the heavy chain (Fc). Each Fab fragment contains one antigen-combining site. The remaining portion of the IgG molecule is a dimer known as "Fc". Similarly, pepsin cleavage at pH 4 results in the so-called F(ab')₂ fragment.

Single chain antibodies or Fv fragments are polypeptides that consist of the variable region of the heavy chain of the antibody linked to the variable region of the light chain, with or without an interconnecting linker. Thus, the Fv comprises an antibody combining site.

Hybrid antibodies may be employed. Hybrid antibodies have constant regions derived substantially or exclusively from human antibody constant regions and variable regions derived substantially or exclusively from the sequence of the variable region of a monoclonal antibody from each stable hybridoma.

Methods for preparation of fragments of antibodies are known to those skilled in the art. See, Goding, "Monoclonal Antibodies Principles and Practice", Academic Press (1983), p. 119-123. Fragments of the monoclonal antibodies containing the antigen binding site, such as Fab and F(ab')₂ fragments, may be preferred in therapeutic applications, owing

to their reduced immunogenicity. Such fragments are less immunogenic than the intact antibody, which contains the immunogenic Fc portion. Hence, as used herein, the term “antibody” includes intact antibody molecules and fragments thereof that retain antigen binding ability.

5 When the antibody used in the practice of the invention is a polyclonal antibody (IgG), the antibody is generated by inoculating a suitable animal with a protein contemplated within the invention, or a fragment thereof. Antibodies produced in the inoculated animal that specifically bind to a protein contemplated within the invention are then isolated from fluid obtained from the animal. Antibodies may be generated in this
10 manner in several non-human mammals such as, but not limited to, goat, sheep, horse, rabbit, and donkey. Methods for generating polyclonal antibodies are well known in the art and are described, for example in Harlow et al. (In: Antibodies, A Laboratory Manual, 1988, Cold Spring Harbor, NY). These methods are not repeated herein as they are commonly used in the art of antibody technology.

15 When the antibody used in the methods used in the practice of the invention is a monoclonal antibody, the antibody is generated using any well-known monoclonal antibody preparation procedures such as those described, for example, in Harlow *et al.* (In: Antibodies, A Laboratory Manual, 1988, Cold Spring Harbor, NY) and Tuszynski *et al.* (Blood 1988, 72: 109-115). Given that these methods are well known in the art, they are not
20 replicated herein. Generally, monoclonal antibodies directed against a desired antigen are generated from mice immunized with the antigen using standard procedures as referenced herein. Monoclonal antibodies directed against full length or fragments of target structure may be prepared using the techniques described in Harlow et al. (In: Antibodies, A Laboratory Manual, 1988, Cold Spring Harbor, NY).

25 The skilled artisan would further appreciate, based upon the disclosure provided herein, that the invention is not limited to the use of an antibody as the binding element for a protein contemplated within the invention. The invention also allows for the use of an non-antibody molecule as the element that binds to one or more of the proteins that are contemplated in the invention. The non-antibody molecule may bind to the protein or a
30 fragment of the protein. Preferred non-antibody molecules within the invention are aptamers. Aptamers are oligonucleic acid (also referred to as nucleic acid) molecules or peptide molecules that bind a specific target molecule. Nucleic acid aptamers are nucleic acid species that have been engineered through repeated rounds of in vitro selection or equivalently, SELEX (systematic evolution of ligands by exponential enrichment), to bind to

various molecular targets such as small molecules, proteins, nucleic acids, and even cells, tissues and organisms. Aptamers are useful in biotechnological and therapeutic applications as they offer molecular recognition properties that rival that of the commonly used antibodies. In addition to their discriminate recognition, aptamers offer advantages over
5 antibodies as they can be engineered completely in a test tube, are readily produced by chemical synthesis, possess desirable storage properties, and elicit little or no immunogenicity in therapeutic applications. See Ellington & Szostak, 1990, *Nature* 346(6287):818-22; Bock, *et al.*, 1992, *Nature* 355(6360):564-6; Drabovich, *et al.*, 2006, *Anal. Chem.* 78(9):3171-8, all of which are incorporated herein by reference in their entireties.
10 Aptamers useful within the invention may be selected and/or prepared according to the teachings of the art.

The binding of the antibody to the protein contemplated within the invention may be analyzed using any appropriate immunoassay available and/or known to those skilled in the art. Immunoassays are based on specific binding of an antibody to its antigen (in this
15 particular case, the protein contemplated within the invention). Detecting the interaction of the antibody with the antigen may be achieved using a variety of methods, of which one of the most common is to label either the antigen or antibody, and monitor the change in environment of the label upon binding. The label may comprise an enzyme (wherein binding is monitored by enzyme immunoassay or EIA), colloidal gold (wherein binding is monitored
20 by lateral flow assays), radioisotopes such as ¹²⁵I radioimmunoassay (wherein binding is monitored by radiometric methods), magnetic labels (wherein binding is monitored by magnetic immunoassay or MIA) or fluorescence. Other techniques include, but are not limited to, agglutination, nephelometry, turbidimetry and Western Blot. All of these methods are known to those of skill in the art. See *e.g.* Harlow, *et al.*, 1988, "Antibodies: A
25 Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; Harlow, *et al.*, 1999, "Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press", Cold Spring Harbor, NY.

Immunoassays may be divided into those that involve non-labelled reagents and those that involve labelled reagents. Immunoassays that involve labelled reagents are
30 divided into homogenous immunoassays and heterogeneous immunoassays (the latter require an extra step to remove unbound antibody or antigen from the site, usually using a solid phase reagent).

Heterogeneous immunoassays may be competitive or non-competitive. In a competitive immunoassay, the antigen in the unknown sample competes with labeled antigen

to bind with antibodies. The amount of labeled antigen bound to the antibody site is then measured. In this method, the response will be inversely proportional to the concentration of antigen in the unknown, since a large response indicates that there is little antigen in the unknown to compete with the labeled antigen. In noncompetitive immunoassays, also referred to as the “sandwich assay,” antigen in the unknown is bound to the antibody site, then labeled antibody is bound to the antigen. The amount of labeled antibody on the site is then measured. Unlike the competitive method, the results of the noncompetitive method are directly proportional to the concentration of the antigen, since the labeled antibody will not bind if the antigen is not present in the unknown sample.

In certain embodiments, the immunoassay is selected from the group consisting of immunoturbidimetry, immunonephelometry, an ELISA assay, radioimmunoassay, chemiluminescence immunoassay, immunofluorescence, immunoprecipitation, immunoelectrophoresis, and flow cytometry-based immunoassay.

One skilled in the art will recognize that optimization studies may be easily performed to determine which chemical reagent(s) present in solution do, or do not, significantly interfere with the selective binding of the antibody to the antigen. The optimization studies may involve the use of two samples, one comprising the protein of interest and the chemical reagent, and the second comprising the protein of interest but devoid of the chemical reagent. The two samples are separately incubated with the antibody.

Non-limiting examples of such chemical reagents are surfactants, non-ionic surfactants, divalent cation salts, dextran salts, PEG, α -cyclodextrin salts, EDTA, and azide salts. Following incubations, an immunoassay is used to determine the degree of antibody binding for each sample, and this information is used to determine the effect of the chemical reagent on the antibody-antigen binding. This evaluation follows standard methodologies used in analytical sciences and should not require unwarranted experimentation from those skilled in the art.

The immunoassay used to detect the interaction of the antibody with the protein of interest may also be used to quantitate the concentration of the protein in the sample. In a typical procedure included in the invention, a series of standard solutions containing known concentrations of the protein of interest are prepared and analyzed by an immunoassay. The readings obtained for each standard solution are used to create a calibration curve. The unknown sample is then analyzed by the same immunoassay and its reading is compared to the standard curve in order to obtain a corresponding concentration of the protein of interest in the sample. This concentration may be used to calculate the actual

concentration of the protein of interest in the biological fluid, taking into account the dilutions that the biological sample was subjected to for the preparation of the test sample.

Use of the calibration curve, as described above, allows the concentration of the protein to be determined in the same units used to express the concentration of the standard solutions. In some instances, the standard solutions have their component concentrations identified in mass/volume units (such as mg/dL units, for example). The concentration of the protein of interest in the biological sample, determined as mg/dL from the calibration curve, may be converted to a concentration of moles/volume (such as nmol/L) based on the molecular weight of the protein of interest.

As will be understood by one of skill in the art, when armed with the disclosure set forth herein, a set of reference proteins or equivalents (also referred to as "calibration samples") may be used to create a calibration curve for a certain method and/or instrument. By way of a non-limiting example, the set of reference proteins or equivalents may be used in a one- or two-point calibration assay. In another embodiment of the invention, the set of reference proteins or equivalents may be used in a three-, four-, five- or six point calibration assay. In one aspect, the set of reference proteins or equivalents may include as many or as few reference points as determined to be necessary to establish a valid and accurate reference curve.

Numerous calibration schemes may be used in the clinical laboratory. Some methods, often manually performed, employ several concentration levels throughout the assay range and typically plot the instrumental response versus concentration or use linear regression to calculate patient analyte values. With the increasing use and availability of computer technology, methods often use one or two calibrator points to achieve the same results. Quite often, the one or two set point method incorporates a saline or distilled water blank as an additional set point, this latter function being dictated by the instrument or reagent manufacturer. For non-linear chemistries, the traditional approach provides five or six levels of calibrator, usually set in a non-linear fashion dictated by the mathematical model used in the final calculation of patient result. A more recent trend for non-linear chemistries is to use one calibrator containing the highest concentration of analyte measured in the assay. Using this method, the analytical system is then directed to perform the necessary dilutions of this high concentration value to generate the predetermined calibration set points on the fly when the system calibrates the analyte. A four- or five-parameter logit/log calibration curve is typically used for automated immunoassays.

Therefore, in an aspect of the present invention, there is provided a method that features the use of multiple calibrator points in order to generate a reference curve. In one embodiment, the method features the use of more than one point. In another embodiment, one of the multiple points is a zero point. In yet another embodiment, the zero point is not included as one of the multiple points, but may be included separately in a reference curve. In another embodiment, the method features the use of a single calibration point, as described in detail elsewhere herein. In yet another embodiment, the method features the use of a zero point in addition to a single calibration point.

By way of a series of non-limiting examples, the method of the invention may use a reference curve based on a single concentration for calibration, a reference curve based on a single concentration plus a zero concentration point for calibration, a reference curve based on at least two concentrations for calibration, or a reference curve based on at least two concentrations plus a zero concentration point for calibration. In one embodiment of the invention, the concentration of a calibration sample is known. In yet another embodiment of the invention, the concentration of at least one calibration sample in a mixture containing at least two calibration samples is known.

Kits

The invention includes various kits that comprise a set of protein antibodies, or equivalents thereof, an applicator, and instructional materials that describe the use of the kit to perform the methods of the invention. Although exemplary kits are described below, the contents of other useful kits will be apparent to the skilled artisan in light of the present disclosure. Each of these kits is included within the invention. The kit is used pursuant to the methods disclosed in the invention.

In certain embodiments, the invention includes a kit for measuring the concentration of at least one protein contemplated in the invention in a biological sample of a patient. In other embodiments, the biological sample comprises urine. The kit may comprise reagents, such as antibodies or equivalents thereof, that allow for the determination of the at least one protein contemplated in the invention. The kit further comprises an applicator and instructional material for the use of the kit.

The kit may further comprise an applicator useful for administering the reagents for use in the relevant assay. The particular applicator included in the kit will depend on, *e.g.*, the method used to assay the protein, as well as the particular analyzer equipment used, and such applicators are well-known in the art and may include, among

other things, a pipette, a syringe, a dropper bottle, and the like. Moreover, the kit may comprise an instructional material for the use of the kit.

Further, the invention includes a kit comprising at least one reference composition comprising a known value of a known constituent, which may be a protein, a derivative thereof or a fragment thereof. Such kits may be used to create a calibration curve for quantitation of the protein. Thus, the invention encompasses a kit comprising at least one reference composition. While the invention is not limited to any particular set, certain combinations of reference compositions are exemplified elsewhere herein.

In certain embodiments, the invention includes a kit for assessing or monitoring systemic HIV viral load in an HIV-infected human patient, the kit comprising an antibody or aptamer that binds to at least one protein contemplated within the invention; an applicator; and, an instructional material for the use of the kit, wherein the instructional material comprises instructions for analyzing a test sample comprising urine from the patient for the presence or concentration of the at least one protein.

In certain embodiments, the kit further comprises a test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample. In other embodiments, the control sample comprises an urine sample from an untreated HIV-infected control human and/or an HIV-negative control human and/or an HIV-infected control human with controlled infection.

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, numerous equivalents to the specific procedures, embodiments, claims, and examples described herein. Such equivalents were considered to be within the scope of this invention and covered by the claims appended hereto. For example, it should be understood, that modifications in reaction conditions, including but not limited to reaction times, reaction size/volume, and experimental reagents, such as solvents, catalysts, pressures, atmospheric conditions, *e.g.*, nitrogen atmosphere, and reducing/oxidizing agents, with art-recognized alternatives and using no more than routine experimentation, are within the scope of the present application.

It is to be understood that wherever values and ranges are provided herein, all values and ranges encompassed by these values and ranges, are meant to be encompassed within the scope of the present invention. Moreover, all values that fall within these ranges, as well as the upper or lower limits of a range of values, are also contemplated by the present application.

The following examples further illustrate aspects of the present invention. However, they are in no way a limitation of the teachings or disclosure of the present invention as set forth herein.

5

EXAMPLES

The invention is now described with reference to the following Examples. These Examples are provided for the purpose of illustration only, and the invention is not limited to these Examples, but rather encompasses all variations that are evident as a result of the teachings provided herein.

10

Methods and Materials

Sample Collection and Processing

Subjects were asked to refrain from consuming alcohol and nonprescription drugs for 24 hours prior to sample collection but were allowed to maintain a normal diet otherwise. Subjects provided their second void of the day after approximately 5 mL of urine had been passed. Samples were promptly placed on ice, centrifuged at 2000 x g for 20 minutes at 4 °C to remove any cells that may have been extraneously passed, and stored at -70 °C.

Protein Isolation and Digestion

Urine solutions were brought to 8 M urea, 10 mM dithiothreitol, 100 mM Tris HCl, pH 7.6, and concentrated using a 30-kD Amicon molecular-weight cutoff (MWCO) device (Millipore, Billerica, MA). Concentrated proteins were depleted of albumin using a Cibacron blue–based method (Pierce, Rockford, IL). Immunoglobulins were depleted using the “top 2” abundant-protein depletion column from Thermo Pierce (<http://www.piercenet.com/product/abundant-protein-depletion-spin-columns>).

A volume of urine containing 500 µg of total protein was buffer exchanged to 10 mM PBS and 0.15 M NaCl using a 3-kD MWCO spin filter (Millipore) and loaded to the depletion column. The sample was incubated in the column for 30 minutes, reverse transcribed, and mixed at 500 rpm (MixMate, Eppendorf, Hamburg, Germany). Following incubation the column was spun and the depleted sample collected for further processing. Depleted protein samples were transferred to a 30-kD Amicon MWCO device (Millipore) and centrifuged at 3,000 x g for 30 minutes. The remaining sample was buffer exchanged with 6 M urea, 100 mM Tris HCl, pH 7.6, then alkylated with 55 mM iodoacetamide. Concentrations were measured using a Qubit fluorometer (Invitrogen, Carlsbad, CA).

Trypsin was added at a ratio of 1:40 enzyme to substrate and the sample incubated overnight on a heat block at 37 °C. The device was centrifuged at 3,000 x g for 30 minutes and the filtrate collected.

Peptide Desalting

5 Digested peptides were desalted using C18 stop-and-go extraction (STAGE) tips. For each sample, a C18 STAGE tip was activated with methanol, then conditioned with 60% acetonitrile/0.5% acetic acid, followed by 5% acetonitrile/0.5% acetic acid. Samples were loaded onto the tips and desalted with 0.5% acetic acid. Peptides were eluted with 60% acetonitrile/0.5% acetic acid and lyophilized in a SpeedVac (Thermo Savant) to dryness, for
10 approximately 2 h.

Liquid Chromatography–Tandem Mass Spectrometry

Each fraction was analyzed by reverse-phase liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). LC was performed on a Thermo Easy NanoLC II system. Mobile phase A included 94.5% Milli-Q water (Millipore) and 5%
15 acetonitrile/0.5% acetic acid. Mobile phase B included 80% acetonitrile, 19.5% Milli-Q water, and 0.5% acetic acid. The 120-minute LC gradient ran from 0% B to 35% B over 90 minutes, with the remaining time used for sample loading and column regeneration. Samples were loaded to a 2 cm × 100-μm inside-diameter trap column. The analytical column was 13 cm × 75 μm inside-diameter fused silica with a pulled tip emitter. Both trap and analytical
20 columns were packed with 3.5-μm C18 resin (Magic C18AQ, Michrom, Fremont, CA). The LC was interfaced to a dual-pressure linear ion trap mass spectrometer (LTQ Velos, Thermo Fisher) via nanoelectrospray ionization. An electrospray voltage of 1.8 kV was applied to a precolumn tee. The mass spectrometer was programmed to acquire, by data-dependent acquisition, tandem mass spectra from the top 15 ions in the full scan from 400 to 1400 m/z.

Data Processing and Library Searching

Mass spectrometer RAW data files were converted to Mascot generic format (MGF) using msconvert. All searches required strict tryptic cleavage, 0 or 1 missed
cleavages, fixed modification of cysteine alkylation, variable modification of methionine oxidation, and expectation value scores of 0.01 or lower. MGF files were searched using
30 X!Hunter against the latest spectral library available in the Global Proteome Machine database at the time. X!!Tandem and OMSSA (Open Mass Spectrometry Search Algorithm) searches used Ensembl protein sequence libraries. The human sequence library used in this analysis was the Ensembl Genome Browser (“Human”) (http://useast dot ensembl dot org/Homo_sapiens/Info/Index). MGF files were searched using X!!Tandem using both the

native and k-score8 scoring algorithms and OMSSA. All searches were performed on Amazon (Seattle, WA) Web Services-based Cluster Compute instances using the Proteome Cluster interface. XML output files were parsed and nonredundant protein sets were determined using in-house scripts. Proteins were required to have 1 or more unique peptides with peptide E-value scores of ≤ 0.01 from X!Tandem, ≤ 0.01 from OMSSA, ≤ 0.001 and theta values of ≥ 0.5 from X!Hunter searches, and protein E-value scores of ≤ 0.0001 from X!Tandem and X!Hunter.

Proteins identified in ≥ 3 HIV-infected urine samples were then compared with published studies of the human urinary proteome to assess potential uniqueness to the urinary proteome of the HIV-infected. Unique urine proteins in the HIV-infected were searched for in the HIV-1, Human Protein Interaction Database and Host Proteins in HIV-1 database in order to report known relevance in HIV biology. Gene ontology information was derived from www.ebi.ac.uk/UniProt.

Example 1:

Study Population

Subjects from the Drexel University College of Medicine HIV clinic were enrolled in this single-center study. Eligible patients included those aged ≥ 18 years with clade B chronic HIV-1 infection free of baseline resistance based on genotype or phenotype testing, with fewer than 2 weeks of intervening antiretroviral therapy, and an HIV-1 serum viral load $\geq 50,000$ copies/mL in the prior 30 days.

Exclusion criteria were:

chronic hepatitis B virus (HBV) or hepatitis C virus (HCV) infection as defined by positive results from serology for HBV surface antigen or detectable HCV viral load by polymerase chain reaction, respectively;

evidence of active infection in the prior 2 weeks;

treatment for acute opportunistic infection, including *Pneumocystis jirovecii* pneumonia, *Toxoplasma gondii* encephalitis, cryptosporidiosis, microsporidiosis, *Mycobacterium tuberculosis* disease, disseminated *Mycobacterium avium* complex disease, bacterial pneumonia, bacterial enteric disease, bartonellosis, syphilis, mucocutaneous candidiasis, cryptococcosis, histoplasmosis, coccidioidomycosis, aspergillosis, cytomegalovirus disease, herpes simplex virus disease, varicella zoster virus disease, human herpesvirus-8 disease, or progressive multifocal leukoencephalopathy caused by JC virus; hematuria on screening urinalysis in the past 30 days;

chemotherapy, radiotherapy, or immunotherapy in the past 30 days except for topical or inhaled steroids;

positive nucleic acid amplification testing of genitourinary tract for *Neisseria gonorrhoeae* or *Chlamydia trachomatis* in the prior 2 weeks; or

5 any other medical condition that rendered the subject unable to complete the study, interfered with participation, or produced significant risk to the subject.

Example

Urine samples from 19 subjects with clade B chronic HIV-1 infection having serum viral loads $\geq 50,000$ copies/mL in the prior 30 days were collected and frozen for
10 subsequent analysis (characteristics of study population are illustrated in Fig. 1). Albumin is generally the major protein constituent of urine and thus may prevent proteomic identification of lower-abundance HIV proteins or unique host biomarkers of HIV infection. Thus, urine samples were depleted of albumin.

HIV infection is associated with a chronic inflammatory state, and thus
15 anticipating high levels of immunoglobulin in the urine (which might also hinder identification of potential lower-abundance HIV peptides or host biomarkers), IgG was depleted from the urine samples. Raw data queried against HIV sequence databases did not identify any HIV-specific peptides. In searches against the human Fasta sequence database, combined analysis of all 19 samples (two of which were analyzed twice using the same LC-
20 MS/MS method) identified a total of 37,886 peptides corresponding to 1794 human-unique proteins. Compared to studies that have sought to comprehensively characterize the human urinary proteome, 22 proteins unique to HIV-infected urine were identified (Fig. 2).

Example 2:

25 The subjects had a mean age of 41 years. The subjects were 60% male, 32% female, and 8% transgender; were 88% Black, 8% Hispanic, and 4% White; had a median serum HIV viral load of 108,960 copies/mL; and a median CD4 count of 340 cells/ μ L.

Urine samples were collected from 20 adults with wild type clade B HIV-1 infection and an HIV-1 serum viral load $\geq 50,000$ copies/mL within 30 days.

30 Subjects were free of *Neisseria gonorrhoeae* or *Chlamydia trachomatis* urethritis, active or opportunistic infection, and hematuria. Samples were centrifuged to remove cellular debris and then frozen to -70°C . Thawed samples were concentrated then depleted of albumin $\pm$ immunoglobulins.

100µg of each sample were lyophilized and suspended in denaturing buffer before reduction, alkylation, and enzymatic digestion with sequencing grade trypsin. Samples underwent strong cation exchange before liquid chromatography coupled to tandem mass spectrometry (MS) with CID fragmentation. Datasets were searched against HIV and
5 fasta human protein databases with Bioworks Sequest algorithm and Protein Prospector. Sequest X-correct scores of 2.5 for doubly charged and 3 for triply charged, and Protein Prospector scores of 20 were used as initial thresholds for peptide identification. Spectral counts corresponding to peptide identifications were used to reflect relative abundance. Unique HIV urine peptide and protein signatures were identified through comparison with
10 reported urine proteomes from non-HIV infected persons.

About 1,500 peptides of about 400 unique proteins were identified in the urine samples (Fig. 3). HIV-derived peptides were not observed. In all cases, a non-immunoglobulin specific protein identified in more than two of the HIV urine samples was also found in reported non-HIV urine proteomes. Several urine markers appeared to be
15 significantly more abundant in HIV urine, including prostaglandin D2, which was found in every HIV urine sample and represented about the 6th most abundant protein (as compared to about the 100th most abundant protein in non-HIV urine samples). Other markers were unique to only the HIV-urine proteomes, such as L-selectin (10 of 20 samples) and lymphatic vessel endothelial hyaluronan receptor 1 (20 of 20 samples).

20 HIV-derived peptides were not identified by MS in the urine of subjects with uncontrolled HIV replication, but a clear increase in inflammatory markers and markers unique to HIV-urine were present, potentially offering insight into the pathogenesis and/or monitoring of HIV infection.

The disclosures of each and every patent, patent application, and publication
25 cited herein are hereby incorporated herein by reference in their entirety. While this invention has been disclosed with reference to specific embodiments, it is apparent that other embodiments and variations of this invention may be devised by others skilled in the art without departing from the true spirit and scope of the invention. The appended claims are intended to be construed to include all such embodiments and equivalent variations.

Table 1.

Q8TD57	10	20	30	40	50	60
SEQ ID NO:1	MGATGRLELT	LAAPPHGPA	FQRSKARETQ	GEEEGSEMQUI	AKSDSIHHMS	HSQGQPELPP
	70	80	90	100	110	120
	LPASANEEPS	GLYQTVMSHS	FYPPLMQRTS	WTLAAPFKEQ	HHHRGPSDSI	ANNYSLMAQD
	130	140	150	160	170	180
	LKLKDLLKVY	QPATISVPRD	RTGQGLPSSG	NRSSSEPMRK	KTKFSSRNKE	DSTRIKLAFK
	190	200	210	220	230	240
	TSIFSPMKKE	VKTSLTFFGS	RPMSPEQQLD	VMLQQEMEME	SKEKKPSESD	LERYYYYLTN
	250	260	270	280	290	300
	GIRKDMIAPE	EGEVMVRISK	LISNTLLTSP	FLEPLMVVLV	QEKENDYYCK	LMKSIVDYIL
	310	320	330	340	350	360
	MDPMERKRLE	IESIPRLFPQ	RVIRAPVFWH	SVYRSACKWN	EEHLHTVNP	MLRLKELWFA
	370	380	390	400	410	420
	EFRDLRFVRT	AEILAGKLPL	QPQEFWDVIQ	KHCLEAHQTL	LNKWIPTCAQ	LFTSRKEHWI
	430	440	450	460	470	480
	HFAPKSNYDS	SRNIEEYFAS	VASFMSLQLR	ELVIKSLEDL	VSLFMIHKDG	NDFKEPYQEM
	490	500	510	520	530	540
	KFFIPQLIMI	KLEVSEPIIV	FNPSFDGCWE	LIRDSFLEII	KNSNGIPKLE	YIPLKFSFTA
	550	560	570	580	590	600
	AAADRQCVKA	AEPGEPSMHA	AATAMAELKG	YNLLLGTVNA	EEKLVSDFLI	QTFKVFQKNQ
	610	620	630	640	650	660
	VGPCCKYLVY	KKYVDLLDNT	AEQNIAAFLK	ENHDIDDFVT	KINAIKKRRN	EIASMNITVP
	670	680	690	700	710	720
	LAMFCLDATA	LNHDLCEAQA	NLKDHLIQFQ	VDVNRDTNTS	ICNQYSHIAD	KVSEVPANTK
	730	740	750	760	770	780
	ELVSLIEFLK	KSSAVTVFKL	RRQLRDASER	LEFLMDYADL	PYQIEDIFDN	SRNLLLHKRD
	790	800	810	820	830	840
	QAEMDLIKRC	SEFELRLEGY	HRELESFRKR	EVMTTEEMKH	NVEKLNELSK	NLNRAFAEFE
	850	860	870	880	890	900
	LINKEEELLE	KEKSTYPLLQ	AMLKNKVPYE	QLWSTAYEFS	IKSEEWNGP	LFLLNAEQIA
	910	920	930	940	950	960
	EEIGNMWRTT	YKLIKTLSDV	PAPRRLAENV	KIKIDKFKQY	IPILSISCNP	GMKDRHWQQI
	970	980	990	1000	1010	1020
	SEIVGYEIKP	TETTCLSNML	EFGFGKFVEK	LEPIGAAASK	EYSLEKNLDR	MKLDWVNVTF
	1030	1040	1050	1060	1070	1080
	SFVKYRDTDT	NILCAIDDIQ	MLLDDHVIKT	QTMCGSPFIK	PIEAECRKWE	EKLIRIQDNL
	1090	1100	1110	1120	1130	1140
	DAWLKCQATW	LYLEPIFSSE	DIIAQMPEEG	RKFGIVDSYW	KSLMSQAVKD	NRILVAADQP
	1150	1160	1170	1180	1190	1200
	RMAEKLQEAN	FLLEDIQKGL	NDYLEKKRLE	FPRFFFLSND	ELLEILSETK	DPLRVQPHLK
	1210	1220	1230	1240	1250	1260
	KCFEGIAKLE	FTDNLEIVGM	ISSEKETVPF	IQKIYPANAK	GMVEKWLQQV	EQMMLASMRE
	1270	1280	1290	1300	1310	1320
	VIGLGIEAYV	KVPRNHWVLQ	WPGQVVICVS	SIFWTQEVSQ	ALAENTLLDF	LKKSNDQIAQ
	1330	1340	1350	1360	1370	1380
	IVQLVRGKLS	SGARLTGAL	TVIDVHARDV	VAKLSEDRVS	DLNDFQWISQ	LRYVWAKDV
	1390	1400	1410	1420	1430	1440
	QVQIITTEAL	YGYEYLGNSP	RLVITPLTDR	CYRTLGMALK	LNLGGAPEGP	AGTGKTETTK
	1450	1460	1470	1480	1490	1500
	DLAKALAKQC	VVFNCSDGLD	YKAMGKFFKG	LAQAGAWACF	DEFNRIEVEV	LSVVAQQILS
	1510	1520	1530	1540	1550	1560
	IQQAIIRKLE	TFIFEGTELS	LNPTCAVFIT	MNPGYAGRAE	LPDNLKALFR	TVAMMVPDYA
	1570	1580	1590	1600	1610	1620
	LIGEISLYSM	GFLDSRSLAQ	KIVATYRLCS	EQLSSQHMYD	YGMRAVKSVL	TAAGNLKLEK
	1630	1640	1650	1660	1670	1680
	PEENESVLLL	RALLDVNLAK	FLAQDVPLFQ	GIISDLFPGV	VLPKPDYEVF	LKVLNDNIKK
	1690	1700	1710	1720	1730	1740
	MKLQPVWFIF	GKIIQIYEMM	LVRHGYMIVG	DPMGGKTSAY	KVLAAALGDL	HAANQMEEFA
	1750	1760	1770	1780	1790	1800
	VEYKIINPKA	ITMGQLYGCF	DQVSHWMDG	VLANAFREQA	SSLSDDRKWI	IFDGPVDAIW
	1810	1820	1830	1840	1850	1860
	IENMNTVLDD	NKKLCLMSGE	IIQMNSKMSL	IFEPADLEQA	SPATVSRGCM	IYMEPHQLGW
	1870	1880	1890	1900	1910	1920

KPLKDSYMDT	LPSSLTKEHK	ELVNDMFMWL	VQPCLEFGRL	HCKFVVQTSP	IHLAFSMMRL
1930	1940	1950	1960	1970	1980
YSSLLDEIRA	VEEEEEMELGE	GLSSQQIFLW	LQGLFLFSLV	WTVAGTINAD	SRKKFDVFFR
1990	2000	2010	2020	2030	2040
NLIMGMDDNH	PRPKSVKLT	NNIFPERGSI	YDFYFIKQAS	GHWETWTQYI	TKEEEKVPAG
2050	2060	2070	2080	2090	2100
AKVSELIIPT	METARQSFFL	KTYLDHEIPM	LFVGPTGTGK	SAITNNFLLH	LPKNTYLPNC
2110	2120	2130	2140	2150	2160
INFSAARTSAN	QTQDIIMSKL	DRRRKGLFGP	PIGKKAVVFV	DDLNMPAKEV	YGAQPPIELL
2170	2180	2190	2200	2210	2220
RQWIDHGYWF	DKKDTTRLDI	VDMLLVMTAMG	PPGGGRNDIT	GRFTRHLNII	SINAFEDDIL
2230	2240	2250	2260	2270	2280
TKIFSSIVDW	HFGKGFDMVF	LRYGKMLVQA	TKTIYRDAVE	NFLPTPSKSH	YVFNLRDFSR
2290	2300	2310	2320	2330	2340
VIQGVLLCPH	THLQDVEKCI	RLWIHEVYRV	FYDRLIDKED	RQVFFNMVKE	TTSNCFKQTI
2350	2360	2370	2380	2390	2400
EKVLIIHLSPT	GKIVDDNIRS	LFFGDYFKPE	SDQKIYDEIT	DLKQLTVVME	HYLEEFNNIS
2410	2420	2430	2440	2450	2460
KAPMSLVMFR	FAIEHISRIC	RVLKQDKGHL	LLVGIGGSGR	QSAAKLSTFM	NAYELYQIEI
2470	2480	2490	2500	2510	2520
TKNYAGNDWR	EDLKKIILQV	GVATKSTVFL	FADNQIKDES	FVEDINMLLN	TGDVPNIFPA
2530	2540	2550	2560	2570	2580
DEKADIVEKM	QTAARTQGEK	VEVTPLSMYN	FFIERVINKI	SFSLAMSPIG	DAFRNRLRMF
2590	2600	2610	2620	2630	2640
PSLINCCTID	WFQSWPTDAL	ELVANKFLED	VELDDNIRVE	VVSMCKYFQE	SVKKLSLDYY
2650	2660	2670	2680	2690	2700
NKLRRHNYVT	PTSYLELILT	FKTLLNSKRQ	EVAMMRNRYL	TGLQKLDFAA	SQVAVMQREL
2710	2720	2730	2740	2750	2760
TALQPQLILT	SEETAKMMVK	IEAETREADG	KKLLVQADEK	EANVAAIAQ	GIKNECEGDL
2770	2780	2790	2800	2810	2820
AEAMPALEAA	LAALDTLNPA	DISLVKSMQN	PPGPVKLVME	SICIMKGMKP	ERKPDPSGSG
2830	2840	2850	2860	2870	2880
KMIEDYWGVS	KKILGDLKFL	ESLKYTDKDN	IPPLTMKRIR	ERFINHPEFQ	PAVIKNVSSA
2890	2900	2910	2920	2930	2940
CEGLCKWVRA	MEVYDRVAKV	VAPKRERLRE	AEGKLAAQMQ	KLNQKRAELK	LVVDRQLQALN
2950	2960	2970	2980	2990	3000
DDFEEMNTTK	KDLEENIEIC	SQKLVRAEKL	ISGLGGEKDR	WTEAARQLGI	RYTNLTGDVL
3010	3020	3030	3040	3050	3060
LSSGTVAYLG	AFTVDYRVQC	QNQWLAECKD	KVIPGFSDFS	LSHTLGDPK	IRAWQIAGLP
3070	3080	3090	3100	3110	3120
VDSFSIDNGI	IVSNSRRWAL	MIDPHGQANK	WIKNMEKANK	LAVIKFSDSN	YMRMLLENALQ
3130	3140	3150	3160	3170	3180
LGTPVLIENI	GEELDASIEP	ILLKATFKQQ	GVEYMRLGEN	IIEYSRDFKL	YITTRLRNPH
3190	3200	3210	3220	3230	3240
YLPEVAVKVC	LLNFMITPLG	LQDQLLGIVA	AKEKPELEEK	KNQLLIVESAK	NKKHLKEIED
3250	3260	3270	3280	3290	3300
KILEVLMSK	GNILEDETAI	KVLSSSKVLS	EEISEKQKVA	SMTETQIDET	RMGYKPVAVH
3310	3320	3330	3340	3350	3360
SATIFFCISD	LANIEPMYQY	SLTWFINLYM	HSLTHSTKSE	ELNLRIKYII	DHFTLSIYNN
3370	3380	3390	3400	3410	3420
VCRSLFEKDK	LLFSLLLTIG	IMKQKKEITE	EVWYFLLTGG	IALDNYPYNP	APQWLSEKAW
3430	3440	3450	3460	3470	3480
AEIVRASALP	KLHGLMEHLE	QNLGEWKLIY	DSAWPHEEQL	PGSWKFSQGL	EKMVILRCLR
3490	3500	3510	3520	3530	3540
PDKMVPVAVRE	FIAEHMGKLY	IEAPTDFDLQ	SYNDSSCCAP	LIFVLSPSAD	PMAGLLKFAD
3550	3560	3570	3580	3590	3600
DLGMMGGTRTQ	TISLGQGQGP	IAAKMINNAI	KDGTWVVLQN	CHLAASWMPT	LEKICEEVIV
3610	3620	3630	3640	3650	3660
PESTNARFRL	WLTSYPSEKF	PVSILQNGIK	MTNEPPKGLR	ANLLRSYLN	PISDPVFFQS
3670	3680	3690	3700	3710	3720
CAKAVMWQKM	LFGLCFFHAV	VQERRNFGPL	GWNIPYEFNE	SDLRISMWQI	QMFLNDYKEV
3730	3740	3750	3760	3770	3780
PFDALTYLTG	ECNYGGRVTD	DKDRRLLSL	LSMFYCKEIE	EDYYSAPAGD	TYIIPPHGSY
3790	3800	3810	3820	3830	3840
QSYIDYLRNL	PITAHPEVFG	LHENADITKD	NQETNQLFEG	VLLTLPRQSG	GSGKSPQEVV
3850	3860	3870	3880	3890	3900
EELAQDILSK	LPRDFDLEEV	MKLYPVVYEE	SMNTVLRQEL	IRFNRLTKVV	RRSLINLGRA
3910	3920	3930	3940	3950	3960

	IKGQVLMSSSE	LEEVFNMSLV	GKVPAMWAAK	SYPSLKPLGG	YVADLLARLT	FFQEWIDKGP
	3970	3980	3990	4000	4010	4020
	PVVFWISGFY	FTQSFLTGV	QNYARKYTIP	IDHIGFEFEV	TPQETVMENN	PEDGAYIKGL
	4030	4040	4050	4060	4070	4080
	FLEGARWDRK	TMQIGESLPK	ILYDPLPIIW	LKPGESAMFL	HQDIYVCPVY	KTSARRGTLS
	4090	4100	4110			
	TTGHSTNYVL	SIELPTDMPQ	KHWINRGVAS	LCQLDN		
Q18PE1	10	20	30	40	50	60
SEQ ID NO:2	MTEAALVEGQ	VKLRDGGKWK	SRWLVLKPKS	PVADCLLMV	YKDKSERIKG	LRERSSLTLE
	70	80	90	100	110	120
	DICGLEPGLP	YEGLVHTLAI	VCLSQAIMLG	FDSHEAMCAW	DARIRYALGE	VHRFHVTVAP
	130	140	150	160	170	180
	GTKLESGPAT	LHLCNDVLVL	ARDIPPAVTG	QWKLSDLRRY	GAVPSGFIFE	GGTRCGYWAG
	190	200	210	220	230	240
	VFFLSSAEGE	QISFLFDCIV	RGISPTKGP	GLRPVLPDPS	PPGPSTVEER	VAQEALETLO
	250	260	270	280	290	300
	LEKRLSLLSH	AGRPGSGGDD	RSLSSSSSEA	SHLDVSASSR	LTAWPEQSSS	SASTSQEGPR
	310	320	330	340	350	360
	PAAAQAAGEA	MVGASRPPPK	PLRPRQLQEV	GRQSSSDSGI	ATGSHSSYSS	SLSSYAGSSL
	370	380	390	400	410	420
	DVWRATDELG	SLLSLPAAGA	PEPSLCTCLP	GTVEYQVPTS	LRAHYDTPRS	LCLAPRDHSP
	430	440	450	460	470	480
	PSQGSFGNSA	ARDSGGQTS	GCPSGWLGR	RRGLVMEAPQ	GSEATLPGPA	PGEFWEAGGP
	490	500				
	HAGPPPAFFS	ACPVCGGLKV	NPPP			
Q8NFH5	10	20	30	40	50	60
SEQ ID NO:3	MAAFAVEPQG	PALGSEPMML	GSPTSPKPGV	NAQFLPGFLM	GDLPAVPTPQ	PRSIGPSVSG
	70	80	90	100	110	120
	VMEMRSPLLA	GGSPQPVPV	AHKDKSGAPP	VRSIYDDISS	PGLGSTPLTS	RRQPNISVMQ
	130	140	150	160	170	180
	SPLVGVTSTP	GTGQSMFSPA	SIGQPRKTTL	SPAQLDPFYT	QGDSLTSSEH	LDDSWVTVFG
	190	200	210	220	230	240
	FPQASASYIL	LQFAQYGNIL	KHVMSNTGNW	MHIRYQSKLQ	ARKALSKDGR	IFGESIMIGV
	250	260	270	280	290	300
	KPCIDKSVME	SSDRCALSSP	SLAFTPPIKT	LGTPTQPGST	PRISTMRPLA	TAYKASTSDY
	310	320				
	QVISDRQTPK	KDESLVSKAM	EYMFGW			
Q8WYL5	10	20	30	40	50	60
SEQ ID NO:4	MALVTLQRSP	TPSAASSSAS	NSELEAGSEE	DRKLNLSLSE	SFFMVKGAAAL	FLQQGSSPQG
	70	80	90	100	110	120
	QRSLOHPPKH	AGDLPQHLQV	MINLLRCEDR	IKLAVRLESA	WADRVRYMVV	VYSSGRQDTE
	130	140	150	160	170	180
	ENILLGVDFS	SKESKSTIG	MVRLWSDTK	IHLGDGGGFS	VSTAGRMHIF	KPVSVQAMWS
	190	200	210	220	230	240
	ALQVLHKACE	VARRHNYFPG	GVALIWATYY	ESCISSEQSC	INEWNAMQDL	ESTRPDSPAL
	250	260	270	280	290	300
	FVDKPTGER	TERLIKAKLR	SIMMSQDLEN	VTSKEIRNEL	EKQMNCLKE	LKEFIDNEML
	310	320	330	340	350	360
	LILGQMDKPS	LIFDHLYLGS	EWNASNLEEL	QSGVDYILN	VTREIDNFFP	GLFAYHNIRV
	370	380	390	400	410	420
	YDEETDILLA	HWNEAYHFIN	KAKRNHSHCL	VHCKMGVSR	ASTVIAYAMK	EFGWPLEKAY
	430	440	450	460	470	480
	NYVKQKRSIT	RPNAGFMRQL	SEYEGILDAS	KQRHNKLWRQ	QTDSSLQQP	DDPAGPGDFL
	490	500	510	520	530	540
	PETPDGTPES	QLPFLDDAAQ	PGLGPPLPCC	FRRLSDPLL	SPEDETGSV	HLEDPEREAL
	550	560	570	580	590	600
	LEEAAAPAEV	HRPARQPQQG	SGLCEKDVKK	KLEFGSPKGR	SGSLLQVEET	EREGLGAGR
	610	620	630	640	650	660
	WGQLPTQLDQ	NLLNSENINN	NSKRSCPNGM	EDDAIFGILN	KVKPSYKSCA	DCMYPTASGA
	670	680	690	700	710	720
	PEASRERCED	PNAPAICTQP	AFLPHITSSP	VAHLASRSRV	PEKPASGPT	PPFPLPPAGS
	730	740	750	760	770	780
	RRADTSGPGA	GAALEPPASL	LEPSRETQKV	LPKSLLKNS	HCDKNPPSTE	VVIKEESSPK
	790	800	810	820	830	840
	KDMKPAKDLR	LLFSNESEKP	TTNSYLMQHQ	ESIIQLQKAG	LVRKHTKELE	RLKSVPADPA

	850	860	870	880	890	900
	PPSRDGPASR	LEASIPEESQ	DPAALHELGP	LVMPSEQAGSD	EKSEAAPASL	EGGSLKSPPP
	910	920	930	940	950	960
	FFYRLDHTSS	FSKDFLKTIC	YTPTSSSMSS	NLTRSSSSDS	IHSVRGKPGI	VKQRTQEIET
	970	980	990	1000	1010	1020
	RLRLAGLTVS	SPLKRSHSLA	KLGLSLTFSTE	DLSSEADPST	VADSQDITLS	ESSFLHEPQG
	1030	1040				
	TPRDPAAATSK	PSGKPAPENL	KSPSWMSKS			
Q8IYD8	10	20	30	40	50	60
SEQ ID NO:5	MSGRQRTLQ	TWGSSISRSS	GTPGCSSGTE	RPQSPGSSKA	PLPAAAEAL	ESDDDVLLVA
	70	80	90	100	110	120
	AYEAERQLCL	ENGGFCTSG	ALWIYPTNCP	VRDYQLHISR	AALFCNTLVC	LPTGLGKTFI
	130	140	150	160	170	180
	AAVVMYNFYR	WFPSGKVVM	APTKPLVTQQ	IEACYQVMGI	PQSHMAEMTG	STQASTRKEI
	190	200	210	220	230	240
	WCSKRVLFLT	PQVMVNDLSR	GACPAAEIKC	LVIDEAHKAL	GNAYACQVVR	ELVKYTNHFR
	250	260	270	280	290	300
	ILALSATPGS	DIKAVQQVIT	NLLIGQIELR	SEDSPIILTY	SHERKVEKLI	VPLGEELAAI
	310	320	330	340	350	360
	QKTYIQILES	FARSLIQNV	LMRRDIPNLT	KYQIILARDQ	FRKNPSPNIV	GIQQGIIEGE
	370	380	390	400	410	420
	FAICISLYHG	YELLQOMGMR	SLYFFLCGIM	DGTMKMTSRK	NELGRNEDFM	KLYNHLECMF
	430	440	450	460	470	480
	ARTRSTSANG	ISAIQQGDKN	KKFVYSHPKL	KKLEEVVIEH	FKSWNAENTT	EKKRDETRVM
	490	500	510	520	530	540
	IFSSFRDSVQ	EIAEMLSQHQ	PIIRVMTFVG	HASGKSTKGF	TQKEQLEVVK	QFRDGGYNTL
	550	560	570	580	590	600
	VSTCVGEEGL	DIGEVDLIIC	FDSQKSPIRL	VQRMGRTGRK	RQGRIVILS	EGREERIYNQ
	610	620	630	640	650	660
	SQSNKRSIYK	AISSNRQVLH	FYQRSRPMVP	DGINPKLHKM	FITHGVYEPE	KPSRNLQRKS
	670	680	690	700	710	720
	SIFSRYDGM	QSSSLKDWFL	SEEEFKLWNR	LYRLRDSDEI	KEITLPQVQF	SSLQNEENKP
	730	740	750	760	770	780
	AQESTTGIHQ	LSLSEWRLWQ	DHPLPTHQVD	HSDRCRHFIF	LMQMIEGMRH	EEGECSEYLE
	790	800	810	820	830	840
	VESYLQMEDV	TSTFIAPRNE	SNNLASDTFI	THKKSSFIKN	INQSSSSSVI	ESDEECAEIV
	850	860	870	880	890	900
	KQTHIKPTKI	VSLKKKVSKE	IKKDQLKKEN	NHGIIDSVDN	DRNSTVENIF	QEDLPNDKRT
	910	920	930	940	950	960
	SDTDEIAATC	TINENVIKEP	CVLLTECQFT	NKSTSSLAGN	VLDSGYNSFN	DEKSVSSNLF
	970	980	990	1000	1010	1020
	LPFEEELYIV	RTDDQFYNCH	SLTKEVLANV	ERFLSYSPPP	LSGLSDLEYE	IAKGTALENL
	1030	1040	1050	1060	1070	1080
	LFLPCAHLR	SDKCTCLLSH	SAVNSQQNLE	LNSLKCINYP	SEKSCLYDIP	NDNISDEPSL
	1090	1100	1110	1120	1130	1140
	CDCDVHKNQ	NENLVPNNRV	QIHRSPAQNL	VGENNHDVDN	SDLPVLSTDQ	DESLLLFEDV
	1150	1160	1170	1180	1190	1200
	NTEFDDVSL	PLNSKSESLP	VSDKTAISET	PLVSQFLISD	ELLLDNNSSEL	QDQITRDANS
	1210	1220	1230	1240	1250	1260
	FKSRDQRGVQ	EEKVKNHEDI	FDCSRDLFSV	TFDLGFCSPD	SDDEILEHTS	DSNRPLDDLY
	1270	1280	1290	1300	1310	1320
	GRYLEIKEIS	DANYVSNQAL	IPRDHSKNFT	SGTVIIPSNE	DMQNPNIYVH	PLSAAKNEEL
	1330	1340	1350	1360	1370	1380
	LSPGYSQFSL	PVQKKVMSTP	LSKSNLTNSF	SKIRKEILKT	PDSSKEKVNL	QRFKEALNST
	1390	1400	1410	1420	1430	1440
	FDYSEFSLEK	SKSSGPMYH	KSCHSVEDGQ	LLTSNESEDD	EIFRRKVKRA	KGNVLNSPED
	1450	1460	1470	1480	1490	1500
	QKNSEVDSPL	HAVKKRRFPI	NRSELSSSDE	SENFPPKPCSQ	LEDFKVCNGN	ARRGIKVPKR
	1510	1520	1530	1540	1550	1560
	QSHLKHVARK	FLDDEAELSE	EDAIEYVSDE	NDESENEQDS	SLLDFLNDDET	QLSQAINDSE
	1570	1580	1590	1600	1610	1620
	MRAIYMKSLR	SPMMNNKYKM	IHKTHKNINI	FSQIPEQDET	YLEDSEFCVDE	EESCKGQSSE
	1630	1640	1650	1660	1670	1680
	EEVCVDFNLI	TDDCFANSKK	YKTRRAVMLK	EMMEQNCAHS	KKKLSRIILP	DDSSEENNV
	1690	1700	1710	1720	1730	1740
	NDKRESNIAV	NPSTVKKNKQ	QDHCLNSVPS	GSSAQSKVRS	TPRVNPLAKQ	SKQTSNLNLD
	1750	1760	1770	1780	1790	1800

	TISEVSDFKP QNHNEVQSTT PPFTTVDSQK DCRKFPVPQK DGSALSDSST SGASCSKSRP 1810 1820 1830 1840 1850 1860 HLAGTHTSLR LPQEGKGTCTI LVGGHEITSG LEVISSLRAI HGLQVEVCPL NGCDYIVSNR 1870 1880 1890 1900 1910 1920 MVVERRSQSE MLNSVNKNKF IEQIQHLQSM FERICVIVEK DREKTGDTSR MFRRTKSYS 1930 1940 1950 1960 1970 1980 LLTTLIGAGI RILFSSCQEE TADLLKELSL VEQRKNVGIH VPTVVNSNKS EALQFYLSIP 1990 2000 2010 2020 2030 2040 NISYITALNM CHQFSSVKRM ANSSLQEISM YAQVTHQKAE EIYRYIHYVF DIQMLPNDLN QDRLKSDI
O14654 SEQ ID NO:6	10 20 30 40 50 60 MASCSTTRDQ ATRRLRGAAA AAAAALAAVV TTPLLSSGTP TALIGTGSSC PGAMWLSTAT 70 80 90 100 110 120 GSRSDSESEE EDLPVGEEVC KRGYLRKQKH GHRRYFVLKL ETADAPARLE YYENARKFRH 130 140 150 160 170 180 SVRAAAAAA AAASGAAIPP LIPPRRVITL YQCFSVSQRA DARYRHLIAL FTQDEYFAMV 190 200 210 220 230 240 AENESEQESW YLLLSRLILE SKRRRCGTLG AQPDPGEPAAL AAAAAAEPF YKDVWQVIVK 250 260 270 280 290 300 PRGLGHRKEL SGVFRCLCLTD EEVVFVRLNT EVASVVVQLL SIRRCGHSEQ YFFLEVGRST 310 320 330 340 350 360 VIGPGELWMQ VDDCVVAQNM HELFLEKMRA LCADEYRARC RSYISISGAH LLTLLSARRH 370 380 390 400 410 420 LGLVPLEPGG WLRRSRFEQF CHLRAIGDGE DEMLFTRRFV TPSEPVAHSR RGRHLHPRGR 430 440 450 460 470 480 RSRRRAVSVPA SFFRRLAPSP ARPRHPAEAP NNGARLSSEV SGSGSGNFGE EGNPQKGEDQ 490 500 510 520 530 540 EGSGGDYMPM NNWGSNGRG SGGGQGSNGQ GSSSHSSGGN QCSGEGQGSR GGQGSNGQGS 550 560 570 580 590 600 GGNQCSRDOG GTAGGHGSGG GQRPGGGHGS GGGQPGDGH GSGGKNSGG KGSGSGKGS 610 620 630 640 650 660 DGDGERGKSL KKRSYFGKLT QSKQQQMPPP PPPPPPPPA GGTGGKGKSG GRFRLYFCVD 670 680 690 700 710 720 RGATKECKEA KEVKDAEIP EGAARGPHRAR AFDEDEDDPY VPMRPGVATP LVSSSDYMPM 730 740 750 760 770 780 APQNVASAKK RHRSPFEDS RGYMMMFPRV SPPAPSPPK APDTNKEDDS KDNDSESDYM 790 800 810 820 830 840 FMAPGAGAIP KNPRNPQGS SSKSWSSYFS LPNPFSSPL QQNDNSEYVP MLPKGFLGRG 850 860 870 880 890 900 LDKEVSYNWD PKDAASKPSG EGSFSKPGDG GSPSKPSDHE PPKNKAKRPN RLSFITKGYK 910 920 930 940 950 960 IKPKPQKPTH EQREADSSD YVNMDFTKRE SNTAPSTQG LPDSWGIIAE PRQSASFNYV 970 980 990 1000 1010 1020 NVEFGVPFPN PANDLSDLLR AIPRANPLSL DSARWPLPPL PLSATGSNAI EEEDGYIEVI 1030 1040 1050 1060 1070 1080 FNSAMTPAMA LADSAIRYDA ETGRIYVVD FSECCMDISL SPSRCSEPPP VARLLQEEEQ 1090 1100 1110 1120 1130 1140 ERRRPQSRQ SFFAAARA AV SAFPTDSLRL DLSPSSAPAV ASAAEPTLAL SQVVAASAL 1150 1160 1170 1180 1190 1200 AAAPGIGAAA AAAGFDSASA RWFQPVANAA DAEAVRGAQD VAGGSNPGAH NPSANLARGD 1210 1220 1230 1240 1250 NQAGGAAAAA AAPEPPPSR RVPRPPERED SDNDDTHVR MDFARRDNQF DSPKRGR
Q96AP4 SEQ ID NO:7	10 20 30 40 50 60 MLSCNICGET VTSEPDMAH LIVHMESEII CPFCKLSGVN YDEMCFHIE AHFEQNTLER 70 80 90 100 110 120 NFERINTVQY GTSDNKKDNT LQCGMEVNSS ILSGCASNHP KNSAQNLTKD STLKHEGFYS 130 140 150 160 170 180 ENLTESRKFL KSREKQSSLT EIKGSVYETT YSPPECPCFG KIEEHSEDME THVKTKHANL 190 200 210 220 230 240 LDIPLEDQDQ PLYDCPMCGL ICTNYHILQE HVDLHLEENS FQQGMDRVQC SGDLQLAHQL 250 260 270 280 290 300 QQEEDRKRRS EESRQEIEEF QKLQRQYGLD NSGGYKQQQL RNMEIEVNRG RMPPSEFHRR 310 320 330 340 350 360 KADMESLAL GFDDGKTKTS GIIEALHRY YQNAATDVRRV WLSSVVDHFFH SSLGDKGWGC

	370	380	390	400	410	420
	GYRNFQMLLS	SLNQNDAYND	CLKGMLIPCI	PKIQSMIEDA	WKEGFDPQGA	SQLNNRLQGT
	430	440	450	460	470	480
	KAWIGACEVY	ILLTSLRVKC	HIVDFHKSTG	PLGTHPRLFE	WILNYYSSG	EGSPKVVCTS
	490	500	510	520	530	540
	KPPIYLQHQQ	HSRTVIGIEE	KKNRTLCLLI	LDPGCPSPREM	QKLLKQDIEA	SSLKQLRKSM
	550	560	570			
	GNLKHKQYQI	LAVEGALSLE	EKLARRQASQ	VFTAEEKIP		
Q9UQ35 SEQ ID NO:8	10	20	30	40	50	60
	MYNGIGLPTP	RSGGTNGYVQ	RNLSLVRGRR	GERPDYKGEE	ELRRLEAALV	KRPNPDILDH
	70	80	90	100	110	120
	ERKRRVELRC	LELEEMMEEQ	GYEEQQIQEK	VATFRLMLLE	KDVNPGGKEE	TPGQRPVATE
	130	140	150	160	170	180
	THQLAELNEK	KNERLRAAFG	ISDSYVDGSS	FDPQRRAREA	KQPAPEPPKP	YSLVRESSSS
	190	200	210	220	230	240
	RSPTPKQKKK	KKKKDRGRRS	ESSSPRRERK	KSSKKKKHRS	ESESKKRKHR	SPTPKSKRKS
	250	260	270	280	290	300
	KDKKRKRKRS	TTPAPKSRRA	HRSTSADSAS	SSDTSRSPSR	SAAAKTHTTA	LAGRSPSPAS
	310	320	330	340	350	360
	GRRGEGDAPF	SEPGTTSTQR	PSSPETATKQ	PSSPYEDKDK	DKKEKSATRP	SPSPERSSTG
	370	380	390	400	410	420
	PEPPAPTPLL	AERHGGSPQP	LATTPLSQEP	VNPPSEASPT	RDRSPPKSPE	KLPQSSSSSES
	430	440	450	460	470	480
	SPPSPQPTKV	SRHASSSPES	PKPAPAPGSH	REISSSPTSK	NRSHGRAKRD	KSHSHTPSRR
	490	500	510	520	530	540
	MGRSRSPATA	KRGRSRSRTP	TKRGHSRSRS	PQWRRSRSAQ	RWGRSRSPQR	RGRSRSPQRP
	550	560	570	580	590	600
	GWSRSRNTQR	RGRSRSAARR	RSHSRSPATR	GRSRSRTPAR	RGRSRSRTPA	RRRSRSTRPT
	610	620	630	640	650	660
	RRRSRSTRPA	RRGRSRSRTP	ARRRSRTRSP	VRRRSRSPSR	ARRSGRSRSP	TPARRGRSRS
	670	680	690	700	710	720
	RTPARRGRSR	SRTPARRSGR	SRSRTPARRG	RSRSTRPRRG	RSRSTRSLVRR	GRSHSRTPQR
	730	740	750	760	770	780
	RGRSGSSSER	KNKSRTSQRR	SRSNSSPEMK	KSRISSRRSR	SLSSPRSKAK	SRLSLRRSLS
	790	800	810	820	830	840
	GSSPCPKQKS	QTPPRRSRSG	SSQPKAKSRT	PPRRSRSSSS	PPPKQKSKTP	SRQSHSSSSP
	850	860	870	880	890	900
	HPKVKSGTPP	RQGSITSPQA	NEQSVTPQRR	SCFESSPDPE	LKSRTPSRHS	CSGSSPPRVK
	910	920	930	940	950	960
	SSTPPRQSPS	RSSSPQPKVK	AIISPRQRSH	SGSSSPSPSR	VTSRTTPRRS	RSVSPCSNVE
	970	980	990	1000	1010	1020
	SRLLPYRSHS	GSSSPDTKVK	PETPPRQSHS	GSISPYPKVK	AQTPPGPSLS	GSKSPCPQEK
	1030	1040	1050	1060	1070	1080
	SKDSLQVQSCP	GSLSLCAGVK	SSTPPGESYF	GVSSLQLKGQ	SQTSPDHRSD	TSSPEVRQSH
	1090	1100	1110	1120	1130	1140
	SESPSLQSKS	QTSPKGGRRS	SSSPVTELAS	RSPIRQDRGE	FSASPMLKSG	MSPEQSRFQS
	1150	1160	1170	1180	1190	1200
	DSSSYPTVDS	NSLLGQSRLE	TAESKEKMAL	PPQEDATASP	PRQKDKFSPF	PVQDRPESSL
	1210	1220	1230	1240	1250	1260
	VFKDTLRTTP	RERSGAGSSP	ETKEQNSALP	TSSQDEELME	VVEKSEEPAG	QILSHLSSEL
	1270	1280	1290	1300	1310	1320
	KEMSTSNFES	SPEVEERPAP	SLTLDQSQSQ	ASLEAVEVPS	MASSWGPHF	SPEHKELSNS
	1330	1340	1350	1360	1370	1380
	PLRENSFGSP	LEFRNSGPLG	TEMNTGFSSE	VKEDLNGPFL	NQLETDPPLD	MKEQSTRSSG
	1390	1400	1410	1420	1430	1440
	HSSSELSPDA	VEKAGMSSNQ	SISSPVLDV	PRTPSRERSS	SASSPEMKDG	LPRTPSRRSR
	1450	1460	1470	1480	1490	1500
	SGSSPGLRDG	SGTPSRHSLS	GSSPGMKDIP	RTPSRGRSEC	DSSPEPKALP	QTPRPRSRSP
	1510	1520	1530	1540	1550	1560
	SSPELNNKCL	TPQRRERSGSE	SSVDQKTIVAR	TPLGQRRSRG	SSQELDVKPS	ASPQERSESD
	1570	1580	1590	1600	1610	1620
	SSPDSKAKTR	TPLRQRRSRG	SSPEVDSKSR	LSPRRSRSGS	SPEVKDKPRA	APRAQSGSDS
	1630	1640	1650	1660	1670	1680
	SPEPKAPAPR	ALPRRSRSGS	SSKGRGPSPE	GSSSTESSPE	HPPKSRTARR	GSRSSEPCKT
	1690	1700	1710	1720	1730	1740
	KSRTPPRRRS	SRSSPELTRK	ARLSRRSRSA	SSSPETRST	PPRHRRSPSV	SSPEPAEKSR
	1750	1760	1770	1780	1790	1800

	SSRRRRSASS	PRTKTTSSRG	RSPSPKPRGL	QRSRSRSRRE	KTRTTRRRDR	SGSSQSTSR
	1810	1820	1830	1840	1850	1860
	RQSRSRSRV	TRRRRGSGY	HSRSPARQES	SRTSSRRRG	RSRTPTSRK	RSRSRTSPAP
	1870	1880	1890	1900	1910	1920
	WKRSRSRASP	ATHRRSRRT	PLISRRRSRS	RTSPVSRRS	RSRTSVTRRR	SRSRASPVSR
	1930	1940	1950	1960	1970	1980
	RRSRRTTPPV	TRRRSRRTTP	TTRRRSRRT	PPVTRRRRS	RTPPVTRRRS	RSRTSPITRR
	1990	2000	2010	2020	2030	2040
	RSRSRTSPVT	RRRSRTSP	VTRRRSRRT	SPVTRRRRS	RTPPAIRRRS	RSRTPLLPK
	2050	2060	2070	2080	2090	2100
	RSRSRSPAL	RRRSRTTPR	TARGKRSR	SPPAIRRRSA	SGSSSDRSR	ATPPATRNHS
	2110	2120	2130	2140	2150	2160
	GSRTPPVALN	SSRMSCFSRP	SMSPTPLDRC	RSPGMLEPLG	SSRTPMSVLQ	QAGGSMDGP
	2170	2180	2190	2200	2210	2220
	GPRIPDHQRT	SVPENHAQSR	IALALTAISL	GTARPPPSMS	AAGLAARMSQ	VPAPVPLMSL
	2230	2240	2250	2260	2270	2280
	RTAPAAANLAS	RIPAASAAAM	NLASARTPAI	PTAVNLADSR	TPAAAAAMNL	ASPRTAVAPS
	2290	2300	2310	2320	2330	2340
	AVNLADPRTP	TAPAVNLAGA	RTPAALAALS	LTGSGTPPTA	ANYPSSSRTP	QAPASANLVG
	2350	2360	2370	2380	2390	2400
	PRSAHATAPV	NIAGSRTAAA	LAPASLTSAR	MAPALSGANL	TSPRVPLSAY	ERVSGRTSPP
	2410	2420	2430	2440	2450	2460
	LLDRARSRT	PSAPSQSRMT	SERAPSPSSR	MGQAPSQSL	PPAQDQPRSP	VPSAFSDQSR
	2470	2480	2490	2500	2510	2520
	CLIAQTTPVA	GSQSLSSGAV	ATTSSAGDH	NGMLSVAPAG	VPHSDVGEPP	ASTGAQQPSA
	2530	2540	2550	2560	2570	2580
	LAALQPAKER	RSSSSSSSSS	SSSSSSSSSS	SSSSSSGSSS	SDSEGSSLPV	QPEVALKRPV
	2590	2600	2610	2620	2630	2640
	SPTPAPKEAV	REGRPPEPTP	AKRKRSSSSS	SSSSSSSSSS	SSSSSSSSSS	SSSSSSSSSS
	2650	2660	2670	2680	2690	2700
	SSSSSSSSPS	PAKPGPQALP	KPASPKKPPP	GERRSRSPRK	PIDSLRDSRS	LSYSPVERRR
	2710	2720	2730	2740	2750	
	PSPQPSPRDQ	QSSSSERGSR	RGQRGDSRSP	SHKRRRETPS	PRPMHRSSR	SP
Q8N6W0	10	20	30	40	50	60
SEQ ID NO:9	MARLSEAR	RQQQQLQPR	PSPVGSSGPE	PPGGQPDGMK	DLDAIKLFVG	QIPRHLDEKD
	70	80	90	100	110	120
	LKPLFEQFGR	IYELTVLKDP	YTMHKGCAF	LYCARDIAI	KAQTALHEQK	TLPGMARPIQ
	130	140	150	160	170	180
	VKPADSESRG	GRDRKLFVGM	LNKQQSEEDV	LRLFQPFQVI	DECTVLRGPD	GSSKGCAFVK
	190	200	210	220	230	240
	FSSHTAQAA	IHALHGSQTM	PGASSSLVVK	FADTDKERTL	RRMQQMVQQL	GILTPSLTLP
	250	260	270	280	290	300
	FSPYSAYAQA	LMQQQTTVLS	TSGSYLSPGV	AFSPCHIQI	GAVSLNLGLP	TPIAPASGLH
	310	320	330	340	350	360
	SPPLLGTTAV	PGLVAPITNG	FAGVVPFPGG	HPALETVYAN	GLVPYPAQSP	TVAETLHPAF
	370	380	390	400	410	420
	SGVQQYTAMY	PTAAITPIAH	SVPQPPPLQ	QQQREGPEGC	NLFYIHLPE	FGDTELTQMF
	430	440	450	460	470	480
	LPFGNISSK	VFMDRATNQS	KCFGFVSFDN	PASQAQAIQA	MNGFQIGMKR	LKVQLKRPKD
	PGHPY					
Q9H792	10	20	30	40	50	60
SEQ ID NO:10	MSACNTFTEH	VWKPGECKNC	FKPKSLHQLP	PDPEKAPITH	GNVKTANANHS	NNHRIRNTGN
	70	80	90	100	110	120
	FRPPVAKKPT	IAVKPTMIVA	DGQSICGELS	IQEHCENKPV	IIGWNRNRRA	LSQKPLNNNN
	130	140	150	160	170	180
	EDDEGISHVP	KPYGNNDASAK	KMSDNNGGLT	EVLKEIAGLD	TAPQIRGNET	NSRETFLGRI
	190	200	210	220	230	240
	NDCYKRSR	KLPPSCMIGG	IKETQGHVI	LSGSTEVISN	EGGRFCYPEF	SSGEESEEDV
	250	260	270	280	290	300
	LFSNMEEEHE	SWDESDEELL	AMEIRMRGQP	RFANFRANTL	SPVRFFVDKK	WNTIPLRNKS
	310	320	330	340	350	360
	LQRICAVD	DSYDEILNGY	EENSVMVSYGQ	GSIQSMVSSD	STSPDSSLTE	ESRSETASSL
	370	380	390	400	410	420
	SQKICNGGLS	PGNPGDSKDM	KEIEPNYESP	SSNNQDKDSS	QASKSSIKVP	ETHKAVLALR

	430	440	450	460	470	480
LEEKDGKIAV	QTEKEESKAS	TDVAGQAVTI	NLVPTEEQAK	PYRVVNLEQP	LCKPYTVVDV	
490	500	510	520	530	540	
SAAMASEHLE	GPVNSPKTKS	SSSTPNSPVT	SSSLTPGQIS	AHFQKSSAIR	YQEVWTSSTS	
550	560	570	580	590	600	
PRQKIPKVEL	ITSGTGNVP	PRKNCHKSAP	TSPTATNISS	KTIPVKSPNL	SEIKFNSYNN	
610	620	630	640	650	660	
AGMPFPPIII	HDEPTYARSS	KNAIKVPIVI	NPAYDNLAI	YKSFLGTSGE	LSVKEKTTSV	
670	680	690	700	710	720	
ISHTYEEIET	ESKVPDNTTS	KTDCLOTKG	FSNSTEHLRG	SVAQKVQEFN	NCLNRGQSSP	
730	740	750	760	770	780	
QRSYSSSHSS	PAKIQRATQE	PVAKIEGTQE	SQMVGSSTR	EKASTVLSQI	VASIQQPPQSP	
790	800	810	820	830	840	
PETPQSGPKA	CSVEELYAIP	PDADVAKSTP	KSTPVRPKSL	FTSQPSGEAE	APQTTDSPTT	
850	860	870	880	890	900	
KVQKDPSIKP	VTPSPSKLVT	SPQSEPPAPF	PPRSTSSPY	HAGNLLQRHF	TNWTKPTSPT	
910	920	930	940	950	960	
RSTEAEVSLH	SEGSRRADA	KPKRWISFKS	FFRRRKTDEE	DDKEKEREGK	KLVGLDGTVI	
970	980	990	1000	1010	1020	
HMLPPPPVQR	HHWFTEAKGE	SSEKPAIVFM	YRCDPAQQQL	SVDQSKARTD	QAAVMEKGRA	
1030	1040	1050	1060	1070	1080	
ENALLQDSEK	KRSHSSPSQI	PKKILSHMTH	EVTEDFSRPR	PRTVVGKQDG	RGCTSVTTAL	
1090	1100	1110	1120	1130	1140	
SLPELEREDG	KEDISDPMDP	NPCSATYSNL	GQSRAAMIPP	KQPRQPKGAV	DDAIAFGGKT	
1150	1160	1170	1180	1190	1200	
DQEAPNASQP	TPPPLPKMI	IRANTEPISK	DLQKSMESL	CVMANPTYDI	DPNWDASSAG	
1210	1220	1230	1240	1250	1260	
SSISYELKGL	DIESYDSLRL	PLRKERPVP	AANSISLTT	LSIKDRFSNS	MESLSSRRGP	
1270	1280	1290	1300	1310	1320	
SCRQGRGIQK	PQRQALYRGL	ENREEVVGKI	RLHTDALKK	LAVKCEDLFM	AGQKDQLRFG	
1330	1340	1350	1360	1370	1380	
VDSWSDFRLT	SDKPCCEAGD	AVYYTASYAK	DPLNNYAVKI	CKSKAKESQQ	YYHSLAVRQS	
1390	1400	1410	1420	1430	1440	
LAVHFNIQD	CGHFLAEVNP	RLLPWEDPDD	PEKDEDDMEE	TEEDAKGETD	GKNPKPCSEA	
1450	1460	1470	1480	1490	1500	
ASSQKENQGV	MSKKQRSHVV	VITREVPCLT	VADFVRDSL	QHGKSPDLYE	RQVCLLLLQL	
1510	1520	1530	1540	1550	1560	
CSGLEHLKPY	HVTHCDLRLE	NLLLVHYQPG	GTAQGFQPAE	PSPTSSYPTR	LIVSNFSQAK	
1570	1580	1590	1600	1610	1620	
QKSHLVDPEI	LRDQSRLAPE	IITATQYKKC	DEFQTGILY	EMLHLPNPF	ENPELKEREY	
1630	1640	1650	1660	1670	1680	
TRADLPRIPF	RSPYSRGLQQ	LASCLLNPNP	SERILISDAK	GILQCLLWGP	REDLFQTF	
1690	1700	1710	1720	1730	1740	
CPSLVQRNTL	LQNWLDIKRT	LLMIKFAEKS	LDREGGISLE	DWLCAQYLAF	ATTDSLSCIV	
KILQHR						
Q9H497	10	20	30	40	50	60
SEQ ID NO:11	MLRGPWRQLW	LFFLLLLPGA	PEPRGASRPW	EGTDEPGSAW	AWPGFQRLQE	QLRAAGALSK
	70	80	90	100	110	120
	RYWTLFSCQV	WPDDCDEDEE	AATGPLGWRL	PLLGQRYLDL	LTTWYCSFKD	CCPRGDCRIS
	130	140	150	160	170	180
	NNFTGLEWDL	NVRLHGQHLV	QQLVLRTRVG	YLETPOPEKA	LALSFGHWSG	TGKNFVARM
	190	200	210	220	230	240
	VENLYRDGLM	SDCVRMFIAI	FHFPHPKYVD	LYKEQLMSQI	RETQQLCHQT	LFIFDEAEKL
	250	260	270	280	290	300
	HPGLLEVLGP	HLERRAPEGH	RAESPWTIFL	FLSNLRGDII	NEVVLKLLKA	GWSREEITME
	310	320	330	340	350	360
	HLEPHLQAEI	VETIDNGFGH	SRLVKENLID	YFIPFLPLEY	RHVRLCARDA	FLSQELLYKE
	370	380	390			
	ETLDEIAQMM	VYVPKEEQLF	SSQGCKSISQ	RINYFLS		
Q9UE35	10	20	30	40	50	60
SEQ ID NO:12	MTMTLHTKAS	GMALLHQIQG	NELEPLNRPQ	LKIPLERPLG	EVYLDSSKPA	VYNYPEGAAY
	70	80	90	100	110	
	EFNAAAAANA	QVYGQTGLPY	GPGSEAAAFG	SNGLGGFPPL	NSVSPSPLML	LHPPP

O00743 SEQ ID NO:13	10	20	30	40	50	60
	MAPLDDLKYV	EIARLCKYLP	ENDLKRLCDY	VCDLLEESN	VQPVSTPVTV	CGDIHQFYD
	70	80	90	100	110	120
	LCELFRITGGQ	VPDTNYIFMG	DFVDRGYISL	ETFTYLLALK	AKWPDRITLL	RGNHESRQIT
	130	140	150	160	170	180
	QVYGFYDECQ	TKYGNANAWR	YCTKVFDMLT	VAALIDEQIL	CVHGGSLSPDI	KTLDQIRTIE
	190	200	210	220	230	240
	RNQEIPHKGA	FCDLVWSDPE	DVDTWAISPR	GAGWLFQAKV	TNEFVHINNL	KLICRAHQLV
	250	260	270	280	290	300
	HEGYKFMFDE	KLVTVWSAPN	YCYRCGNIAS	IMVFKDVNTR	EPKLFRAVPD	SERVIPRRTT
	TPYFL					
Q8WXF8 SEQ ID NO:14	10	20	30	40	50	60
	MALSGSTPAP	CWEEDECLDY	YGMLSLHRMF	EVVGGQLTEC	ELELLAFLLD	EAPGAAGGLA
	70	80	90	100	110	120
	RARSGLELLL	ELERRGQCDE	SNLRLLGQLL	RVLARHDLLP	HLARKRRRPP	SPERYSYGTS
	130	140	150	160	170	180
	SSSKRTEGSC	RRRRQSSSSA	NSQQGQWETG	SPPTKRQRRS	RGRPSGGARR	RRRGAPAAPQ
	190	200	210	220	230	240
	QQSEPARPSS	EGKVTCDIRL	RVRAEYCEHG	PALEQGVASR	RPQALARQLD	VFGQATAVLR
	250	260	270	280	290	300
	SRDLGSVVCD	IKFSELSYLD	AFWGDYLSGA	LLQALRGVFL	TEALREAVGR	EAVRLLVSV
	310	320				
	EADYEAGR	RR	LLMEEEGGR	RPTEAS		
P81274 SEQ ID NO:15	10	20	30	40	50	60
	MEENLISMRE	DHSFHVRYRM	EASCLELAL	GERLCKSGDC	RAGVSFFFAA	VQVGTEDLKT
	70	80	90	100	110	120
	LSAIYSQIGN	AYFYLHDYAK	ALEYHHHDLT	LARTIGDQLG	EAKASGNLGN	TLKVLGNFDE
	130	140	150	160	170	180
	AIVCCQRHLD	ISRELNDKVG	EARALYNLGN	VYHAKGKSFG	CPGPQDVGEF	PEEVVDALQA
	190	200	210	220	230	240
	AVDFYEENLS	LVTALGDRAA	QGRAFGNLGN	THYLLGNFRD	AVIAHEQRL	IAKEFGDKAA
	250	260	270	280	290	300
	ERRAYSNLGN	AYIFLGEFET	ASEYYKKTLL	LARQLKDRAV	EAQSCYSLGN	TYTLLQDYEK
	310	320	330	340	350	360
	AIDYHLKHLA	IAQELNDRIG	EGRACWSLGN	AYTALGNHDQ	AMHFAEKHLE	ISREVGDKSG
	370	380	390	400	410	420
	ELTARLNLS	D	LQMVGLSYS	TNNSIMSENT	EIDSSLNGVR	PKLGRRHSM
	430	440	450	460	470	480
	KVQNWNSIL	AKQKPLIAKP	SAKLLFVNRL	KGKKYKTNS	TKVLQDASNS	IDHRIPNSQR
	490	500	510	520	530	540
	KISADTIGDE	GFFDLLSRFQ	SNRMDDQRCC	LQEKNCHTAS	TTTSSTPPKM	MLKTSSVPVV
	550	560	570	580	590	600
	SPNTDEFLDL	LASSQSRRLD	DQRASFSNLP	GLRLTQNSQS	VLSHLMTNDN	KEADEDFFDI
	610	620	630	640	650	660
	LVKCQGSRLD	DQRCAPPPAT	TKGPTVPDED	FFSLILRSQG	KRMDEQRVLL	QRDQNRDTDF
	670	680				
	GLKDFLQNN	A	LLEFKNSGKK	SADH		
Q8NG08 SEQ ID NO:16	10	20	30	40	50	60
	MARSSPYLRQ	LQGPLLPPRD	LVEEDDDYLN	DDVEEDES	FIDAEELCSG	GVKAGSLPGC
	70	80	90	100	110	120
	LRVSICDENT	QETCKVFGFR	PITGAWWRVK	VQVKPVVGS	SYQYQVQGF	SYFLQSDMSP
	130	140	150	160	170	180
	PNQKHICALF	LKECEVSSDD	VNKFLTWVKE	VSNYKNLNFE	NLRETLRTFH	KETGRKDQKQ
	190	200	210	220	230	240
	PTQNGQEELF	LDNEMSLPLE	NTIPFRNVMT	ALQFPKIMEF	LPVLLPRHFK	WIIGSGSKEM
	250	260	270	280	290	300
	LKEIEEILGT	HPWKLGFSKI	TYREWKLRLC	EASWIAFCQC	ESLLQLMTDL	EKNALIMYSR
	310	320	330	340	350	360
	LKQICREDGH	TYVEVNDLTL	TLSNHMSFHA	ASESLKFLKD	IGVVTYEKSC	VFPYDLYHAE
	370	380	390	400	410	420
	RAIAFSICDL	MKKPPWHLCV	DVEKVLASIH	TKPENSSDD	ALNESKPDEV	RLENPVDVVD
	430	440	450	460	470	480
	TQDNGDHIWT	NGENEINAEI	SEVQLDQDQV	EVPLDRDQVA	ALEMICSNPV	TVISGKGGCG
	490	500	510	520	530	540

	KTIVSRLFK	HIEQLEEREV	KKACEDFEQD	QNASEEWITF	TEQSQLEADK	AIEVLLTAPT
	550	560	570	580	590	600
	GKAAGLLRQK	TGLHAYTLCQ	VNYSFYSWTQ	TMMTTNKPWK	FSSVRVLVVD	EGSLVSVGIF
	610	620	630	640	650	660
	KSVLNLCEH	SKLSKLIILG	DIRQLPSIEP	GNLLKDLFET	LKSRNCAIEL	KTNHRAESQL
	670	680	690	700	710	720
	IVDNATRISR	RQFPKFDAEL	NISDNPTLPI	SIQDKTFIFV	RLPEEDASSQ	SSKTNHHSCL
	730	740	750	760	770	780
	YSAVKTLQEQ	NNLQNAKTSQ	FIAFRRQDCD	LINDCCCKHY	TGHLTKDHQS	RLVFGIGDKI
	790	800	810	820	830	840
	CCTRDAYLSD	LLPENISGSQ	QNNDLASSE	DFSGTLPDFA	KNKRDFESNV	RLCNGEIFFI
	850	860	870	880	890	900
	TNDVTDVTFG	KRRSLTINNM	AGLEVTVDK	KLMKYCRIKH	AWARTIHTFQ	GSEEQTVVYV
	910	920	930	940	950	960
	VGKAGRQHWQ	HVYTAVTRGR	CRVYVIAEES	QLRNAIMKNS	FPRKTRLKHF	LQSKLSSSGA
	970	980	990	1000	1010	1020
	PPADFPSPRK	SSGDSGGPST	PSASPLPVVT	DHMTNDVTW	SEASSPDERT	LTFAERWQLS
	1030	1040	1050	1060	1070	1080
	SPDGVDTDDD	LPKSRASKRT	CGVNDDSPS	KIFMVGESPO	VSSRLQNLRL	NNLIPRQLFK
	PTDNQET					
Q96AE7	10	20	30	40	50	60
SEQ ID NO:17	MAAAVGVRGR	YELPPCSGPG	WLLSLSALLS	VAARGAFATT	HWVVTEDGKI	QQQVDSMPNL
	70	80	90	100	110	120
	KHPHDLVILM	RQEATVNYLK	ELEKQLVAQK	IHIEENEDRD	TGLEQRHNKE	DPDCIKAKVP
	130	140	150	160	170	180
	LGDLDLYDGT	YITLESKDIS	PEDYIDTESP	VPPDPEQPDC	TKILELPYSI	HAFQHLRGVQ
	190	200	210	220	230	240
	ERVNLSAPLL	PKEDPIFTYL	SKRLGRSIDD	IGHLIHEGLQ	KNTSSWVLN	MASFYWRICK
	250	260	270	280	290	300
	EPYQVVECAM	RALHFSSRHN	KDIALVNLAN	VLHRAHFSAD	AAVVVHAALD	DSDDFTSYT
	310	320	330	340	350	360
	LGNIYAMLGE	YNHSLVLCYD	ALQARPGFEQ	AIKRKHAVLC	QQKLEQKLEA	QHRSLQRTLN
	370	380	390	400	410	420
	ELKEYQKQHD	HYLRQOEILE	KHKLIQEEQI	LRNIIHETQM	AKEAQLGNHQ	ICRLVNQQHS
	430	440	450	460	470	480
	LHCQWDQVPR	YHRGDIFENV	DYVQFGEDSS	TSSMMSVNF	VQSNQSDIND	SVKSSPVAHS
	490	500	510	520	530	540
	ILWIWGRDSD	AYRDKQHILW	PKRADCTESY	PRVPVGGELP	TYFLPPENKG	LRIHELSSDD
	550	560	570	580	590	600
	YSTEEREAQTP	DCSITDFRKS	HTLSYLVKEL	EVRMDLKAKM	PDDHARKILL	SRINNYTIPE
	610	620	630	640	650	660
	EEIGSFLFHA	INKPNAPIWL	ILNEAGLYWR	AVGNSTFAIA	CLQRALNLAP	LQYQDVPLVN
	670	680	690	700	710	720
	LANLLIHYGL	HLDATKLLQ	ALAINSSEPL	TFLSLGNAYL	ALKNISGALE	AFRQALKLTT
	730	740	750	760	770	780
	KCPECENSLK	LIRCMQFYPF	LYNITSSVCS	GTVVEESNGS	DEMENSDETK	MSEEILALVD
	790	800	810	820	830	840
	EFQQAWPLEG	FGGALEMKGR	RLDLQGIRVL	KKGPDGVAR	SSCYGDCRSE	DDEATEWITF
	850	860	870	880	890	900
	QVKRVKPKG	DHKKTPGKKV	ETGQIENGHR	YQANLEITGP	KVASPGPGK	KRDYQRLGWP
	910	920	930	940	950	960
	SPDECLKLRL	VELTAIVSTW	LAVSSKNIDI	TEHIDFATPI	QQPAMEPLCN	GNLPTSMHTL
	970	980	990	1000	1010	1020
	DHLHGVSNNR	SLHYTGESQL	TEVLQNLGKD	QYPQQSLEQI	GTRIAKVLEK	NQTSWVLSSM
	1030	1040	1050	1060	1070	1080
	AALYWRVKGQ	GKKAIDCLRQ	ALHYAPHQMK	DVPLISLANI	LHNAKLWDA	VIVATMAVEI
	1090	1100	1110	1120	1130	1140
	APHFAVNHFT	LGNVYVAMEE	FEKALVWYES	TLKLQPEFVP	AKNRIQTIQC	HLMLKKGRRS
	P					
Q9BZM4	10	20	30	40	50	60
SEQ ID NO:18	MAAAASPAIL	PRLAILPYLL	FDWSGTGRAD	AHSLWYNFTI	IHLPRHGQQW	CEVQSQVDQK
	70	80	90	100	110	120
	NFLSYDCGSD	KVLSMGHLEE	QLYATDAWGK	QLEMLREVGO	RLRLELADTE	LEDFTPSGPL
	130	140	150	160	170	180

	TLQVRMSCEC	EADGYIRGSW	QFSFDGRKFL	LFDSNNRKWT	VVHAGARRMK	EKWEKDSGLT
	190	200	210	220	230	240
	TFFKMVSMRD	CKSWLRDFLM	HRKKRLEPTA	PPTMAPGLAQ	PKAIATTLSL	WSFLIILCFI
	LPGI					
Q5T2D3	10	20	30	40	50	60
SEQ ID NO:19	MSRKQAAKSR	PGSGSRKAEA	ERKRDERAAR	RALAKERRNR	PESGGGGGCE	EEFVSFANQL
	70	80	90	100	110	120
	QALGLKLREV	PGDGNCLFRA	LGDQLEHGR	NHLKHRQETV	DYMIKQREDF	EPFVEDDIPF
	130	140	150	160	170	180
	EKHVASLAKP	GTFAGNDAIV	AFARNHQLNV	VIHQLNAPLW	QIRGTEKSSV	RELHIAYRYG
	190	200	210	220	230	240
	EHYDSVRRIN	DNSEAPAHQ	TDFQMLHQDE	SNKREKIKTK	GMDSEDDLRL	EVEDAVQKVC
	250	260	270	280	290	300
	NATGCSDFNL	IVQNLEAENY	NIESAIIAVL	RMNQGKRNN	EENLEPSGRV	LKQCGPLWEE
	310	320	330	340	350	360
	GGSGARIFGN	QGLNEGRTEN	NKAQASPSEE	NKANKNQLAK	VTNKQRREQQ	WMEKKKRQEE
	370	380	390			
	RHRHKALESR	GSHRDNNRSE	AEANTQVTLV	KTFAALNI		
Q8IXT5	10	20	30	40	50	60
SEQ ID NO:20	MAVVIRLLGL	PFIAGPVDIR	HFFTGLTIPD	GGVHIIGGEI	GEAFIIFATD	EDARRAISRS
	70	80	90	100	110	120
	GGFIKDSSVE	LFLSSKAEMQ	KTIEMKRTDR	VGRGRPGSGT	SGVDSLNSFI	ESVKEEASNS
	130	140	150	160	170	180
	GYGSSINQDA	GFHTNGTGHG	NLRPRKTRPL	KAENPYLFLR	GLPYLVNEDD	VRVFFSGLCV
	190	200	210	220	230	240
	DGVIFLKHHD	GRNNGDAIVK	FASCV DASGG	LKCHRSFMGS	RFIEVMQSGE	QQWIEFGGNA
	250	260	270	280	290	300
	VKEGDVLRRS	EEHSPPRGIN	DRHFRKRSHS	KSPRRTRRSR	PLGFYVHLKN	LSLSIDERDL
	310	320	330	340	350	360
	RNFFRGTDLT	DEQIRFLYKD	ENRTRYAFVM	FKTLKDYN TA	LSLHKTVLQY	RPVHIDPISR
	370	380	390	400	410	420
	KQMLKFIARY	EKKRSGSLER	DRPGHVSQKY	SQEGNSGQKL	CIYIRNFPFD	VTKVEVQKFF
	430	440	450	460	470	480
	ADFLLAEDDI	YLLYDDKGVG	LGEALVKFKS	EEQAMKAERL	NRRRFLGTVE	LLRLISEAQI
	490	500	510	520	530	540
	QEFGVNF SVM	SSEKMQARSQ	SRERGDHSHL	FDSKDPPIYS	VGAFENFRHQ	LEDLRQLDNF
	550	560	570	580	590	600
	KHPQRDFRQP	DRHPPEDFRH	SSEDFRFPPE	DFRHSPEDFR	RPREEDFRFP	SEEDFRFPWE
	610	620	630	640	650	660
	EDFRPPEDD	FRHPREEDWR	RPLEEDWRRP	LEEDFRRSPT	EDFRQLPEED	FRQPPEEDLR
	670	680	690	700	710	720
	WLPEEDFRFP	PEEDWRRPPE	EDFRRLQGE	WRRPPEDDFR	RPPEEDFRHS	PEEDFRQSPQ
	730	740	750	760	770	780
	EHFRPPQEH	FRRPPPEHFR	RPPPEHFRRP	PPEHFRRPPP	EHFRPPPEH	FRRPPPEHFR
	790	800	810	820	830	840
	RPPQEHFRFP	PQEHFRRSRE	EDFRHPPDED	FRGPPDEDFR	HPPDEDFRSP	QEEDFRCPSD
	850	860	870	880	890	900
	EDFRQLPEED	LREAPEEDPR	LPDNFRPPGE	DFRSPPDDFR	SHRPFVNFGR	PEGGKFDGFK
	910	920	930	940	950	960
	HNMGSFP EGR	FMPDPKINCG	SGRVTPIKIM	NLPFKANVNE	ILDFFHGYRI	IPDSVSIQYN
	970	980	990	1000		
	EQGLPTGEAI	VAMINYNEAM	AAIKDLNDRP	VGPRKVKLTL	L	
Q9P225	10	20	30	40	50	60
SEQ ID NO:21	MSSKAEEKQR	LSGRGSSQAS	WSGRATRAAV	ATQEQGNAPA	VSEPELQAEL	PKEEPEPRLE
	70	80	90	100	110	120
	GPQAQSEESV	EPEADV KPLF	LSRAALTGLA	DAVWTQEHDA	ILEHFAQDPT	ESILTIFIDP
	130	140	150	160	170	180
	CFGLKLELGM	PVQTQNQLVY	FIRQAPVPIT	WENFEATVQF	GTVRGPYIPA	LLRLLGGVFA
	190	200	210	220	230	240
	PQIFANTGWP	ESIRNH FASH	LHKFLACLT	TRYKLEGHTV	LYIPAEAMNM	KPEMVIKDKE
	250	260	270	280	290	300
	LVQRLETSMI	HWTRQIK EML	SAQETVETGE	NLGPLEEIEF	WRNRCMDLSG	ISKQLVKKGV
	310	320	330	340	350	360
	KHVESILHLA	KSSYLAPFMK	LAQQIQD GSR	QAQSNLTFLS	ILKEPYQELA	FMKPKDISSK

370	380	390	400	410	420
LPKLISLIRI	IWVNSPHYNT	RERLTSLFRK	VCDCQYHFAR	WEDGKQGPLP	CFFGAQGPQI
430	440	450	460	470	480
TRNLLEIEDI	FHKNLHTLRA	VRGGILDVKN	TCWHEDYNKF	RAGIKDLEVM	TQNLITSAFE
490	500	510	520	530	540
LVRDVPHGVL	LLDTFHRLAS	REAIKRTYDK	KAVDLYMLFN	SELALVNRER	NKKWPDLEPY
550	560	570	580	590	600
VAQYSGKARW	VHILRRRIDR	VMTCLAGAHF	LPRIGTGKES	VHTYQQMVQA	IDELVRKTFQ
610	620	630	640	650	660
EWTSLLDKDC	IRRLDTPLLR	ISQEKAGMLD	VNFDKSLIL	FAEIDYWERL	LFETPHYVVN
670	680	690	700	710	720
VAERAEDLRI	LRENLLVAR	DYNRIIAMLS	PDEQALFKER	IRLLDKKIHP	GLKKLHWALK
730	740	750	760	770	780
GASAFFITEC	RIHASKVQMI	VNEFKASTLT	IGWRAQEMSE	KLLVRISGKR	VYRDLEFEED
790	800	810	820	830	840
QREHRAAVQQ	KLMNLHQDVV	TIMTNSYEVF	KNDGPEIQQQ	WMLYMIRLDR	MMEDALRLNV
850	860	870	880	890	900
KWSLLELSKA	INGDGKTSPL	PLFQVLVILK	NDLQGSVAQV	EFSPTLQTLA	GVVNDIGNHL
910	920	930	940	950	960
FSTISVFCFL	PDILTKRKLH	REPIQTVVEQ	DEDIKKIQTQ	ISSGMTNNAS	LLQNYLKTWD
970	980	990	1000	1010	1020
MYREIWEINK	DSFIHRYQRL	NPPVSSFVAD	IARYTEVANN	VQKEETVTNI	QFVLLDCSHL
1030	1040	1050	1060	1070	1080
KFSLVQHCNE	WQNKFATLLR	EMAAGRLEEL	HTYLKENAEK	ISRPPQTLLE	LGVSLQLVDA
1090	1100	1110	1120	1130	1140
LKHDLANVET	QIPPIHEQFA	ILEKYEVPVE	DSVLEMLDSL	NGEWWVFQQT	LLDSKQMLKK
1150	1160	1170	1180	1190	1200
HKEKFKTGLI	HSADDFKKKA	HTLLEDFEFK	GHFTSNVGYM	SALDQITQVR	AMLAMAREEE
1210	1220	1230	1240	1250	1260
NSLRANLGIF	KIEQPPSKDL	QNLEKELDAL	QQIWEIARDW	EENWNEWKTG	RFLILQTETM
1270	1280	1290	1300	1310	1320
ETTAHGLFRR	LTKLAKEYKD	RNWEIIETTR	SKIEQFKRTM	PLISDLRNP	LRERHWDQVR
1330	1340	1350	1360	1370	1380
DEIQREFDQE	SESFTLEQIV	ELGMDQHVEK	IGEISASATK	ELAIEVALQN	IAKTWDVTQL
1390	1400	1410	1420	1430	1440
DIVPYKDKGH	HRLRGTEEVF	QALDNQVAL	STMKASRFVK	AFEKDVDHWE	RCLSLILEVI
1450	1460	1470	1480	1490	1500
EMILTVQRQW	MYLENIFLGE	DIRKQLPNES	TLFDQVNSNW	KAIMDRMNKD	NNALRSTHHP
1510	1520	1530	1540	1550	1560
GLLDTLIEMN	TILEDIQKSL	DMYLETKRHI	FPRFYFLSND	DLLEILGQSR	NPEAVQPHLK
1570	1580	1590	1600	1610	1620
KCFDNIKLLR	IQKVGGPSSK	WEAVGMFSGD	GEYIDFLHSV	FLEGPVESWL	GDVEQTMVRT
1630	1640	1650	1660	1670	1680
LRDLLRNCHL	ALRKFLNKRD	KWVKEWAGQV	VITASQIQWT	ADVTKCLLTA	KERADKKILK
1690	1700	1710	1720	1730	1740
VMKKNQVSIL	NKYSEAIRGN	LTKIMRLKIV	ALVTIEIHAR	DVLEKLYKSG	LMDVNSFDWL
1750	1760	1770	1780	1790	1800
SQLRFYWEKD	LDDCVIRQTN	TQFQYNYEYL	GNSGRLVITP	LTDRCYMTLT	TALHLHRGGS
1810	1820	1830	1840	1850	1860
PKGPAGTGKT	ETVKDLGKAL	GIYVIVVNCS	EGLDYKSMGR	MYSGLAQTGA	WGCDEFNRI
1870	1880	1890	1900	1910	1920
NIEVLSVVAH	QILCILSALA	AGLTHFHFDG	FEINLVWSCG	IFITMNPGYA	GRTELPENLK
1930	1940	1950	1960	1970	1980
SMFRPIAMVV	PDSTLIAEII	LFGEFGNCK	ILAKKVYTTY	SLAVQQLSRQ	DHYDFGLRAL
1990	2000	2010	2020	2030	2040
TSLRLYAGKK	RRLQPDLTDE	EVLLLSMRDM	NIAKLTSVDA	PLFNAIVQDL	FPNIELPVID
2050	2060	2070	2080	2090	2100
YGKLRETVEQ	EIRDMGLQST	PFTLTQVFQ	YETKNSRHST	MIVGCTGSGK	TASWRILQAS
2110	2120	2130	2140	2150	2160
LSSLCRAGDP	NFNIVREFPL	NPKALS LGEL	YGEYDLSTNE	WTDGILSSVM	RTACADEKPD
2170	2180	2190	2200	2210	2220
EKWILFDGPV	DTLWIENMNS	VMDDNKVLT	INGERIAMPE	QVSLLEFEED	LAMASPATVS
2230	2240	2250	2260	2270	2280
RCGMVYTDYA	DLGWKPYVQS	WLEKRPKAEV	EPLQRMFEKL	INKMLAFKKD	NCKELVPLPE
2290	2300	2310	2320	2330	2340
YSGITSLCKL	YSALATPENG	VNPADGENYV	TMVEMTFVFS	MIWSVCASVD	EEGRKRIDSY
2350	2360	2370	2380	2390	2400
LREIEGSFPN	KDTVYEFVFD	PKIRSWTSFE	DKLPKSWRYP	PNAPFYKIMV	PTVDTVRYNY

2410	2420	2430	2440	2450	2460
LVSSLVANQN	PILLVGPVGT	GKTSIAQSVL	QSLPSSQWSV	LVVNMSAQTT	SNNVQSIIES
2470	2480	2490	2500	2510	2520
RVEKRTKGVY	VPFGGKSMIT	FMDDLNMMPAK	DMFGSQPPLE	LIRLWIDYGF	WYDRTKQTIK
2530	2540	2550	2560	2570	2580
YIREMFLMAA	MGPPGGGRTV	ISPRLSRFN	IINMTFPTKS	QIIRIFGTMI	NQKLQDFEEE
2590	2600	2610	2620	2630	2640
VKPIGNVVTE	ATLDMYNTVV	QRFLPTPTKM	HYLFNLRDIS	KVFQGMRLAN	KDFHDTKSSI
2650	2660	2670	2680	2690	2700
TRLWIHECFR	VFSDRLVDAA	DTEAFMGIIS	DKLGSFFDLT	FHHLCPSKRP	PIFGDFLKEP
2710	2720	2730	2740	2750	2760
KVYEDLTDLT	VLKTVMETAL	NEYNLSPSVV	PMQLVLFREA	IEHITRIVRV	IGQPRGNMLL
2770	2780	2790	2800	2810	2820
VGIGGSGRQS	LARLASSICD	YTTFQIEVTK	HYRKQEFRDD	IKRLYRQAGV	ELKTTSFIFV
2830	2840	2850	2860	2870	2880
DTQIADESFL	EDINNILSSG	EVPNLYKPDE	FEEIQSHIID	QARVEQVPES	SDSLFAYLIE
2890	2900	2910	2920	2930	2940
RVQNNLHIVL	CLSPMGDPFR	NWIRQYPALV	NCTTINWFSE	WPQEALLEVA	EKCLIGVDLG
2950	2960	2970	2980	2990	3000
TQENIHRKVA	QIFVTMHSV	AQYSQKMLLE	LRRHNYVTPT	KYLELLSGYK	KLLGEKRQEL
3010	3020	3030	3040	3050	3060
LAQANKLRTG	LFKIDETREK	VQVMSLELED	AKKKVAEFQK	QCEEYLVIIV	QQKREADEQQ
3070	3080	3090	3100	3110	3120
KAVTANSEKI	AVEEIKCQAL	ADNAQKDLEE	ALPALEEAMR	ALESLNKKDI	GEIKSYGRPP
3130	3140	3150	3160	3170	3180
AQVEIVMQAV	MILRGNEPTW	AEAKRQLGEQ	NFIKSLINFD	KDNISDKVLK	KIGAYCAQPD
3190	3200	3210	3220	3230	3240
FQPDIIIGRVS	LAAKSLCMWV	RAMELYGRLY	RVVEPKRIRM	NAALAQLEK	QAALAEAEQK
3250	3260	3270	3280	3290	3300
LREVAEKLEM	LKKQYDEKLA	QKEELRKKSE	EMELKLERAG	MLVSGLAGEK	ARWEETVQGL
3310	3320	3330	3340	3350	3360
EEDLGYLVGD	CLLAAAFLSY	MGPFLTNYRD	EIVNQIWIWK	IWELQVPCSP	SFAIDNFLCN
3370	3380	3390	3400	3410	3420
PTKVRDWNQ	GLPSDAFSTE	NGIIVTRGNR	WALMIDPQAA	ALKWIKNMEG	GQGLKIIDLQ
3430	3440	3450	3460	3470	3480
MSDYLRILEH	AIHFGYPVLL	QNVQEYLDPT	LNPMLNKSV	RIGGRLLMRI	GDKEVEYNTN
3490	3500	3510	3520	3530	3540
FRFYITTKLS	NPHYSPETSA	KTTIVNFAVK	EQGLEAQLLG	IVVRKERPEL	EEQKDSLVIN
3550	3560	3570	3580	3590	3600
IAAGKRKLKE	LEDEILRLLN	EATGSLDDV	QLVNTLHTSK	ITATEVTEQL	ETSETTEINT
3610	3620	3630	3640	3650	3660
DLAREAYRPC	AQRASILFFV	LNDMGCIDPM	YQFSLDAYIS	LFILSIDKSH	RSNKLEDRID
3670	3680	3690	3700	3710	3720
YLNDYHTYAV	YRYTCRTLFE	RHKLLFSFHM	CAKILETSBK	LNMDEYNFFL	RGGVVLDREG
3730	3740	3750	3760	3770	3780
QMDNPCSSWL	ADAYWDNITE	LDKLTNFHGL	MNSFEQYPRD	WHLWYTNAAP	EKAMLPGEWE
3790	3800	3810	3820	3830	3840
NACNEMQRM	IVRSLRQDRV	AFCVTSFIIT	NLGSRFIEPP	VLNMKSVLED	STPRSPLVFI
3850	3860	3870	3880	3890	3900
LSPGVDPTSA	LLQLAEHMG	AQRFHALSLG	QGQAPIAARL	LREGVTQGHW	VFLANCHLSL
3910	3920	3930	3940	3950	3960
SWMPNLDKLV	EQQLQVEDPHP	SFRLWLSSIP	HPDFPISILQ	VSIKMTTEPP	KGLKANMTRL
3970	3980	3990	4000	4010	4020
YQLMSEPQFS	RCSKPAKYKK	LLFSLCFFHS	VLLERKKFLQ	LGWNIIYGFN	DSDFEVSENL
4030	4040	4050	4060	4070	4080
LSLYLDEYEE	TPWDALKYLI	AGINYGGHVT	DDWDRRLTT	YINDYFCDQS	LSTPFHRLSA
4090	4100	4110	4120	4130	4140
LETYFIPKDG	SLASYKEYIS	LLPGMDPPEA	FGQHPNADVA	SQITEAQTLF	DTLLSLQPQI
4150	4160	4170	4180	4190	4200
TPTRAGGQTR	EEKVLELAAD	VKQKIPEMID	YEGTQKLLAL	DPSPLNVVLL	QEIQRYNTLM
4210	4220	4230	4240	4250	4260
QTILFSLTDL	EKGIQGLIVM	STSLEEIFNC	IFDAHVPPLW	GKAYPSQKPL	AAWTRDLAMR
4270	4280	4290	4300	4310	4320
VEQFELWASR	ARPPVIFWLS	GFTFPTGFLT	AVLQSSARQN	NVSVDLSWE	FIVSTVDDSN
4330	4340	4350	4360	4370	4380
LVYPPKDGWV	VRGLYLEGAG	WDRKNSCLVE	AEPMLVLCLM	PTIHFRPAES	RKKSAGMYS
4390	4400	4410	4420		
CPCYYYPNRA	GSSDRASFI	GIDLRSGAMT	PDHWIKRGTA	LLMSLDS	

Q9Y2I9 SEQ ID NO:22	10	20	30	40	50	60
	MDVLPTGGGR	PGLRTELEFR	GGGGEARLES	QEEETIPAAP	PAPRLRGAAE	RPRRSRDTWD
	70	80	90	100	110	120
	GDEDEPGEA	CGGRTSRTAS	LVSGLLNELY	SCTEEEEAAG	GGRGAEGRRR	RRDSLDSSTE
	130	140	150	160	170	180
	ASGSDVVLGG	RSGAGDSRVL	QELQERPSQR	HQMLYLRQKD	ANELKILRE	LKYRIGIQSA
	190	200	210	220	230	240
	KLLRHLKQKD	RLHKKVQRNC	DIVTACLQAV	SQKRRVDTKL	KFTLEPSLGQ	NGFQQWYDAL
	250	260	270	280	290	300
	KAVARLSTGI	PKEWRRKVWL	TLADHYLHSI	AIDWDKTMRF	TFNERSNPDD	DSMGIQIVKD
	310	320	330	340	350	360
	LHRTGCSSYC	GQEAQDRVV	LKRVLLAYAR	WNKTVGYCQG	FNILAALILE	VMEGNEGDAL
	370	380	390	400	410	420
	KIMIYLIDKV	LPESYFVNNL	RALSVDMAVF	RDLLRMKLPE	LSQHLDLQOR	TANKESGGGY
	430	440	450	460	470	480
	EPPLTNVFTM	QWFLTLFATC	LPNQTVLKIW	DSVFFEGSEI	ILRVSLAIWA	KLGEQIECCE
	490	500	510	520	530	540
	TADEFYSTMG	RLTQEMLEND	LLQSHELMQT	VYSMAFPFPF	QLAELREKYT	YNITPFPATV
	550	560	570	580	590	600
	KPTSVSGRHS	KARDSDEEND	PDDEDAVVNA	VGCLGPFSGF	LAPELQKYQK	QIKEPNEEQS
	610	620	630	640	650	660
	LRSNNIAELS	PGAINSCRSE	YHAAFNSMMM	ERMTTDINAL	KRQYSRIKKK	QQQQVHQVYI
	670	680	690	700	710	720
	RADKGPVTISI	LPSQVNSSPV	INHLLLGKKM	KMTNRAAKNA	VIHIPGHTGG	KISPVVPYEDL
	730	740	750	760	770	780
	KTKLNSPWRT	HIRVHKKNMP	RTKSHPGCGD	TVGLIDEQNE	ASKTNGLGAA	EAFPSGCTAT
	790	800	810	820	830	840
	AGREGSSPEG	STRRTIEGQS	PEPVFGDADV	DVSAVQAKLG	ALELNQRDAA	AETELRVHPP
	850	860	870	880	890	900
	CQRHCPEPPS	APEENKATSK	APQGSNSKTP	IFSPFPSVKP	LRKSATARNL	GLYGPTERTP
	910	920				
	TVHFPQMSRS	FSKPGGGNSG	TKKR			

Table 2.

P41222 (PTGDS) SEQ ID NO:23	10	20	30	40	50	60
	MATHHTLWMG	LALLGLVGLDL	QAAPEAQVSV	QPNFQQDKFL	GRWFSAGLAS	NSSWLREKKA
	70	80	90	100	110	120
	ALSMCKSVVA	PATDGGLNLT	STFLRKNQCE	TRTMLLQPAG	SLGSSYSRSP	HWGSTYSVSV
	130	140	150	160	170	180
	VETDQYDQYAL	LYSQGSKGPG	EDFRMATLYS	RTQTPRAELK	EKFTAFCKAQ	GFTEDTIVFL
	190					
	PQTDKCMTEQ					
P14151 (SELL) SEQ ID NO:24	10	20	30	40	50	60
	MIFPWKCQST	QRDLWNIFKL	WGWTMLCCDF	LAHHGTDCWT	YHYSEKPMNW	QRARRFCRDN
	70	80	90	100	110	120
	YTDLVAIQNK	AEIEYLEKTL	PFSRSYYWIG	IRKIGGIWTW	VGTKNSLTEE	AENWGDGEPN
	130	140	150	160	170	180
	NKKNKEDCVE	IYIKRNKDAG	KWNDDACHKL	KAALCYTASC	QPWSCSGHGE	CVEIINNYTC
	190	200	210	220	230	240
	NCDVGGYGPQ	CQFVIQCEPL	EAPELGTMDC	THPLGNFSFS	SQCAFSCSEG	TNLTGIEETT
	250	260	270	280	290	300
	CGPFGNWSSP	EPTCQVIQCE	PLSAPDLGIM	NCSHPLASFS	FTSACTFICS	EGTELIGKKK
	310	320	330	340	350	360
	TICSSSGIWS	NPSPICQKLD	KSFMSMIKEG	YNPLFIPVAV	MVTAFSGLAF	IIWLARRLKK
	370					
	GKKSKRSMND	PY				
Q06418 (TYRO3) SEQ ID NO:25	10	20	30	40	50	60
	TVEGTRANLT	GWDPQKDLIV	RVCVSNAVGC	GPWSQPLVVS	SHDRAGQQGP	PHSRTSWVPV
	70	80	90	100	110	120
	VLGVLTALVT	AAALALILLR	KRRKETRFQ	AFDSVMARGE	PAVHFRAARS	FNRRERPERIE
	130	140	150	160	170	180
	ATLDSLGISD	ELKEKLEDVL	IPEQQFTLGR	MLGKGEFGSV	REAQLKQEDG	SFVKVAVKML
	190	200	210	220	230	240
	KADIASSDI	EEFLREAACM	KEFDHPHVAK	LVGVSLRSRA	KGRLPIMVI	LPFMKHGDHL
	250	260	270	280	290	300
	AFLASRIGE	NPFNLPLQTL	IRFMVDIACG	MEYLSSRNFI	HRDLAARNCM	LAEDMTVCVA
	310	320	330	340	350	360
	DFGLSRKIYS	GDYYRQGCAS	KLPVKWLAL	SLADNLYTVQ	SDVWAFGVMT	WEIMTRGQTP
	370	380	390	400	410	420
	YAGIENAEIY	NYLIGGNRLK	QPPECMEDVY	DLMYQCWSAD	PKQRPSFTCL	RMELENILGQ
	430	440	450	460	470	480
	LSVLSASQDP	LYINIERAEE	PTAGGSLELP	GRDQPYSGAG	DGSGMGAVGG	TPSDCRYILT
	490	500	510			
	PGGLAEQPGQ	AEHQPEspln	ETQRLLLLLQ	GLLPHSSC		
P52306 (RAPIGDS1) SEQ ID NO:26	10	20	30	40	50	60
	MDNLSDTLKK	LKITAVDKTE	DSLEGCLDCL	LQALAQNNTE	TSEKIQASGI	LQLFASLLTP
	70	80	90	100	110	120
	QSSCKAKVAN	IIAEVAKNEF	MRIPCVDAGL	ISPLVQLLNS	KDQEVLLQTG	RALGNICYDS
	130	140	150	160	170	180
	HEGRSAVDQA	GGAQIVIDHL	RSLCSITDPA	NEKLLTVFCG	MLMNYSNEND	SLQAQLINMG
	190	200	210	220	230	240
	VIPTLVKLLG	IHCQNAALTE	MCLVAFGNLA	ELESSKEQFA	STNIAEELVK	LFKKQIEHDK
	250	260	270	280	290	300
	REMIFEVLAP	LAENDAIKLQ	LVEAGLVECL	LEIVQQKVDS	DKEDDITELK	TGSDLMVLLL
	310	320	330	340	350	360
	LGDESMQKLF	EGGKGSVFQR	VLSWIPSNNH	QLQLAGALAI	ANFARNANC	IHMVDNGIVE
	370	380	390	400	410	420
	KLMDLLDRHV	EDGNVTVQHA	ALSALRNLA	PVINKAKMLS	AGVTEAVLKF	LKSEMPVQF
	430	440	450	460	470	480
	KLLGTLRMLI	DAQAEAAEQ	GKNVKLVERL	VEWCEAKDHA	GVMGESNRL	SALIRHSSK
	490	500	510	520	530	540
	DVIKTIVQSG	GIKHLVTMAT	SEHVIMQNEA	LVALALIAAL	ELGTAEKDLE	SAKLVIQLHR
	550	560	570	580	590	600
	LLADERSAPE	IKYNSMVLIC	ALMGSECLHK	EVQDLAFLDV	VSKLRSHENK	SVAQQASLTE

	QRLTVES					
Q9Y5Y7 (LYVE1) SEQ ID NO:27	10	20	30	40	50	60
	MARCFSLVLL	LTSIWTTTRL	VQGSRAEEL	SIQVSCRIMG	ITLVSKKANQ	QLNFTEAKEA
	70	80	90	100	110	120
	CRLGLSLAG	KDQVETALKA	SFETCSYGWV	GDGFVVISRI	SPNPKCGKNG	VGVLWKVPV
	130	140	150	160	170	180
	SRQFAAYCYN	SSDTWTNSCI	PEIITTKDPI	FNTQTATQTT	EFIVSDSTYS	VASPYSTIPA
	190	200	210	220	230	240
	PTTTPPAPAS	TSIPRRKKLI	CVTEVFMETS	TMSTETEPFV	ENKAAFKNEA	AGFGGVPTAL
	250	260	270	280	290	300
	LVLALLFFGA	AAGLGFCYVK	RYVKAFPFTN	KNQOKEMIET	KVVKEEKAND	SNPNEESKKT
	310	320				
	DKNPEESKSP	SKTTVRCLEA	EV			

CLAIMS

What is claimed is:

1. A method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient, the method comprising the steps of:
analyzing a test sample comprising urine from the patient for the presence or concentration of at least one protein, whereby a test data set is obtained; and,
comparing the test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample;
whereby the HIV viral load in the patient is assessed or monitored.
2. The method of claim 1, wherein the patient has received or is receiving a first anti-HIV medication.
3. The method of claim 1, wherein the patient is a new-born human or an infant younger than about 18 months of age.
4. The method of claim 1, wherein the test sample is prepared by a method comprising subjecting urine from the patient to at least one procedure selected from the group consisting of protein isolation and protein digestion.
5. The method of claim 1, wherein the test sample is analyzed using mass spectrometry, a quantum dot assay or a chromophore assay.
6. The method of claim 1, wherein the test sample is analyzed using a method comprising contacting the test sample with an antibody or aptamer.
7. The method of claim 6, wherein the antibody is at least one selected from the group consisting of a polyclonal antibody, monoclonal antibody, Fv, Fab, F(ab)₂, single chain antibody, human antibody, humanized antibody, and fragments and derivatives thereof.
8. The method of claim 6, wherein the antibody or aptamer is used in an immunoassay.

9. The method of claim 8, wherein the immunoassay comprises at least one selected from the group consisting of immunoturbidimetry, immunonephelometry, ELISA assay, radioimmunoassay, chemiluminescence immunoassay, immunofluorescence, immunoprecipitation, immunoelectrophoresis, and flow cytometry-based immunoassay.

10. The method of claim 1, wherein the control sample comprises an urine sample from an untreated HIV-infected control human.

11. The method of claim 10, wherein the untreated HIV-infected control human is the human patient before receiving anti-HIV medication.

12. The method of claim 1, wherein the control sample comprises an urine sample from an HIV-uninfected control human.

13. The method of claim 1, wherein the control sample comprises an urine sample from an HIV-infected control human with controlled infection.

14. The method of claim 1, wherein the concentration of the protein in the patient's urine is higher by at least a multiplicity factor than the concentration of the protein in the urine from an HIV-uninfected control human or from an HIV-infected control human with controlled infection, wherein the patient is identified as having uncontrolled HIV infection, whereby the patient is prescribed a second anti-HIV medication that is distinct from the first anti-HIV medication.

15. The method of claim 14, wherein the multiplicity factor is selected from the group consisting of about 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 30, 40, 50, 75, 100, 125, 250, 500 and 1,000.

16. The method of claim 1, wherein the concentration of the protein in the patient's urine is lower by at least a multiplicity factor than the concentration of the protein in the urine from an HIV-uninfected control human or from an HIV-infected control human with controlled infection, wherein the patient is identified as having controlled HIV infection, whereby the patient continues to be prescribed the first anti-HIV medication.

17. The method of claim 16, wherein the multiplicity factor is selected from the group consisting of about 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 30, 40, 50, 75, 100, 125, 250, 500 and 1,000.

18. The method of claim 1, wherein the concentration of the protein in the patient's urine is equal to or greater than a multiplicity factor of the concentration of the protein in the urine from an untreated HIV-positive control human, wherein the patient is identified as having uncontrolled HIV infection, whereby the patient is prescribed a second anti-HIV medication which is distinct from the first anti-HIV medication.

19. The method of claim 18, wherein the multiplicity factor is selected from the group consisting of about 1, 0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.025, 0.01, 0.005, 0.0025, 0.001, 0.0005, 0.00025, 0.0001, 0.00005 and 0.00001.

20. The method of claim 1, wherein the concentration of the protein in the patient's sample is lower than a multiplicity factor of the concentration of the protein in the urine from an untreated HIV-positive control human, wherein the patient is identified as having controlled HIV infection, whereby the patient continues to be prescribed the first anti-HIV medication.

21. The method of claim 20, wherein the multiplicity factor is selected from the group consisting of about 0.95, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.025, 0.01, 0.005, 0.0025, 0.001, 0.0005, 0.00025, 0.0001, 0.00005 and 0.00001.

22. The method of claim 1, wherein the at least one protein has an accession number selected from the group consisting of Q8TD57 (SEQ ID NO:1), Q18PE1 (SEQ ID NO:2), Q8NFH5 (SEQ ID NO:3), Q8WYL5 (SEQ ID NO:4), Q8IYD8 (SEQ ID NO:5), O14654 (SEQ ID NO:6), Q96AP4 (SEQ ID NO:7), Q9UQ35 (SEQ ID NO:8), Q8N6W0 (SEQ ID NO:9), Q9H792 (SEQ ID NO:10), Q9H497 (SEQ ID NO:11), Q9UE35 (SEQ ID NO:12), O00743 (SEQ ID NO:13), Q8WXF8 (SEQ ID NO:14), P81274 (SEQ ID NO:15), Q8NG08 (SEQ ID NO:16), Q96AE7 (SEQ ID NO:17), Q9BZM4 (SEQ ID NO:18), Q5T2D3 (SEQ ID NO:19), Q8IXT5 (SEQ ID NO:20), Q9P225 (SEQ ID NO:21), and Q9Y2I9 (SEQ ID NO:22).

23. The method of claim 1, wherein the at least one protein has an accession number selected from the group consisting of P41222 (PTGDS) (SEQ ID NO:23), P14151 (SELL) (SEQ ID NO:24), Q06418 (TYRO3) (SEQ ID NO:25), P52306 (RAP1GDS1) (SEQ ID NO:26), and Q9Y5Y7 (LYVE1) (SEQ ID NO:27).

24. A kit for assessing or monitoring systemic HIV viral load in an HIV-infected human patient, the kit comprising

an antibody or aptamer that binds to at least one protein with an accession number selected from the group consisting of Q8TD57 (SEQ ID NO:1), Q18PE1 (SEQ ID NO:2), Q8NFH5 (SEQ ID NO:3), Q8WYL5 (SEQ ID NO:4), Q8IYD8 (SEQ ID NO:5), O14654 (SEQ ID NO:6), Q96AP4 (SEQ ID NO:7), Q9UQ35 (SEQ ID NO:8), Q8N6W0 (SEQ ID NO:9), Q9H792 (SEQ ID NO:10), Q9H497 (SEQ ID NO:11), Q9UE35 (SEQ ID NO:12), O00743 (SEQ ID NO:13), Q8WXF8 (SEQ ID NO:14), P81274 (SEQ ID NO:15), Q8NG08 (SEQ ID NO:16), Q96AE7 (SEQ ID NO:17), Q9BZM4 (SEQ ID NO:18), Q5T2D3 (SEQ ID NO:19), Q8IXT5 (SEQ ID NO:20), Q9P225 (SEQ ID NO:21), and Q9Y2I9 (SEQ ID NO:22);

an applicator; and,

an instructional material for the use of the kit, wherein the instruction material comprises instructions for analyzing a test sample comprising urine from the patient for the presence or concentration of the at least one protein.

25. The kit of claim 24, further comprising a test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample.

26. The kit of claim 25, wherein the control sample comprises an urine sample from an untreated HIV-infected control human.

27. The kit of claim 25, wherein the control sample comprises an urine sample from an HIV-uninfected control human.

28. The kit of claim 25, wherein the control sample comprises an urine sample from an HIV-infected control human with controlled infection.

29. A kit for assessing or monitoring systemic HIV viral load in an HIV-infected human patient, the kit comprising

an antibody or aptamer that binds to at least one protein with an accession number selected from the group consisting of P41222 (PTGDS) (SEQ ID NO:23), P14151 (SELL) (SEQ ID NO:24), Q06418 (TYRO3) (SEQ ID NO:25), P52306 (RAP1GDS1) (SEQ ID NO:26), and Q9Y5Y7 (LYVE1) (SEQ ID NO:27);

an applicator; and,

an instructional material for the use of the kit, wherein the instruction material comprises instructions for analyzing a test sample comprising urine from the patient for the presence or concentration of the at least one protein.

30. The kit of claim 29, further comprising a test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample.

31. The kit of claim 30, wherein the control sample comprises an urine sample from an untreated HIV-infected control human.

32. The kit of claim 30, wherein the control sample comprises an urine sample from an HIV-uninfected control human.

33. The kit of claim 30, wherein the control sample comprises an urine sample from an HIV-infected control human with controlled infection.

Fig. 1

Characteristic	Value
Age at enrollment, y	36.7 (range: 23–60)
Sex	11 men, 8 women (including 2 male-to-female transgendered individuals)
Race	African American 18, white 0, Hispanic 1, other 0
HIV-1 viral load, copies/mL	137,164 (range: 54,670–395,060)
CD4 count, cells/ μ L	325 (range: 10–624)

Fig. 2A

No. of HIV Samples in Which Protein Was Found (out of 19)	Protein Accession No. (Uniprot)	Protein Name and Gene Ontology
14	Q8TD57	Dynein heavy-chain 3, axonemal (DNAH3) (ATPase force-generating protein in microtubule-based movement)
14	Q18PE1	Docking protein 7 (DOK7) (roles in neuromuscular synaptogenesis)
12	Q8NFH5	Nucleoporin 35-kD (NUP35) (component of nuclear pore complex regulating nuclear transport)
12	Q8WYL5	Slingshot protein phosphatase 1 (SSH1) (protein phosphatase that regulates actin filament dynamics in cell projection and locomotion)
8	Q8IYD8	Fanconi anemia, complementation group M (FANCM) (ATP-dependent helicase involved in DNA damage repair)
8	O14654	Insulin receptor substrate 4 (IRS4) (tyrosine kinase growth factor receptor binding signal transduction protein)
6	Q96AP4	Zinc finger with UFM1-specific peptidase domain (ZUFSP) (metal ion binding peptidase)
6	Q9UQ35	Serine/arginine repetitive matrix 2 (SRRM2) (RNA-binding component of the active spliceosome in pre-mRNA splicing)
6	Q8N6W0	CUGBP, Elav-like family member 5 (CELF5) (RNA-binding protein implicated in the regulation of pre-mRNA alternative splicing)
5	Q9H792	Pseudopodium-enriched atypical kinase 1 (PEAK1) (non-membrane spanning protein tyrosine kinase activity in cell spreading and migration)
5	Q9H497	Torsin family 3, member A (TOR3A) (ATP binding, protein folding cofactor)
5	Q9UE35	Estrogen receptor (ESR1) (nuclear hormone receptor family)

Fig. 2B

No. of HIV Samples in Which Protein Was Found (out of 19)	Protein Accession No. (Uniprot)	Protein Name and Gene Ontology
5	O00743	Protein phosphatase 6, catalytic subunit (PPP6C) (catalytic subunit of protein phosphatase 6 involved in cell cycle progression)
5	Q8WXF8	Death effector domain containing 2 (DEDD2) (death receptor-induced apoptosis)
4	P81274	G-protein signaling modulator 2 (GPSM2) (modulates activation of G proteins, which transduce extracellular signals received by cell surface receptors into integrated cellular responses)
4	Q8NG08	Helicase (DNA) B (HELB) (catalyzes the unwinding of DNA necessary for DNA replication, repair, recombination, and transcription)
4	Q96AE7	Tetratricopeptide repeat domain 17 (TTC17) (role in primary ciliogenesis by modulating actin polymerization)
4	Q9BZM4	UL16 binding protein 3 (ULBP3) (ligand for the KLRK1/NKG2D receptor in NK cells involved in production of cytokines and chemokines and antigen processing and presentation)
4	Q5T2D3	OTU domain containing 3 (OTUD3) (ubiquitin-specific protease activity in deubiquitination)
4	Q8IXT5	RNA-binding motif protein 12B (RBM12B) (RNA binding)
3	Q9P225	Dynein, axonemal, heavy chain 2 (DNAH2) (ATPase force-generating protein in microtubule-based movement)
3	Q9Y2I9	TBC1 domain family, member 30 (TBC1D30) (GTPase-activating protein with broad specificity for RAB family small G proteins)

Fig. 3A
(Part 1/2)

Gene Name	Protein Description	Accession number	Novel in HIV	Greatly increase in HIV urine
PGLYRP1	peptidoglycan recognition protein 1 [S	ENSP00000007752		
CD9	CD9 molecule [Source:HGNC Symbol;Acc	ENSP00000008938		
		ENSP00000027335		
NGFR	nerve growth factor receptor [Source	ENSP00000028008		
CTSA	cathepsin A [Source:HGNC Symbol;Acc	ENSP00000029717		
LAMP2	lysosomal-associated membrane prot	ENSP00000053243		
PTGDS	prostaglandin D2 synthase 21kDa (br	ENSP000000284240		yes
CETP	cholesteryl ester transfer protein, pla	ENSP00000053867		
CD209	CD209 molecule [Source:HGNC Symbi	ENSP00000058691		
EFNB1	ephrin-B1 [Source:HGNC Symbol;Acc	ENSP00000075322		
APOH	apolipoprotein H (beta-2-glycoprotein	ENSP00000078445		
MADCAM1	mucosal vascular addressin cell adhes	ENSP00000158302		
PIK3IP1	phosphoinositide-3-kinase interacting	ENSP00000162749		
ARSA	arylsulfatase A [Source:HGNC Symbol	ENSP00000170630		
COCH	coagulation factor C homolog, cochlin	ENSP00000172229		
PROCR	protein C receptor, endothelial [Sourc	ENSP00000178510		

Fig. 3A
(Part 2/2)

peptides								spectral counts							
HIV 1	HIV 2	HIV 3	HIV 4	HIV 5	HIV 6	HIV 7	HIV 8	HIV 1	HIV 2	HIV 3	HIV 4	HIV 5	HIV 6	HIV 7	HIV 8
1	3	1	0	2	1	1	0	1	9	2	0	3	2	4	0
1	4	2	3	6	4	6	2	1	11	14	7	78	12	105	6
2	0	0	1	1	1	1	1	6	0	0	2	2	5	2	6
2	0	0	2	0	0	1	0	2	0	0	3	0	0	1	0
2	0	4	2	1	1	2	3	4	0	14	4	2	2	4	11
0	5	4	1	3	5	7	0	0	11	18	1	14	29	29	0
2	0	2	3	1	3	2	3	9	0	7	6	2	9	4	31
14	17	18	12	13	15	17	16	771	316	557	219	852	625	620	315
0	0	0	0	4	1	4	0	0	0	0	0	24	6	36	0
0	0	1	2	0	0	5	0	0	0	9	15	0	0	23	0
0	1	0	0	2	0	2	1	0	2	0	0	5	0	3	1
5	4	6	11	10	8	11	7	18	42	14	101	101	60	167	20
2	0	0	0	1	2	2	0	3	0	0	0	13	6	6	0
6	1	7	6	7	7	6	7	99	1	85	53	67	63	52	87
2	1	6	0	3	1	5	1	5	5	51	0	12	3	29	6
0	0	1	2	1	0	1	0	0	0	1	5	1	0	1	0
5	1	2	2	4	4	3	2	30	3	3	4	27	22	13	8

Fig. 3B
(Part 1/2)

Gene Name	Protein Description	Accession number	Novel in HIV	Greatly increase in HIV urine
CTSZ	cathepsin Z [Source:HGNC Symbol;Acc:ENSP00000190983]	ENSP00000190983		
GLA	galactosidase, alpha [Source:HGNC Sy	ENSP00000196968		
OLFM4	olfactomedin 4 [Source:HGNC Symbo	ENSP00000199764		
NUTF2	nuclear transport factor 2 [Source:HG	ENSP00000200639		
CTSH	cathepsin H [Source:HGNC Symbol;Ac	ENSP00000203005		
CPQ	carboxypeptidase Q [Source:HGNC Sy	ENSP00000204005		
FCGBP	Fc fragment of IgG binding protein [Sc	ENSP00000204801		
PCOLCE	procollagen C-endopeptidase enhanc	ENSP00000204961		
SELL	selectin L [Source:HGNC Symbol;Acc:J	ENSP00000216953	yes	
		ENSP00000205948		
ACTA2	actin, alpha 2, smooth muscle, aorta [	ENSP00000211979		
		ENSP00000212355		
NAGLU	N-acetylglucosaminidase, alpha [Sour	ENSP00000215637		
VTN	vitronectin [Source:HGNC Symbol;Acc	ENSP00000215727		
AFM	afamin [Source:HGNC Symbol;Acc:31	ENSP00000215754		
CTSC	cathepsin C [Source:HGNC Symbol;Ac	ENSP00000215912		
HSPA8	heat shock 70kDa protein 8 [Source:H	ENSP00000216200		
GAPDH	glyceraldehyde-3-phosphate dehydro	ENSP00000216254		
LDHB	lactate dehydrogenase B [Source:HGN	ENSP00000216361		
		ENSP00000216492		

Fig. 3B
(Part 2/2)

peptides								spectral counts							
2	3	0	0	0	1	1	0	8	18	0	0	0	2	2	0
0	1	2	2	0	1	2	0	0	3	4	5	0	4	5	0
2	0	0	0	1	4	0	0	4	0	0	0	2	27	0	0
1	1	1	1	4	3	4	0	1	1	1	1	19	9	30	0
0	0	3	0	0	0	0	0	0	0	6	0	0	0	0	0
0	1	0	0	0	3	2	0	0	2	0	0	0	5	3	0
3	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0
0	1	2	0	2	1	2	0	0	2	4	0	13	2	12	0
2	1	0	0	2	0	1	0	3	2	0	0	2	0	1	0
2	0	2	0	2	0	0	1	2	0	7	0	2	0	0	2
2	0	0	0	2	4	2	1	2	0	0	0	3	5	2	2
3	1	1	3	3	3	3	2	28	2	2	18	54	25	34	18
4	3	3	1	7	17	14	5	5	5	7	2	39	134	132	13
5	3	1	2	2	4	3	1	31	24	23	10	25	41	35	4
14	2	1	1	0	0	1	0	48	10	2	1	0	0	3	0
0	4	0	0	0	1	0	0	0	13	0	0	0	1	0	0
0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	6
0	0	0	0	1	3	3	0	0	0	0	0	1	5	11	0
0	0	0	0	1	3	1	0	0	0	0	0	2	7	2	0
0	0	0	0	1	2	2	0	0	0	0	0	1	4	2	0

Fig. 3C
(Part 1/2)

Gene Name	Protein Description	Accession number	Greatly increase	
			Novel in HIV	in HIV urine
REG1A	regenerating islet-derived 1 alpha [Source:HGNC Symbol;Acc:6]	ENSP00000216649		
ENO1	enolase 1, (alpha) [Source:HGNC Symbol;Acc:6]	ENSP00000216879		
APOA1	apolipoprotein A-I [Source:HGNC Symbol;Acc:6]	ENSP00000216968		
TTR	transthyretin [Source:HGNC Symbol;Acc:6]	ENSP00000217182		
SPP1	secreted phosphoprotein 1 [Source:HGNC Symbol;Acc:6]	ENSP00000217402		
NPC2	Niemann-Pick disease, type C2 [Source:HGNC Symbol;Acc:6]	ENSP00000217421		
CNTFR	ciliary neurotrophic factor receptor [Source:HGNC Symbol;Acc:6]	ENSP00000217423		
SEMG1	semenogelin I [Source:HGNC Symbol;Acc:6]	ENSP00000217426		
TYRO3	TYRO3 protein tyrosine kinase [Source:HGNC Symbol;Acc:6]	ENSP00000233813	yes	
HIST1H2BG	histone cluster 1, H2bg [Source:HGNC Symbol;Acc:6]	ENSP00000218230		
C3	complement component 3 [Source:HGNC Symbol;Acc:6]	ENSP00000218388		
SLURP1	secreted LY6/PLAUR domain containing 1 [Source:HGNC Symbol;Acc:6]	ENSP00000219169		
IGFLR1	IGF-like family receptor 1 [Source:HGNC Symbol;Acc:6]	ENSP00000220166		
DNASE1	deoxyribonuclease I [Source:HGNC Symbol;Acc:6]	ENSP00000221264		
HSPB1	heat shock 27kDa protein 1 [Source:HGNC Symbol;Acc:6]	ENSP00000221515		
GGT1	gamma-glutamyltransferase 1 [Source:HGNC Symbol;Acc:6]	ENSP00000222219		
DSC2	desmocollin 2 [Source:HGNC Symbol;Acc:6]	ENSP00000222271		
KRT33B	keratin 33B [Source:HGNC Symbol;Acc:6]	ENSP00000222374		
KRT1	keratin 1 [Source:HGNC Symbol;Acc:6]	ENSP00000222690		
KRT6C	keratin 6C [Source:HGNC Symbol;Acc:6]	ENSP00000222725		

Fig. 3C
(Part 2/2)

peptides								spectral counts									
4	0	5	4	5	1	5	3	7	0	25	13	20	1	18	4		
0	0	0	0	2	5	2	1	0	0	0	0	6	20	7	2		
0	0	0	6	11	4	3	0	0	0	0	30	52	18	18	0		
0	0	0	0	1	1	0	0	0	0	0	0	1	3	0	0		
6	1	7	10	10	7	8	7	94	1	118	165	127	74	92	135		
0	1	1	1	1	1	2	2	0	2	7	4	6	4	11	8		
2	0	0	0	1	1	0	0	2	0	0	0	2	4	0	0		
0	2	0	0	0	0	0	0	0	9	0	0	0	0	0	0		
0	0	1	0	1	0	0	2	0	0	2	0	1	0	0	10		
0	0	0	0	0	3	0	0	0	0	0	0	0	3	0	0		
0	0	0	5	1	2	3	1	0	0	0	17	1	4	6	1		
1	2	3	2	3	3	2	1	2	16	42	14	24	19	25	14		
0	0	1	0	1	0	0	0	0	0	2	0	1	0	0	0		
5	0	2	3	4	5	4	2	30	0	6	22	42	43	34	16		
0	0	0	0	2	0	1	0	0	0	0	0	2	0	1	0		
0	0	0	0	2	2	0	0	0	0	0	0	2	3	0	0		
3	2	1	2	4	3	3	2	10	8	2	3	20	11	10	2		
2	0	0	2	0	0	0	0	7	0	0	3	0	0	0	0		
14	0	7	22	5	1	0	14	41	0	18	161	8	1	0	56		
2	0	2	4	1	0	2	1	5	0	3	11	4	0	2	2		

Fig. 3D
(Part 1/2)

Gene Name	Protein Description	Accession number	Greatly increase	
			Novel in HIV	in HIV urine
PVRL2	poliovirus receptor-related 2 (herpes)	ENSP00000222847		
APOE	apolipoprotein E [Source:HGNC Symb	ENSP00000223051		
PGLS	6-phosphogluconolactonase [Source:	ENSP00000223061		
ASS1	argininosuccinate synthase 1 [Source:	ENSP00000223214		
NAPSA	napsin A aspartic peptidase [Source:H	ENSP00000223500		
RAP1GDS1	RAP1, GTP-GDP dissociation stimulat	ENSP00000237313	yes	
CA1	carbonic anhydrase I [Source:HGNC S	ENSP00000224324		
LYVE1	lymphatic vessel endothelial hyaluron	ENSP00000225655	yes	
GSTM3	glutathione S-transferase mu 3 (brain	ENSP00000225760		
ACP2	acid phosphatase 2, lysosomal [Sourc	ENSP00000225927		
ANXA1	annexin A1 [Source:HGNC Symbol;Acc	ENSP00000225964		
ACVR1B	activin A receptor, type IB [Source:HG	ENSP00000226218		
GNS	glucosamine (N-acetyl)-6-sulfatase [S	ENSP00000226355		
ORM1	orosomuroid 1 [Source:HGNC Symbo	ENSP00000226877		
HIST1H2AB	histone cluster 1, H2ab [Source:HGNC	ENSP00000227251		
IL18BP	interleukin 18 binding protein [Source	ENSP00000227266		
PI15	peptidase inhibitor 15 [Source:HGNC	ENSP00000227613;		
GGH	gamma-glutamyl hydrolase (conjugas	ENSP00000227666		
GCHFR	GTP cyclohydrolase I feedback regula	ENSP00000228251		
IDH1	isocitrate dehydrogenase 1 (NADP+),	ENSP00000228916		

Fig. 3D
(Part 2/2)

peptides								spectral counts									
0	0	0	0	4	2	2	0	0	0	0	0	9	3	7	0		
0	0	0	0	0	3	0	0	0	0	0	0	0	14	0	0		
0	3	1	2	4	3	4	0	0	8	2	6	20	18	25	0		
0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	0		
0	0	0	0	0	4	2	0	0	0	0	0	0	17	9	0		
0	0	2	1	0	0	0	0	0	0	7	2	0	0	0	0		
0	1	0	0	0	0	2	0	0	3	0	0	0	0	8	0		
5	2	4	3	4	2	4	5	25	5	28	7	24	5	22	16		
0	0	0	0	0	2	1	0	0	0	0	0	0	5	2	0		
0	0	0	0	0	3	0	0	0	0	0	0	0	7	0	0		
0	0	0	0	2	0	1	0	0	0	0	0	4	0	2	0		
0	1	1	1	1	1	2	1	0	2	5	3	9	4	8	2		
3	4	1	0	3	6	6	2	5	25	2	0	9	27	33	3		
13	11	14	11	10	7	11	4	420	145	407	152	176	59	197	36		
1	0	0	0	0	2	0	0	1	0	0	0	0	2	0	0		
4	0	1	0	3	3	1	1	9	0	1	0	9	3	4	1		
0	1	1	0	1	0	0	1	0	1	2	0	1	0	0	1		
2	1	6	0	3	5	7	2	2	1	19	0	4	18	18	5		
0	0	2	0	0	0	0	0	0	0	4	0	0	0	0	0		
2	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0		

Fig. 3E
(Part 1/2)

Gene Name	Protein Description	Accession number	Novel in HIV	Greatly increase in HIV urine
HEXB	hexosaminidase B (beta polypeptide)	ENSP00000228929		
DPEP1	dipeptidase 1 (renal) [Source:HGNC S	ENSP00000229239		
CDH1	cadherin 1, type 1, E-cadherin (epithe	ENSP00000229270		
PDGFRB	platelet-derived growth factor recepto	ENSP00000229707		
NEO1	neogenin 1 [Source:HGNC Symbol;Acc	ENSP00000230538		
CDHR2	cadherin-related family member 2 [Sc	ENSP00000230588		
ASAH1	N-acylsphingosine amidohydrolase (a	ENSP00000230685		
TWSG1	twisted gastrulation homolog 1 (Dros	ENSP00000230990		
FBLN1	fibulin 1 [Source:HGNC Symbol;Acc:3	ENSP00000231751		
LGALS3BP	lectin, galactoside-binding, soluble, 3	ENSP00000232003		
ATRNL	attractin [Source:HGNC Symbol;Acc:8	ENSP00000232447		
SIRPB1	signal-regulatory protein beta 1 [Sour	ENSP00000232907		
A1BG	alpha-1-B glycoprotein [Source:HGNC	ENSP00000233143		
NUCB1	nucleobindin 1 [Source:HGNC Symbol	ENSP00000233156		
CD44	CD44 molecule (Indian blood group) [	ENSP00000233735		
APLP2	amyloid beta (A4) precursor-like prote	ENSP00000233809		
LRP2	low density lipoprotein receptor-relat	ENSP00000234590		
LCAT	lecithin-cholesterol acyltransferase [S	ENSP00000234619		
MCAM	melanoma cell adhesion molecule [So	ENSP00000235933		
NID1	nidogen 1 [Source:HGNC Symbol;Acc:	ENSP00000236671		

Fig. 3E
(Part 2/2)

peptides								spectral counts							
1	0	3	3	1	1	1	0	1	0	17	7	1	4	8	0
0	0	0	0	1	3	0	1	0	0	0	0	2	19	0	4
6	6	4	5	7	4	6	7	136	31	30	55	135	86	96	118
2	0	1	0	1	1	1	1	2	0	2	0	2	1	4	5
0	0	0	0	2	2	1	0	0	0	0	0	6	7	2	0
2	0	3	1	1	2	1	2	4	0	5	3	6	9	2	8
0	0	2	0	1	2	5	0	0	0	20	0	1	3	17	0
1	0	1	1	1	1	2	1	7	0	6	1	10	4	9	1
0	0	0	0	1	1	2	0	0	0	0	0	1	4	5	0
16	9	7	6	13	15	12	9	123	87	38	35	140	148	85	99
9	0	0	0	7	2	2	0	19	0	0	0	12	3	4	0
3	0	0	0	1	0	0	0	4	0	0	0	1	0	0	0
13	20	6	5	14	10	15	6	149	386	45	42	243	113	269	52
0	0	0	0	0	0	1	1	0	0	0	0	0	0	2	1
6	4	6	6	7	4	4	6	109	38	86	56	57	32	47	85
0	0	0	1	0	1	0	0	0	0	0	4	0	1	0	0
5	0	1	5	16	22	12	5	21	0	10	32	60	110	34	27
0	0	0	0	3	1	1	1	0	0	0	0	9	1	3	1
1	0	0	0	4	1	2	0	4	0	0	0	20	2	7	0
6	3	3	2	8	7	7	7	62	10	36	41	119	119	104	105

Fig. 3F
(Part 1/2)

Gene Name	Protein Description	Accession number	Novel in HIV	Greatly increase in HIV urine
YWHAE	tyrosine 3-monooxygenase/tryptophan	ENSP00000237014		
CP	ceruloplasmin (ferroxidase) [Source:HGNC Symbol;Acc:ENSP00000237622]	ENSP00000237622		
MMRN1	multimerin 1 [Source:HGNC Symbol;Acc:ENSP00000237623]	ENSP00000237623		
COL5A3	collagen, type V, alpha 3 [Source:HGNC Symbol;Acc:ENSP00000237858]	ENSP00000237858		
KNG1	kininogen 1 [Source:HGNC Symbol;Acc:ENSP00000238633]	ENSP00000238633		
TNC	tenascin C [Source:HGNC Symbol;Acc:ENSP00000238671]	ENSP00000238671		
AMBP	alpha-1-microglobulin/bikunin precursor [Source:HGNC Symbol;Acc:ENSP00000240618]	ENSP00000240618		
ENPEP	glutamyl aminopeptidase (aminopeptidase) [Source:HGNC Symbol;Acc:ENSP00000241669]	ENSP00000241669		
EGF	epidermal growth factor [Source:HGNC Symbol;Acc:ENSP00000242338]	ENSP00000242338		
FN1	fibronectin 1 [Source:HGNC Symbol;Acc:ENSP00000242464]	ENSP00000242464		
CPVL	carboxypeptidase, vitellogenic-like [Source:HGNC Symbol;Acc:ENSP00000242786]	ENSP00000242786		
CALB1	calbindin 1, 28kDa [Source:HGNC Symbol;Acc:ENSP00000243077]	ENSP00000243077		
HPX	hemopexin [Source:HGNC Symbol;Acc:ENSP00000243219]	ENSP00000243219		
HYAL1	hyaluronoglucosaminidase 1 [Source:HGNC Symbol;Acc:ENSP00000243924]	ENSP00000243924		
ITI4	inter-alpha-trypsin inhibitor heavy chain 4 [Source:HGNC Symbol;Acc:ENSP00000243981]	ENSP00000243981		
C1RL	complement component 1, r subcomponent [Source:HGNC Symbol;Acc:ENSP00000243983]	ENSP00000243983		
CD27	CD27 molecule [Source:HGNC Symbol;Acc:ENSP00000244333]	ENSP00000244333		
LUM	lumican [Source:HGNC Symbol;Acc:ENSP00000244336]	ENSP00000244336		
HEXA	hexosaminidase A (alpha polypeptide) [Source:HGNC Symbol;Acc:ENSP00000244601]	ENSP00000244601		

Fig. 3F
(Part 2/2)

peptides								spectral counts							
2	0	0	0	1	0	1	0	5	0	0	0	1	0	2	0
20	14	3	5	25	11	23	2	161	85	20	19	209	102	249	5
0	0	2	1	0	0	0	1	0	0	3	2	0	0	0	1
2	0	1	2	2	1	0	2	7	0	2	3	5	4	0	3
31	16	12	27	30	28	31	13	329	77	40	241	289	308	209	67
0	0	3	0	2	0	0	1	0	0	8	0	3	0	0	2
38	39	44	30	29	22	43	28	2094	####	2690	1253	####	1123	3436	949
0	1	0	0	2	3	4	0	0	2	0	0	5	6	15	0
20	5	2	11	20	24	21	14	64	16	5	35	111	243	159	48
5	0	10	2	4	5	5	7	20	0	79	5	23	22	14	24
0	0	2	0	1	4	1	0	0	0	2	0	1	9	2	0
0	0	0	4	2	2	0	0	0	0	0	17	7	8	0	0
12	7	7	10	13	9	15	3	81	37	27	56	93	45	112	8
0	0	0	0	1	1	0	0	0	0	0	0	1	2	0	0
8	7	8	8	12	15	13	7	147	77	113	91	578	635	1014	96
5	1	0	1	6	5	7	0	12	6	0	2	15	11	16	0
2	0	2	0	3	1	2	2	21	0	22	0	32	11	21	22
3	2	0	0	5	2	3	0	14	2	0	0	38	6	19	0
1	0	4	4	1	3	3	2	1	0	15	16	1	13	16	3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 14/53676

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - G01N 33/50, G01N 33/52, G01N 33/53, G01N 33/493 (2015.01)

CPC - G01N 33/50, G01N 33/5091, G01N 33/52, G01N 33/53, G01N 33/56988

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) - G01N 33/50, G01N 33/52, G01N 33/53, G01N 33/493 (2015.01)

CPC - G01N 33/50, G01N 33/5091, G01N 33/52, G01N 33/53, G01N 33/56988

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

CPC - G01N 33/493, G01N 33/56983

(keyword limited; terms below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PatBase, Google Scholar; Search terms: Q8TD57, DNAH3, dynein heavy chain 3, dynein, HIV, human immunodeficiency, load, burden, titre, titer, viral, virus, urine, urinary, compare, control, baby, infant, newborn, new-born, child, neonate, toddler, mass spectrometry, quantum dot, chromophore

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	WO 96/24694 A1 (CHAVAN et al.) 15 August 1996 (15.08.1996) p 4, ln 25-32; p 11, ln 11-13; p 17, ln 9-17	1 ----- 2-5, 10-11, 14-15, 18-19, 22
Y	US 2010/0034747 A1 (ROZMYSLOWICZ et al.) 11 February 2010 (11.02.2010) para [0010], [0012], [0040], [0045], [0049], [0050], [0051], [0054], [0057], [0074]	2-5, 10-11, 14-15, 18-19
Y	LEHMANN et al., Intracellular transport of human immunodeficiency virus type 1 genomic RNA and viral production are dependent on dynein motor function and late endosome positioning. J Biol Chem. 22 May 2009, Vol 284, No 21, pages 14572-85. Especially abstract	22
Y	Q8TD57, UniProtKB Accession No. Q8TD57 (DYH3_HUMAN) Dynein heavy chain 3, axonemal, 24 July 2013 [online]. [Retrieved on 22 November 2014]. Retrieved from the internet <URL: http://www.uniprot.org/uniprot/Q8TD57.txt?version=86 > Entire document	22

☐ Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

05 February 2015 (05.02.2015)

Date of mailing of the international search report

25 FEB 2015

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 14/53676

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

---please see extra sheet---

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-5, 10-11, 14-15, 18-19, and 22, limited to Q8TD57 (SEQ ID NO: 1)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 14/53676

Continuation of Box No. III Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I+: Claims 1-23 drawn to a method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient by analyzing a test sample comprising urine from the patient for the presence or concentration of at least one protein, wherein (i) the test sample is analyzed using mass spectrometry, a quantum dot assay or a chromophore assay; or is analyzed using a method comprising contacting the test sample with an antibody or aptamer; (ii) the control sample comprises a urine sample from an untreated HIV-infected control human; or the human patient before receiving anti-HIV medication; or an HIV -uninfected control human; or an HIV-infected control human with controlled infection; (iii) wherein the concentration of the protein in the patient's urine is higher by at least a multiplicity factor than the concentration of the protein in the urine from a control; or is lower by at least a multiplicity factor than the concentration of the protein in the urine from a control; or is equal to or greater than a multiplicity factor of the concentration of the protein in the urine from an untreated HIV-positive control human; or is lower than a multiplicity factor of the concentration of the protein in the urine from an untreated HIV-positive control human; and (iv) wherein the at least one protein has an accession number selected from the group consisting of Q8TD57 (SEQ ID NO: 1), Q18PE1 (SEQ ID NO:2), Q8NHF5 (SEQ ID NO:3), Q8WYL5 (SEQ ID NO:4), Q8IYD8 (SEQ ID NO:5), O14654 (SEQ ID NO:6), Q96AP4 (SEQ ID NO:7), Q9UQ35 (SEQ ID NO:8), Q8N6WO (SEQ ID NO:9), Q9H792 (SEQ ID NO:10), Q9H497 (SEQ ID NO:11), Q9UE35 (SEQ ID NO:12), O00743 (SEQ ID NO:13), Q8WXF8 (SEQ ID NO:14), P81274 (SEQ ID NO:15), Q8NG08 (SEQ ID NO:16), Q96AE7 (SEQ ID NO:17), Q9BZM4 (SEQ ID NO:18), Q5T2D3 (SEQ ID NO:19), Q8IXT5 (SEQ ID NO:20), Q9P225 (SEQ ID NO:21), and Q9Y2I9 (SEQ ID NO:22); or selected from the group consisting of P41222 (PTGDS) (SEQ ID NO:23), P14151 (SELL) (SEQ ID NO:24), Q06418 (TYR03) (SEQ ID NO:25), P52306 (RPIGDSI) (SEQ ID NO:26), and Q9Y5Y7 (LYVE1) (SEQ ID NO:27).

Group I+ will be searched without fee to the extent that it reads on a method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient, wherein (i) the test sample is analyzed using mass spectrometry, (ii) the control sample comprises a urine sample from an untreated HIV-infected control human, (iii) the concentration of the protein in the patient's urine is higher, or equal to or greater than a multiplicity factor of the concentration of the protein in the urine from an untreated HIV-positive control human, and (iv) the protein has accession number Q8TD57 (SEQ ID NO: 1). It is believed that claims 1-5, 10-11, 14-15, 18-19 and 22, limited to Q8TD57 (SEQ ID NO: 1), read on this first named invention. Applicants must indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention being searched/examined. An exemplary election would be: a method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient, wherein (i) the test sample is analyzed using mass spectrometry (ii) the control sample comprises a urine sample from an untreated HIV-infected control human, (iii) the concentration of the protein in the patient's urine is higher by at least a multiplicity factor than the concentration of the protein in the urine from a control, and (iv) the protein has accession number Q18PE1 (SEQ ID NO: 2), i.e. claims 1-5, 10-11, 14-15, and 22, limited to Q18PE1 (SEQ ID NO: 2).

Group II+: Claims 24-33, drawn to a kit for assessing or monitoring systemic HIV viral load in an HIV infected human patient, the kit comprising an antibody or aptamer that binds to at least one protein with an accession number selected from the group consisting of Q8TD57 (SEQ ID NO: 1), Q18PE1 (SEQ ID NO:2), Q8NHF5 (SEQ ID NO:3), Q8WYL5 (SEQ ID NO:4), Q8IYD8 (SEQ ID NO:5), O14654 (SEQ ID NO:6), Q96AP4 (SEQ ID NO:7), Q9UQ35 (SEQ ID NO:8), Q8N6WO (SEQ ID NO:9), Q9H792 (SEQ ID NO:10), Q9H497 (SEQ ID NO:11), Q9UE35 (SEQ ID NO:12), O00743 (SEQ ID NO:13), Q8WXF8 (SEQ ID NO:14), P81274 (SEQ ID NO:15), Q8NG08 (SEQ ID NO:16), Q96AE7 (SEQ ID NO:17), Q9BZM4 (SEQ ID NO:18), Q5T2D3 (SEQ ID NO:19), Q8IXT5 (SEQ ID NO:20), Q9P225 (SEQ ID NO:21), and Q9Y2I9 (SEQ ID NO:22); or selected from the group consisting of P41222 (PTGDS) (SEQ ID NO:23), P14151 (SELL) (SEQ ID NO:24), Q06418 (TYR03) (SEQ ID NO:25), P52306 (RAP1GDS1) (SEQ ID NO:26), and Q9Y5Y7 (LYVE1) (SEQ ID NO:27).

Group II+ could be searched to the extent that it reads on a kit comprising an antibody or aptamer that binds to Q8TD57 (SEQ ID NO: 1), with election as such and payment of an additional fee. It is believed that claims 24-28, limited to Q8TD57 (SEQ ID NO: 1) would read on this first named invention. Applicants must indicate, if applicable, the claims which read on the elected invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention being searched/examined. An exemplary election would be: a kit comprising an antibody or aptamer that binds to Q18PE1 (SEQ ID NO:2), i.e. claims 24-28, limited to Q18PE1 (SEQ ID NO:2).

The inventions listed as Groups I+ and II+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special Technical Features

Group I+ requires method steps of analyzing urine samples for the presence or concentration of at least one protein, not required by Group II+.

Group II+ requires a kit comprising antibody or aptamers, an applicator, and instructional materials, not required by Group I+.

---please see continuation on next extra sheet---

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 14/53676

Continuation of Box No. III Observations where unity of invention is lacking

Common Technical Features

One feature shared by Groups I+ and II+ is a protein having accession numbers selected from:

Q8TD57 (SEQ ID NO: 1), Q18PE1 (SEQ ID NO:2), Q8NFH5 (SEQ ID NO:3), Q8WYL5 (SEQ ID NO:4), Q8IYD8 (SEQ ID NO:5), 014654 (SEQ ID NO:6), Q96AP4 (SEQ ID NO:7), Q9UQ35 (SEQ ID NO:8), Q8N6WO (SEQ ID NO:9), Q9H792 (SEQ ID NO:10), Q9H497 (SEQ ID NO:11), Q9UE35 (SEQ ID NO:12), 000743 (SEQ ID NO:13), Q8WXF8 (SEQ ID NO:14), P81274 (SEQ ID NO:15), Q8NG08 (SEQ ID NO:16), Q96AE7 (SEQ ID NO:17), Q9BZM4 (SEQ ID NO:18), Q5T2D3 (SEQ ID NO:19), Q8IXT5 (SEQ ID NO:20), Q9P225 (SEQ ID NO:21), Q9Y2I9 (SEQ ID NO:22); or P41222 (PTGDS) (SEQ ID NO:23), P14151 (SELL) (SEQ ID NO:24), Q06418 (TYR03) (SEQ ID NO:25), P52306 (RAP1GDS1) (SEQ ID NO:26), and Q9Y5Y7 (LYVE1) (SEQ ID NO:27).

However, these amino acid sequence do not represent a contribution over prior art, because these amino acid sequences were known in the prior art. For example, Q8TD57 (SEQ ID NO: 1) is taught by UniProt Accession No. Q8TD57 (Q8TD57, Dynein heavy chain 3, axonemal, UniProtKB/Swiss-Prot, 24 July 2013 [online]. [Retrieved on 22 November 2014]. Retrieved from the internet <URL: <http://www.uniprot.org/uniprot/Q8TD57.txt?version=86>>]. As the amino acid sequences were known in the art at the time of the invention, they cannot be considered a special technical feature that would otherwise unify the groups.

Another feature shared by Groups I+ and II+ is the assessment or monitoring of systemic HIV viral load by analyzing a test sample comprising urine from the patient for the presence or concentration of the at least one protein.

A further feature shared by the inventions of Groups I+ is (claim 1) a method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient, the method comprising the steps of: analyzing a test sample comprising urine from the patient for the presence or concentration of at least one protein, whereby a test data set is obtained; and, comparing the test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample; whereby the HIV viral load in the patient is assessed or monitored.

A further feature shared by the inventions of Groups II+ is (claims 24 and 29) a kit for assessing or monitoring systemic HIV viral load in an HIV infected human patient, the kit comprising an antibody or aptamer that binds to at least one protein; an applicator; and, an instructional material for the use of the kit, wherein the instruction material comprises instructions for analyzing a test sample comprising urine from the patient for the presence or concentration of the at least one protein.

However, these shared technical features do not represent a contribution over prior art, because the shared technical features are taught by WO 96/24694 A1 to Chavan et al. (hereinafter 'Chavan').

Chavan discloses the assessment or monitoring of systemic HIV viral load (p 4, In 25-32 - "This invention also provides a method for determining in a sample viral load of HIV") by analyzing a test sample comprising urine from the patient (p 11, In 11-13 - "Biological fluids are well-known to those of ordinary skill in the art and include, but are not limited to, urine") for the presence or concentration of the at least one protein (p 4, In 25-32 - "(a) determining the activity of reverse transcriptase in the sample according to the aforementioned method of determining in a sample the activity of an enzyme which catalyzes the synthesis of an RNA-DNA heteroduplex from a nucleic acid template; and (b) calculating the viral load of HIV in the sample as a predetermined function of the activity of reverse transcriptase determined in step (a)."). As the technical feature was known in the art at the time of the invention, it cannot be considered a special technical feature that would otherwise unify the groups.

Chavan further discloses (claim 1) a method of assessing or monitoring systemic HIV viral load in an HIV-infected human patient (p 4, In 25-32 - "This invention also provides a method for determining in a sample viral load of HIV"), the method comprising the steps of: analyzing a test sample comprising urine from the patient (p 11, In 11-13 - "Biological fluids are well-known to those of ordinary skill in the art and include, but are not limited to, urine") for the presence or concentration of at least one protein, whereby a test data set is obtained (p 4, In 25-32 - "(a) determining the activity of reverse transcriptase in the sample according to the aforementioned method of determining in a sample the activity of an enzyme which catalyzes the synthesis of an RNA-DNA heteroduplex from a nucleic acid template; and (b) calculating the viral load of HIV in the sample as a predetermined function of the activity of reverse transcriptase determined in step (a)."; p 17, In 9-11 - "the viral load may be calculated from the actual quantity of reverse transcriptase molecules calculated in the sample"); and, comparing the test data set with a control data set relating to the presence or concentration of the at least one protein in a control sample (p 17, In 12-17 - "As described above, the actual quantity of reverse transcriptase molecules in a sample may be determined by comparing the fluorescence resulting in the sample to the fluorescence produced by a standard predetermined to contain a specific quantity of reverse transcriptase molecules."); whereby the HIV viral load in the patient is assessed or monitored (p 4, In 25-32; p 17, In 4-17). As the technical feature was known in the art at the time of the invention, it cannot be considered a special technical feature that would otherwise unify the groups.

Chavan discloses a kit for assessing or monitoring systemic HIV viral load in an HIV infected human patient (p 19, In 30-32 - "the subject kit is useful for ... determining the viral load of HIV in a subject infected with HIV"). Although Chavan does not specifically teach that the kit comprises an antibody or aptamer that binds to at least one protein; an applicator; and, an instructional material for the use of the kit, wherein the instruction material comprises instructions for analyzing a test sample comprising urine from the patient for the presence or concentration of the at least one protein, this kit composition would have been obvious to one of ordinary skill in the art in view of Chavan's teaching that the viral load can be determined (p 4, In 25-32; p 17, In 4-17) from a test sample comprising urine from the patient (p 11, In 11-13), wherein the concentration of the protein may be determined by comparing the fluorescence resulting in the sample to the fluorescence produced by a control (p 17, In 12-17). Since Chavan further teaches that prior art immunological detection methods were known in the art (p 2, In 22), one of ordinary skill in the art would have found it obvious to have a kit that comprises an antibody or aptamer that binds to at least one protein; an applicator; and, an instructional material for the use of the kit. As the technical feature was known in the art at the time of the invention, it cannot be considered a special technical feature that would otherwise unify the groups.

Groups I+ and II+ therefore lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.