(19)





(11) **EP 1 415 005 B1**

(12)

EUROPEAN PATENT SPECIFICATION

- (45) Date of publication and mention of the grant of the patent:21.11.2012 Bulletin 2012/47
- (21) Application number: 01996222.4
- (22) Date of filing: 07.12.2001

(51) Int Cl.: **C07K 14/15**^(2006.01) **C12Q 1/68**^(2006.01)

C12N 7/00 ^(2006.01) G01N 33/574 ^(2006.01)

- (86) International application number: PCT/US2001/047824
- (87) International publication number:
 WO 2002/046477 (13.06.2002 Gazette 2002/24)

(54) ENDOGENOUS RETROVIRUSES UP-REGULATED IN PROSTATE CANCER

HOCHREGULIERTE ENDOGENE RETROVIREN IN PROSTATAKREBS

RETROVIRUS ENDOGENES REGULES POSITIVEMENT DANS LE CANCER DE LA PROSTATE

- (84) Designated Contracting States:
 AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
 MC NL PT SE TR
- (30) Priority: 07.12.2000 US 251830 P 07.12.2001 US 16604
- (43) Date of publication of application: 06.05.2004 Bulletin 2004/19
- (60) Divisional application: 10176402.5 / 2 336 368 10176900.8 / 2 339 035
- (73) Proprietor: Novartis Vaccines and Diagnostics, Inc.
 Emeryville, CA 94608 (US)
- (72) Inventors:
 - GARCIA, Pablo San Francisco, CA 94131 (US)
 - HARDY, Stephen, F. San Francisco, CA 94121 (US)
 WILLIAMS, Lewis, T.
 - Mill Valley, CA 94902 (US)ESCOBEDO, Jaime
 - Alamo, CA 94507 (US)
- (74) Representative: Marshall, Cameron John et al Carpmaels & Ransford One Southampton Row London WC1B 5HA (GB)

(56) References cited: US-A- 5 858 723

- TÖNJES R R ET AL: "Genome-wide screening, cloning, chromosomal assignment, and expression of full-length human endogenous retrovirus type K." JOURNAL OF VIROLOGY. UNITED STATES NOV 1999, vol. 73, no. 11, November 1999 (1999-11), pages 9187-9195, XP002238923 ISSN: 0022-538X cited in the application
- BARBULESCU M ET AL: "MANY HUMAN ENDOGENOUS RETROVIRUS K (HERV-K) PROVIRUSES ARE UNIQUE TO HUMANS" CURRENT BIOLOGY, CURRENT SCIENCE,, GB, vol. 9, 1999, pages 861-868, XP000953273 ISSN: 0960-9822 cited in the application
- MAYER J ET AL: "Chromosomal assignment of human endogenous retrovirus K (HERV-K) env open reading frames." CYTOGENETICS AND CELL GENETICS, vol. 79, no. 1-2, 1997, pages 157-161, XP009009741 ISSN: 0301-0171
- MAYER JENS ET AL: "An almost-intact human endogenous retrovirus K on human chromosome 7." NATURE GENETICS, vol. 21, no. 3, March 1999 (1999-03), pages 257-258, XP002239083 ISSN: 1061-4036 cited in the application
- HUANG HAOJIE ET AL: "FRA7G extends over a broad region: Coincidence of human endogenous retroviral sequences (HERV-H) and small polydispersed circular DNAs (spcDNA) and fragile sites." ONCOGENE, vol. 16, no. 18, 7 May 1998 (1998-05-07), pages 2311-2319, XP009009633 ISSN: 0950-9232

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

- LOEWER R ET AL: "The viruses in all of us: Characteristics and biological significance of human endogenous retrovirus sequences" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 93, no. 11, May 1996 (1996-05), pages 5177-5184, XP002233322 ISSN: 0027-8424 cited in the application
- MAGIN CHRISTINE ET AL: "Corf, the Rev/Rex homologue of HTDV/HERV-K, encodes an arginine-rich nuclear localization signal that exerts a trans-dominant phenotype when mutated." VIROLOGY, vol. 274, no. 1, 15 August 2000 (2000-08-15), pages 11-16, XP002239089 ISSN: 0042-6822 cited in the application
- SMITH R D ET AL: "The human endogenous retrovirus HERV-K in prostate cancer." AMERICAN JOURNAL OF HUMAN GENETICS, vol. 69, no. 4 Supplement, October 2001 (2001-10), page 275 XP009009588 51st Annual Meeting of the American Society of Human Genetics;San Diego, California, USA; October 12-16, 2001 ISSN: 0002-9297
- SMITH RICHARD DARON ET AL: "Human endogenous retrovirus HERV-K expression in prostate cancer." JOURNAL OF UROLOGY, vol. 165, no. 5 Supplement, May 2001 (2001-05), pages 136-137, XP009009589 Annual Meeting of the American Urological Association, Inc.;Anaheim, California, USA; June 02-07, 2001 ISSN: 0022-5347

Description

TECHNICAL FIELD

⁵ **[0001]** The present invention relates to the diagnosis of cancer, particularly prostate cancer. In particular, it relates to a subgroup of human endogenous retroviruses (HERVs) which show up-regulated expression in tumors, particularly prostate tumors.

BACKGROUND ART

10

[0002] Prostate cancer is the most common type of cancer in men in the USA. Benign prostatic hyperplasia (BPH) is the abnormal growth of benign prostate cells in which the prostate grows and pushes against the urethra and bladder, blocking the normal flow of urine. More than half of the men in the USA between the ages of 60 and 70 and as many as 90 percent between the ages of 70 and 90 have symptoms of BPH. Although this condition is seldom a threat to life, it may require treatment to relieve symptoms.

- ¹⁵ may require treatment to relieve symptoms. [0003] Cancer that begins in the prostate is called primary prostate cancer (or prostatic cancer). Prostate cancer may remain in the prostate gland, or it may spread to nearby lymph nodes and may also spread to the bones, bladder, rectum, and other organs. Prostate cancer is diagnosed by measuring the levels of **pmstate-specific** antigen (PSA) and prostatic acid phosphatase (PAP) in the blood. The level of PSA in blood may rise in men who have prostate cancer, BPH, or an
- ²⁰ infection in the prostate. The level of PAP rises above normal in many prostate cancer patients, especially if the cancer has spread beyond the prostate. However, one cannot diagnose prostate cancer with these tests alone because elevated PSA or PAP levels may also indicate other, non-cancerous problems.

[0004] In order to help determine whether conditions of the prostate are benign or malignant further tests such as transrectal ultrasonography, intravenous pyelogram, and cystoscopy are usually performed. If these test results suggest

- that cancer may be present, the patient must undergo a biopsy as the only sure way to diagnose prostate cancer. Consequently, it is desirable to provide a simple and direct test for the early detection and diagnosis of prostate cancer without having to undergo multiple rounds of cumbersome testing procedures. It is also desirable and necessary to provide compositions and methods for the prevention and/or treatment of prostate cancer.
- [0005] It is an object of the invention to provide materials that can be used in the prevention, treatment and diagnosis of prostate cancer. It is a further object to provide improvements in the prevention, treatment and diagnosis of prostate cancer.

DISCLOSURE OF THE INVENTION

³⁵ **[0006]** It has been found that human endogenous retroviruses (HERVs) of the HML-2 subgroup of the HERV-K family show up-regulated expression in prostate tumors. This finding can be used in prostate cancer screening, diagnosis and therapy.

[0007] The invention provides a method for diagnosing cancer, especially prostate cancer, the method comprising the step of detecting the presence or absence of an expression product of a HML-2 endogenous retrovirus in a patient

- 40 sample, wherein the patient sample contains prostate cells and for wherein the patient is suspected of having prostate cancer wherein up-regulation of expression of at least 150% relative to a negative control is indicative of prostate cancer. [0008] The HML-2 expression product which is detected is either a mRNA transcript or a polypeptide translated from such a transcript These expression products may be detected directly or indirectly. A direct test uses an assay which detects HML-2 RNA or polypeptide in a patient sample. An indirect test uses an assay which detects biomolecules which
- ⁴⁵ are not directly expressed *in vivo* from HML-2 *e.g.* an assay to detect cDNA which has been reverse-transcribed from a HML-2 mRNA, or an assay to detect an antibody which has been raised in response to a HML-2 polypeptide.

A - THE PATIENT SAMPLE

- ⁵⁰ **[0009]** Where the diagnostic method of the invention is based on HML-2 mRNA, the patient sample will generally comprise cells, preferably, prostate cells. These may be present in a sample of tissue, preferably, prostate tissue, or may be cells, preferably, prostate cells which have escaped into circulation (*e.g.* during metastasis). Instead of or as well as comprising prostate cells, the sample may comprise virions which contain mRNA from HML-2.
- [0010] Where the diagnostic method of the invention is based on HML-2 polypeptides, the patient sample may comprise cells, preferably, prostate cells and/or virions (as described above for mRNA), or may comprise antibodies which recognize HML-2 polypeptides. Such antibodies will typically be present in circulation.

[0011] In general, therefore, the patient sample is tissue sample (*e.g.* a biopsy), preferably, a prostate sample (*e.g.* a biopsy) or a blood sample.

[0012] The patient is generally a human, preferably human male, and more preferably an adult human male.

[0013] Expression products may be detected in the patient sample itself, or it may be detected in material derived from the sample (*e.g.* the supernatant of a cell lysate, or a RNA extract, or cDNA generated from a RNA extract, or polypeptides translated from a RNA extract, or cells derived from culture of cells extracted from a patient *etc.*). These are still considered to be "patient samples" within the meaning of the invention.

[0014] Methods of the invention can be conducted *in vitro* or *in vivo*.

[0015] Other possible sources of patient samples include isolated cells, whole tissues, or bodily fluids (e.g. blood, plasma, serum, urine, pleural effusions, cerebro-spinal fluid, *etc.*)

10 B - THE mRNA EXPRESSION PRODUCT

5

[0016] Where the diagnostic method of the invention is based on mRNA detection, it typically involves detecting a RNA comprising six basic regions. From 5' to 3', these are:

- 15 1. A sequence which has at least 75% identity to SEQ ID 155 (e.g. 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity); or a sequence which has at least 50% identity to SEQ ID 155 (e.g. 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 20 98%, 99%, 99.5%, 99.9%, 100% identity) and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, etc., 25 contiguous nucleotides) of SEQ ID 155; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25,30,35,40,45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, etc., contiguous nucleotides) of SEQ ID 155 and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with 30 at least a 95% confidence level. This sequence will typically be at the 5' end of the RNA. SEQ ID 155 is the nucleotide sequence of the start of R region in the LTR of the 'ERVK6' HML-2 virus [ref. 1]. This portion of the R region is found in all full-length HML-2 transcripts.
- 2. A downstream region comprising a sequence which has at least 75% sequence identity to SEQ ID 156 (e.g. 76%, 35 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity); or a sequence which has at least 50% identity to SEQ ID 156 (e.g. 51 %, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) and is expressed at least 1.5 fold 40 (e.g. 2, 2.5, 5, 10, 20, S0, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 45 235, 240, 245, 250, 255, etc., contiguous nucleotides) of SEQ ID 156; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 245, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, etc., contiguous nucleotides) of SEQ ID 156 and is 50 expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level. SEQ ID 156 is the nucleotide sequence of the RU₅ region downstream of SEQ ID 155 in the ERVK6 LTR. This region is found in full-length HML-2 transcripts, but may not be present in all mRNAs transcribed from a HML-2 LTR promoter.
- ⁵⁵ 3. A downstream region comprising a sequence which has at least 75% sequence identity to SEQ ID 6 (*e.g.* 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity); or a sequence which has at least 50% identity to SEQ ID 6 (*e.g.* 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%,

71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, etc., contiguous nucleotides) of SEQ ID 6; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, etc., contiguous nucleotides) of SEQ m 6 and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level. SEQ ID 6 is the nucleotide sequence of the region of the ERVK6 virus between the U_5 region and the first 5' splice site. This region is found in full-length HML-2 transcripts, but has been lost by some variants and, like region 2 above, may not be present in all mRNAs transcribed from a HML-2 LTR promoter.

15

5

10

4. A downstream region comprising any RNA sequence. This region will typically comprise the coding sequence of one or more HML-2 polypeptides, but may alternatively comprise: a mutant viral coding sequence; a viral or nonviral non-coding sequence; or a non-viral coding sequence. Transcription of any of these sequences can come under the control of a HML-2 LTR.

20

30

35

40

5. A downstream region comprising a sequence which has at least 75% sequence identity to SEQ ID 5 (e.g. 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity); or a sequence which has at least 50% identity to SEQ ID 5 (e.g. 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 25 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (*i.e.*, non cancerous) cell with at least a 95% confidence level; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g. 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, etc., contiguous nucleotides) of SEQ ID 5; or a sequence which has at least 80% identity (e.g. 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.9%, 100% identity) to at least a 20 contiguous nucleotide fragment (e.g., 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, etc., contiguous nucleotides) of SEQ ID 5 and is expressed at least 1.5 fold (e.g. 2, 2.5, 5, 10, 20, 50, etc., fold) higher level relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level. SEQ ID 5 is the nucleotide sequence of the U₃R region in the 3' end of ERVK6. This sequence will typically be near the 3' end of the RNA, immediately preceding any polyA tail.

6. A 3' polyA tail.

- 45 [0017] The percent identity of the sequences described above are determined by the Smith-Waterman algorithm using the default parameters: open gap penalty = -20 and extension penalty =-5. [0018] These mRNA molecules are referred to below as "PCA-mRNA" molecules ("prostate cancer associated mRNA"), and endogenous viruses which express these PCA-mRNAs are referred to as PCAVs ("prostate cancer associated viruses"). Nevertheless, said PCAVs may also be associated with other types of cancer.
- 50 [0019] Although some PCA-mRNAs include all six of these regions, most HERVs are defective in that they have accumulated multiple stop codons, frameshifts, or larger deletions etc. This means that many PCA-mRNAs do not include all six regions. As all PCA-mRNAs are transcribed under the control of the same group of LTRs, however, transcription of all PCA-mRNAs is up-regulated in prostate tumors even though the mRNA may not encode functional polypeptides. [0020] Where a mRNA to be detected is driven by 5' LTR of HML-2 in genomic DNA, the first of these regions will
- 55 always be present, but the remaining five are optional. Conversely, where a mRNA to be detected is controlled by 3' LTR of HML-2, the fifth of these regions will always be present, but the remaining five are optional.

[0021] In general, therefore, the mRNA to be detected has the formula $N_1-N_2-N_3-N_4-N_5$ -polyA, wherein:

- N₁ has at least 75% sequence identity to SEQ ID 155; or has at least 50% identity to SEQ ID 155 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID 155; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID 155 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level;
- N₂ has at least 75% sequence identity to SEQ ID 156; or has at least 50% identity to SEQ ID 156 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID 156; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID 156 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level;
- N₃ has at least 75% sequence identity to SEQ ID 6; or has at least 50% identity to SEQ ID 6 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID 6; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID 6 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level;
 - N₄ comprises any RNA sequence;
 - N₅ has at least 75% sequence identity to SEQ ID 5; or has at least 50% identity to SEQ ID 5 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; or has at least 80% identity to at least a 20 contiguous nucleotide fragment of SEQ ID 5; or has at least 80% identity
- 20 to at least a 20 contiguous nucleotide fragment of SEQ ID 5 and is expressed at least 1.5 fold higher relative to expression in a normal (i.e., non cancerous) cell with at least a 95% confidence level; and
 - at least one of N_1 , N_2 , N_3 , N_4 or N_5 is present, but polyA is optional.
- **[0022]** Although only at least one of N_1 , N_2 , N_3 , N_4 or N_5 needs to be present, it is preferred that two, three, four or five of these regions are present. It is preferred that at least one of N_1 and/or N_5 is present.

[0023] N₁ is preferably present in the mRNA to be detected (*i.e.* the invention is preferably based on the detection of mRNA driven by a 5' LTR). More preferably, at least N₁-N₂ is present.

- [0024] Where N₁ is present, it is preferably at the 5' end of the mRNA (*i.e.* 5'- N₁-...).
- [0025] Where N₅ is present, it is preferably immediately before a 3' polyA tail (*i.e.*... -N₅-polyA-3').
- **[0026]** Where N_4 is present, it preferably comprises a polypeptide-coding sequence (*e.g.* encoding a HML-2 polypeptide). Examples of HML-2 polypeptide-coding sequences are described below.
 - [0027] The RNA will generally have a 5' cap.

B.1 -Enriching RNA in a sample

35

30

5

10

15

- multiple PCAV DNA templates, whereas increased PCA-mRNA levels are only found in cancerous cells. As an alternative, a RNA-specific assay can be used which is not affected by the presence of homologous DNA.
 [0029] Methods for extracting RNA from biological samples are well known [e.g. refs. 2 & 8] and include methods based on guanidinium buffers, lithium chloride, SDS/potassium acetate *etc.* After total cellular RNA has been extracted, mRNA may be enriched e.g. using oligo-dT techniques.
- ⁴⁵ **[0030]** Methods for removing DNA from biological samples without removing mRNA are well known [*e.g.* appendix C of ref. 2] and include DNase digestion.

[0031] Methods for removing DNA, but not RNA, comprising PCA-mRNA sequences will use a reagent which is specific to a sequence within a PCA-mRNA *e.g.* a restriction enzyme which recognizes a DNA sequence within SEQ ID 4, but which does not cleave the corresponding RNA sequence.

⁵⁰ **[0032]** Methods for specifically purifying PCA-mRNAs from a sample may also be used. One such method uses an affinity support which binds to PCA-mRNAs. The affinity support may include a polypeptide sequence which binds to the PCAV-mRNA *e.g.* the cORF polypeptide, which binds to the LTR of HERV-K mRNAs in a sequence-specific manner, or HIV Rev protein, which has been shown to recognize the HERV-K LTR [3].

55 <u>B.2 - Direct detection of RNA</u>

[0033] Various techniques are available for detecting the presence or absence of a particular RNA sequence in a sample [*e.g.* refs. 2 & 8]. If a sample contains genomic PCAV DNA, the detection technique will generally be RNA-

specific; if the sample contains no PCAV DNA, the detection technique may or may not be RNA-specific.

[0034] Hybridization-based detection techniques may be used, in which a polynucleotide probe complementary to a region of PCA-mRNA is contacted with a RNA-containing sample under hybridizing conditions. Detection of hybridization indicates that nucleic acid complementary to the probe is present. Hybridization techniques for use with RNA include Northern blots, in situ hybridization and arrays.

- [0035] Sequencing may also be used, in which the sequence(s) of RNA molecules in a sample are obtained. These techniques reveal directly whether a sequence of interest is present in a sample. Sequence determination of the 5' end of a RNA corresponding to N_1 will generally be adequate.
- [0036] Amplification-based techniques may also be used. These include PCR, SDA, SSSR, LCR, TMA, NASBA, T7 10 amplification etc. The technique preferably gives exponential amplification. A preferred technique for use with RNA is RT-PCR [e.g. see chapter 15 of ref. 2]. RT-PCR of mRNA from prostate cells is reported in references 4, 5, 6 & 7.

B.3 - Indirect detection of RNA

- 15 [0037] Rather than detect RNA directly, it may be preferred to detect molecules which are derived from RNA (i.e. indirect detection of RNA). A typical indirect method of detecting mRNA is to prepare cDNA by reverse transcription and then to directly detect the cDNA. Direct detection of cDNA will generally use the same techniques as described above for direct detection of RNA (but it will be appreciated that methods such as RT-PCR are not suitable for DNA detection and that cDNA is doublo-stranded, so detection techniques can be based on a sequence, on its complement, or on the 20
- double-stranded molecule).

5

C - POLYPEPTIDE EXPRESSION PRODUCT

- [0038] Where the method is based on polypeptide detection, it will involve detecting expression of a polypeptide 25 encoded by a PCAV-mRNA. This will typically involve detecting one or more of the following HML-2 polypeptides: gag, prt, pol, env, cORF. Although some PCA-mRNAs encode all of these polypeptides (e.g. ERVK6 [1]), the polypeptidecoding regions of most HERVs (including PCAVs) contain mutations which mean that one or more coding-regions in the mRNA transcript are either mutated or absent. Thus not all PCAVs have the ability to encode all HML-2 polypeptides. [0039] The transcripts which encode HML-2 polypeptides are generated by alternative splicing of the full-length mRNA
- 30 copy of the endogenous genome [e.g. Figure 4 of ref. 143]. [0040] <u>HML-2 gag polypeptide</u> is encoded by the first long ORF in a complete HML-2 genome [140]. Full-length gag polypeptide is proteolytically cleaved.

[0041] Examples of gag nucleotide sequences are: SEQ IDs 7, 8, 9 & 11 [HERV-K(CH)]; SEQ ID 85 [HERV-K108]; SEQ ID 91 [HERV-K(C7)]; SEQ ID 97 [BFRV-K(II)]; SEQ ID 102 [HERV-K10].

- 35 [0042] Examples of gag polypeptide sequences are: SEQ IDs 46, 47, 48, 49, 56 & 57 [HERV-K(CH)]; SEQ ID 92 [HERV-K(C7)]; SEQ ID 98 [HERV-K(II)]; SEQ IDs 103 & 104 [HERV-K10]; SEQ ID 146 ['ERVK66'].
 - [0043] An alignment of gag polypeptide sequences is shown in Figure 7.
 - [0044] HML-2 prt polypeptide is encoded by the second long ORF in a complete HML-2 genome. It is translated as a gag-prt fusion polypeptide. The fusion polypeptide is proteolytically cleaved to give a protease.
- 40 [0045] Examples of prt nucleotide sequences are: SEQ ID 86 [HERV-K(108)]; SEQ ID 99 [HERV-K(II)]; SEQ ID 105 [HERV-K10].
 - [0046] Examples of prt polypeptide sequences are: SEQ ID 106 [HERV-K10]; SEQ ID 147 ['ERVK6'].
- [0047] HML-2 pol polypeptide is encoded by the third long ORF in a complete HML-2 genome. It is translated as a gag-prt-pol fusion polypeptide. The fusion polypeptide is proteolytically cleaved to give three pol products - reverse 45 transcriptase, endonuclease and integrase [14].
 - [0048] Examples of pol nucleotide sequences are: SEQ ID 87 [HERV-K(108)]; SEQ ID 93 [HERV-K(C7)]; SEQ ID 100 [HERV-K(II)]; SEQ ID 107 [HERV-K10].

[0049] Examples of pol polypeptide sequences are: SEQ ID 94 [HERV-K(C7)]; SEQ ID 108 [HERV-K10]; SEQ ID 148 ['ERVK6'].

50 [0050] An alignment of pol polypeptide sequences is shown in Figure 8.

[0051] HML-2 env polypeptide is encoded by the fourth long ORF in a complete HML-2 genome. The translated polypeptide is proteolytically cleaved.

[0052] Examples of env nucleotide sequences are: SEQ ID 88 [HERV-K(108)]; SEQ ID 95 [HERV-K(C7)]; SEQ ID 101 [HERV-K(II)]; SEQ ID 107 [HERV-K10].

- 55 [0053] Examples of env polypeptide sequences are: SEQ ID 96 [HERV-K(C7)]; SEQ ID 108 [HERV-K10] ; SEQ ID 149 ['ERVK6'].
 - [0054] Alignments of env polynucleotide and polypeptide sequences are shown in Figures 6 and 9.
 - [0055] HML-2 cORP polypeptide is encoded by an ORF which shares the same 5' region and start codon as env.

After amino acid 87, a splicing event removes env-coding sequences and the cORF-coding sequence continues in the reading frame +1 relative to that of env [15,16; see below]. cORF has also been called Rec [17].

- [0056] Examples of cORF nucleotide sequences are: SEQ ID 89 and SEQ ID 90 [HERV-K(108)]
- [0057] Examples of cORF polypeptide sequences are SEQ ID 109.

C.1- Direct detection of HML-2 polypeptides

5

10

15

40

[0058] Various techniques are available for detecting the presence or absence of a particular polypeptides in a sample. These are generally immunoassay techniques which are based on the specific interaction between an antibody and an antigenic amino acid sequence in the polypeptide. Suitable techniques include standard immunohistological methods, immunoprecipitation, immunofluorescence, ELISA, RIA, FIA, *etc.*

[0059] In general, therefore, the invention provides a method for detecting the presence of and/or measuring a level of a polypeptide in a biological sample, wherein the method uses an antibody specific for the polypeptide. The method generally comprises the steps of: a) contacting the sample with an antibody specific for the polypeptide; and b) detecting binding between the antibody and polypeptides in the sample.

- **[0060]** Polypeptides can also be detected by functional assays *e.g.* assays to detect binding activity or enzymatic activity. For instance, a functional assay for cORF is disclosed in references 16,129 & 130. A functional assay for the protease is disclosed in reference 140.
- [0061] Another way for detecting polypeptides is to use standard proteomics techniques *e.g.* purify or separate polypep tides and then use peptide sequencing. For example, polypeptides can be separated using 2D-PAGE and polypeptide spots can be sequenced (*e.g.* by mass spectroscopy) in order to identify if a sequence is present in a target polypeptide.
 [0062] Detection methods may be adapted for use *in vivo* (*e.g.* to locate or identify sites where cancer cells are present). In these embodiments, an antibody specific for a target polypeptide is administered to an individual (*e.g.* by injection) and the antibody is located using standard imaging techniques (*e.g.* magnetic resonance imaging, computed tomography
- ²⁵ scanning, *etc.*). Appropriate labels (*e.g.* spin labels *etc.*) will be used. Using these techniques, cancer cells are differentially labeled.

[0063] An immunofluorescence assay can be easily performed on cells without the need for purification of the target polypeptide. The cells are first fixed onto a solid support, such as a microscope slide or microtiter well. The membranes of the cells are then permeablized in order to permit entry of polypeptide-specific antibody (NB: fixing and permeabilization

³⁰ can be achieved together). Next, the fixed cells are exposed to an antibody which is specific for the encoded polypeptide and which is fluorescently labeled. The presence of this label (*e.g.* visualized under a microscope) identifies cells which express the target PCAV polypeptide. To increase the sensitivity of the assay, it is possible to use a second antibody to bind to the anti-PCAV antibody, with the label being carried by the second antibody. [18]

35 <u>C.2-Indirect detection of HML-2 polypeptides</u>

[0064] Rather than detect polypeptides directly, it may be preferred to detect molecules which are produced by the body in response to a polypeptide (*i.e.* indirect detection of a polypeptide). This will typically involve the detection of antibodies, so the patient sample will generally be a blood sample. Antibodies can be detected by conventional immunoassay techniques *e.g.* using PCAV polypeptides which will typically be immobilized.

[0065] Antibodies against HERV-K polypeptides have been detected in humans [143].

[0066] References to a percentage sequence identity between two amino acid sequences means that, when aligned, that percentage of amino acids are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art, for example those described in section

⁴⁵ 7.7.18 of reference 11. A preferred alignment is determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. The Smith-Waterman homology search algorithm is taught in reference 32.

[0067] The term "polypeptide" refers to amino acid polymers of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids. The terms also encompass an

- ⁵⁰ amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, *etc.*), as well as other modifications known in the art. Polypeptides can occur as single chains or associated chains. Polypeptides of the invention can be naturally or non-
- ⁵⁵ naturally glycosylated (*i.e.* the polypeptide has a glycosylation pattern that differs from the glycosylation pattern found in the corresponding naturally occurring polypeptide).

[0068] Mutants can include amino acid substitutions, additions or deletions. The amino acid substitutions can be conservative amino acid substitutions or substitutions to eliminate non-essential amino acids, such as to alter a glyco-

sylation site, a phosphorylation site or an acetylation site, or to minimize misfolding by substitution or deletion of one or more cysteine residues that are not necessary for function. Conservative amino acid substitutions are those that preserve the general charge, hydrophobicity/hydrophilicity, and/or steric bulk of the amino acid substituted. Variants can be designed so as to retain or have enhanced biological activity of a particular region of the polypeptide (*e.g.* a functional

- ⁵ domain and/or, where the polypeptide is a member of a polypeptide family, a region associated with a consensus sequence). Selection of amino acid alterations for production of variants can be based upon the accessibility (interior vs. exterior) of the amino acid (*e.g.* ref. 33), the thermostability of the variant polypeptide (*e.g.* ref. 34), desired glycosylation sites (*e.g.* ref. 35), desired disulfide bridges (*e.g.* refs. 36 & 37), desired metal binding sites (*e.g.* refs.38 & 39), and desired substitutions with in proline loops (*e.g.* ref. 40). Cysteine-depleted muteins can be produced as disclosed in
- ¹⁰ reference 41.

C.4 -Antibody materials

[0069] Antibodies be polyclonal or monoclonal and may be produced by any suitable means (*e.g.* by recombinant expression).

[0070] Antibodies may include a label. The label may be detectable directly, such as a radioactive or fluorescent label. Alternatively, the label may be detectable indirectly, such as an enzyme whose products are detectable (*e.g.* luciferase, ß-galactosidase, peroxidase *etc.*).

[0071] Antibodies may be attached to a solid support.

[0072] Antibodies be prepared by administering (*e.g.* injecting) a polypeptide of the invention to an appropriate animal (*e.g.* a rabbit, hamster, mouse or other rodent).

[0073] Antigen-binding fragments of antibodies include Fv, scFv, Fc, Fab, F(ab')₂ etc.

[0074] To increase compatibility with the human immune system, the antibodies may be chimeric or humanized [*e.g.* refs. 42 & 43], or fully human antibodies may be used. Because humanized antibodies are far less immunogenic in humans the the original non human menadanal artihodies, they can be used for the treatment of humans with for less

- ²⁵ humans than the original non-human monoclonal antibodies, they can be used for the treatment of humans with far less risk of anaphylaxis. Thus, these antibodies may be preferred in therapeutic applications that involve in vivo administation to a human such as, use as radiation sensitizers for the treatment of neoplastic disease or use in methods to reduce the side effects of cancer therapy.
- 30

20

[0075] Humanized antibodies may be achieved by a variety of methods including, for example:

(1) grafting non-human complementarity determining regions (CDRs) onto a human framework and constant region ("humanizing"), with the optional transfer of one or more framework residues from the non-human antibody; (2) transplanting entire non-human variable domains, but "cloaking" them with a human-like surface by replacement of surface residues ("veneering"). In the present invention, humanized antibodies will include both "humanized" and

³⁵ "veneered" antibodies. [44, 45, 46, 47, 48, 49, 50].

[0076] CDRs are amino acid sequences which together define the binding affinity and specificity of a Fv region of a native immunoglobulin binding site [*e.g.* refs. 51 & 52].

[0077] The phrase "constant region" refers to the portion of the antibody molecule that confers effector functions. In chimeric antibodies, mouse constant regions are substituted by human constant regions. The constant regions of humanized antibodies are derived from human immunoglobulins. The heavy chain constant region can be selected from any of the 5 isotypes: alpha, delta, epsilon, gamma or mu.

[0078] One method of humanizing antibodies comprises aligning the heavy and light chain sequences of a non-human antibody to human heavy and light chain sequences, replacing the non-human framework residues with human framework residues with human framework residues with human framework residues with human framework residues at the non-human framework residues with human framework residues at the non-human framework residues with human framework residues at the non-human framework residues with human framework residues at the non-human framework residues with human framework residues at the non-human framework residues with human framework residues at the non-human fr

residues based on such alignment, molecular modeling of the conformation of the humanized sequence in comparison to the conformation of the non-human parent antibody, and repeated back mutation of residues in the framework region which disturb the structure of the non-human CDRs until the predicted conformation of the CDRs in the humanized sequence model closely approximates the conformation of the non-human CDRs of the parent non-human antibody. Such humanized antibodies may be further derivatized to facilitate uptake and clearance e.g, via Ashwell receptors. [refs. 53 & 54]

[0079] Humanized or fully-human antibodies can also be produced using transgenic animals that are engineered to contain human immunoglobulin loci. For example, ref. 55 discloses transgenic animals having a human Ig locus wherein the animals do not produce functional endogenous immunoglobulins due to the inactivation of endogenous heavy and light chain loci. Ref. 56 also discloses transgenic non-primate mammalian hosts capable of mounting an immune response

⁵⁵ to an immunogen, wherein the antibodies have primate constant and/or variable regions, and wherein the endogenous immunoglobulin-encoding loci are substituted or inactivated. Ref. 57 discloses the use of the Cre/Lox system to modify the immunoglobulin locus in a mammal, such as to replace all or a portion of the constant or variable region to form a modified antibody molecule. Ref. 58 discloses non-human mammalian hosts having inactivated endogenous Ig loci and

functional human lg loci. Ref. 59 discloses methods of making transgenic mice in which the mice lack endogenous heavy claims, and express an exogenous immunoglobulin locus comprising one or more xenogeneic constant regions.

[0080] Using a transgenic animal described above, an immune response can be produced to a PCAV polypeptide, and antibody-producing cells can be removed from the animal and used to produce hybridomas that secrete human monoclonal antibodies. Immunization protocols, adjuvants, and the like are known in the art, and are used in immunization of, for example, a transgenic mouse as described in ref. 60. The monoclonal antibodies can be tested for the ability to inhibit or neutralize the biological activity or physiological effect of the corresponding polypeptide.

D - COMPARISON WITH CONTROL SAMPLES

10

15

5

D.1 - The control

[0081] HML-2 transcripts are up-regulated in tumors, including prostate tumors. To detect such up-regulation, a reference point is needed *i.e.* a control. Analysis of the control sample gives a standard level of RNA and/or protein expression against which a patient sample can be compared.

- **[0082]** A negative control gives a background or basal level of expression against which a patient sample can be compared. Higher levels of expression product relative to a negative control indicate that the patient from whom the sample was taken has, for example, prostate cancer. Typically, for prostate cancer, for example, negative controls would include lifetime baseline levels of expression or the expression level observed in pooled normals. Conversely, equivalent
- 20 levels of expression product indicate that the patient does not have a HML-2-related cancer such as prostate cancer. [0083] A positive control gives a level of expression against which a patient sample can be compared. Equivalent or higher levels of expression product relative to a positive control indicate that the patient from whom the sample was taken has cancer such as prostate cancer. Conversely, lower levels of expression product indicate that the patient does not have a HML-2 related cancer such as prostate cancer.
- ²⁵ **[0084]** For direct or indirect RNA measurement, or for direct polypeptide measurement, a negative control will generally comprise cells which are not from a tumor cell, e.g. a prostate tumor cell. For indirect polypeptide measurement, a negative control will generally be a blood sample from a patient who does not have a prostate tumor. The negative control could be a sample from the same patient as the patient sample, but from a tissue in which HML-2 expression is not up-regulated e.g. a non-tumor non-prostate cell. The negative control could be a prostate cell from the same patient
- 30 as the patient sample, but taken at an earlier stage in the patient's life. The negative control could be a cell from a patient without a prostate tumor. This cell may or may not be a prostate cell. The negative control cell could be a prostate cell from a patient with BPH.

[0085] For direct or indirect RNA measurement, or for direct polypeptide measurement, a positive control will generally comprise cells from a tumor cell *e.g.* a prostate tumor. For indirect polypeptide measurement, a negative control will

³⁵ generally be a blood sample from a patient who has a prostate tumor. The positive control could be a prostate tumor cell from the same patient as the patient sample, but taken at an earlier stage in the patient's life (*e.g.* to monitor remission). The positive control could be a cell from another patient with a prostate tumor. The positive control could be a prostate cell line.

[0086] Other suitable positive and negative controls will be apparent to the skilled person.

40 **[0087]** HML-2 expression in the control can be assessed at the same time as expression in the patient sample. Alternatively, HML-2 expression in the control can be assessed separately (earlier or later).

[0088] Rather than actually compare two samples, however, the control may be an absolute value *i.e.* a level of expression which has been empirically determined from samples taken from prostate tumor patients (*e.g.* under standard conditions).

45

D.2 - Degree of up-regulation

[0089] The up-regulation relative to the control (100%) will usually be at least 150% (*e.g.* 200%, 250%, 300%, 400%, 500%, 600% or more).

50

<u>D.3 - Diagnosis</u>

55

[0090] The invention provides a method for diagnosing prostate cancer. It will be appreciated that "diagnosis" according to the invention can range from a definite clinical diagnosis of disease to an indication that the patient should undergo further testing which may lead to a definite diagnosis. For example, the method of the invention can be used as part of a screening process, with positive samples being subjected to further analysis.

[0091] Furthermore, diagnosis includes monitoring the progress of cancer in a patient already known to have the cancer. Cancer can also be staged by the methods of the invention. The cancer is prostate cancer.

[0092] The efficacy of a treatment regimen (therametrics) of a cancer associated can also monitored by the method of the invention *e.g.* to determine its efficacy.

- **[0093]** Susceptibility to a cancer can also be detected *e.g.* where up-regulation of expression has occurred, but before cancer has developed. Prognostic methods are also encompassed.
- ⁵ **[0094]** All of these techniques fall within the general meaning of "diagnosis" in the present invention.

G- THE HML-2 FAMILY OF HUMAN ENDOGENOUS RETROVIRUSES

[0095] Genomes of all eukaryotes contain multiple copies of sequences related to infectious retroviruses. These endogenous retroviruses have been well studied in mice where both true infectious forms and thousands of defective retrovirus-like elements (e.g. the IAP and Etn sequence families) exist. Some members of the IAP and Etn families are "active" retrotransposons since insertions of these elements have been documented which cause germ line mutations or oncogenic transformation.

- [0096] Endogenous retroviruses were identified in human genomic DNA by their homology to retroviruses of other vertebrates [131, 132]. It is believed that the human genome probably contains numerous copies of endogenous proviral DNAs, but little is known about their function. Most HERV families have relatively few members (1-50) but one family (HERV-H) consists of ~1000 copies per haploid genome distributed on all chromosomes. The large numbers and general transcriptional activity of HERVs in embryonic and tumor cell lines suggest that they could act as disease-causing insertional mutagens or affect adj acent gene expression in a neutral or beneficial way.
- 20 [0097] The K family of human endogenous retroviruses (HERV-K) is well known [133]. It is related to the mouse mammary tumor virus (MMTV) and is present in the genomes of humans, apes and old world monkeys, but several human HERV-K proviruses are unique to humans [134]. The HERV-K family is present at 30-50 full-length copies per haploid human genome and possesses long open reading frames that potentially are translated into viral proteins [135, 136]. Two types of proviral genomes are known, which differ by the presence (type 2) or absence (type 1) of a stretch
- of 292 nucleotides in the overlapping boundary of the pol and env genes [137]. Some members of the HERV-K family are known to code for the gag protein and retroviral particles, which are both detectable in germ cell tumors and derived cell lines [138]. Analysis of the RNA expression pattern of full-length HERV-K has also identified a doubly-spliced RNA that encodes a 105 amino acid protein termed central ORF ('cORF') which is a sequence-specific nuclear RNA export factor that is functionally equivalent to the Rev protein of HIV [139]. HERV-K10 has been shown to encode a full-length
- 30 gag homologous 73 kDa protein and a functional protease [140]. [0098] Patients suffering from germ cell tumors show high antibody titers against HERV-K gag and env proteins at the time of tumor detection [141]. In normal testis and testicular tumors the HERV-K transmembrane envelope protein has been detected both in germ cells and tumor cells, but not in the surrounding tissue. In the case of testicular tumor, correlations between the expression of the env-specific mRNA, the presence of the transmembrane env, cORF and gag
- ³⁵ proteins and antibodies against HERV-K specific peptides in the serum of the patients, have been reported. Reference 142 reports that HERV-K10 gag and/or env proteins are synthesized in seminoma cells and that patients with those tumors exhibit relatively high antibody titers against gag and/or env.

[0099] Gag proteins released in form of particles from HERV-K have been identified in the cell culture supernatant of the teratocarcinoma derived cell line Tera 1. These retrovirus-like particles (termed "human teratocarcinoma derived virus" or HTDV) have been shown to have a 90% sequence homology to the HERV-K10 genome [138, 143].

- **[0100]** While the HERV-K family is present in the genome of every human cell, a high level of expression of mRNAs, proteins and particles is observed only in human teratocarcinoma cell lines [144]. In other tissues and cell lines, only a basal level of expression of mRNA has been demonstrated even using very sensitive methods. The expression of retroviral proviruses is generally regulated by elements of the 5' long terminal repeat (LTR). Furthermore, the activation
- ⁴⁵ of expression of an endogenous retrovirus may trigger the expression of a downstream gene that triggers a neoplastic effect.

[0101] The sequence of HERV-K(II), which locates to chromosome 3, has been disclosed [145].

40

[0102] HML-2 is a subgroup of the HERV-K family [146]. HERV isolates which are members of the HML-2 subgroup include HERV-K10 [137,142], the 27 HML-2 viruses shown in Figure 4 of reference 147, HERV-K(C7) [148], HERV-K

- (II) [145], HERV-K(CH) Table 11 provides a list of all known members of the HML-2 subgroup of the HERV-K family as determined by searching the DoubleTwist database containing all genomic contigs with the sequence AF074086 using the Smith-Waterman algorithm with the default parameters: open gap penalty = -20 and extension penalty = -5.
 [0103] The invention is based on the finding that HML-2 mRNA expression is up-regulated in prostate tumors. Because
 - HML-2 is a well-recognized family, the skilled person will be able to determine without difficulty whether any particular endogenous retroviruses is or is not a HML-2. Preferred members of the HML-2 family for use in accordance with the
- ⁵⁵ endogenous retroviruses is or is not a HML-2. Preferred members of the HML-2 family for use in accordance with the present invention are those whose proviral genome has an LTR which has at least 75% sequence identity to SEQ ID 150 (the LTR sequence from HML-2.HOM [1]). Example LTRs include SEQ IDs 151-154.

H-HERV-K(CH)

[0104] The present invention is based on the discovery of elevated levels of multiple HML-2 polynucleotides in prostate tumor samples as compared to normal prostate tissue. One particular HML-2 whose mRNA was found to be up-regulated is designated herein as 'HERV-K(CH)'.

[0105] Sequences from HERV-K(CH) are shown in SEQ IDs 14-39 and have been deposited with the ATCC (see Table 7). The skilled person will be able to classify any further HERV as HERV-K(CH) or not based on sequence identity to these HERV-K(CH) polynucleotides. Preferably such a comparison is to one or more, or all, of the polynucleotide sequences disclosed herein or of the polynucleotide inserts in the ATCC-deposited isolates. Alternatively, the skilled

¹⁰ artisan can determine the sequence identity based on a comparison to any one or more, or all, of the sequences in SEQ IDs 7-10 and SEQ IDs 14-39 taking into consideration the spontaneous mutation rate associated with retroviral replication. Thus, it will be apparent when the differences in the sequences are consistent with a HERV-K(CH) isolate or consistent with another HERV.

[0106] HERV-K(CH) is therefore a specific member of the HML-2 subgroup which can be used in the invention as described above. It can also be used in methods previously described in relation to HERV-K *e.g.* the diagnosis of testicular cancer [142], autoimmune diseases, multiple sclerosis [149], insulin-dependent diabetes mellitus (IDDM) [150] *etc.*

H.1- HERV-K(CH) Nucleic acids

20 H.1.1-HMV-K(CH) genomic sequences

[0107] The disclosure provides an isolated polynucleotide comprising: (a) the nucleotide sequence of any of SEQ IDs 7-10; (b) the nucleotide sequence of any of SEQ IDs 27-39; (c) the complement of a nucleotide sequence of any of SEQ IDs 7-10; or (d) the complement of the nucleotide sequence of any of SEQ IDs 27-39.

25

5

H.1.2 - HERV-K(CH) fragments

[0108] The disclosure also provides an isolated polynucleotide comprising a fragment of: (a) a nucleotide sequence shown in SEQ IDs 7-10; (b) the nucleotide sequence shown in any of SEQ IDs 27-39; (c) the complement of a nucleotide sequence shown in SEQ IDs 7-10; or (d) the complement of the nucleotide sequence shown in any of SEQ IDs 27-39.
[0109] The fragment is preferably at least x nucleotides in length, wherein x is at least 7 (*e.g.* at least 8, 9, 10,11, 12, 13, 14,15, 16,17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 75, 80, 90,100 *etc.*). The value of *x* may be between about 150 and about 200 or be between about 250 and about 300. The value of *x* may be about 350, about 400, about 450, about 550, about 600, about 650, about 700, or about 750. The value of x may be less than 2000 (*e.g.* less than 1000, 500, 100, or 50).

- [0110] The fragment is preferably neither one of the following sequences nor a fragment of one of the following sequences: (i) the nucleotide sequence shown in SEQ ID 42; (ii) the nucleotide sequence shown in SEQ ID 43; (iii) the nucleotide sequence shown in SEQ ID 43; (iii) the nucleotide sequence shown in SEQ ID 43; (iv) the nucleotide sequence shown in SEQ ID 45; (v) a known polynucleotide; or (vi) a polynucleotide known as of 7th December 2000 (*e.g.* a polynucleotide available in a public database such as GenBank of GeneSeq before 7th December 2000).
- **[0111]** The fragment is preferably a contiguous sequence of one of polynucleotides of (a), (b), (c) or (d) that remains unmasked following application of a masking program for masking low complexity (*e.g.* XBLAST) to the sequence (*i.e.* one would select an unmasked region, as indicated by the polynucleotides outside the poly-n stretches of the masked sequence produced by the masking program).
- ⁴⁵ [0112] These polynucleotides are particularly useful as probes. In general, a probe in which *x*=15 represents sufficient sequence for unique identification. Probes can be used, for example, to determine the presence or absence of a polynucleotide of the invention (or variants thereof) in a sample. By using probes, particularly labeled probes of DNA sequences, one can isolate homologous or related genes. The source of homologous genes can be any species *e.g.* primate species, Particularly human; rodents, such as rats and mice; canines; felines; bovines; ovines; equines; yeast; nematodes; *etc.*
 - **[0113]** Probes from more than one polynucleotide sequence of the disclosure can hybridize with the same nucleic acid if the nucleic acid from which they were derived corresponds to a single sequence (*e.g.* more than one can hybridize to a single cDNA derived from the same mRNA).
- [0114] Preferred fragments (*e.g.* for the identification of HERV-K(CH) polynucleotides associated with cancer) which
 do not correspond identically in their entirety to any portion of the sequence(s) shown in SEQ IDs 42-45 are: SEQ ID
 59 (from gag region), SEQ IDs 60-70 (from pol region) and SEQ IDs 71-82 (from 3' pol region).

[0115] Preferred fragments (*e.g.* for the simultaneous identification of HERV-K(CH) polynucleotides, HERV-KII polynucleotides and/or HERV-K10 polynucleotides) which do correspond identically in their entirety to any portion of the

sequence(s) shown in SEQ IDs 44 & 45 are SEQ IDs 83 & 84 (from gag region).

[0116] Polynucleotide probes unique to HERV-K(CH), HERV-KII and HERV-K10 gag regions are provided in Table 1; polynucleotide probes unique to HERV-K(CH), HERV-KII, and HERV-K10 protease 3' and polymerase 5' regions are provided in Table 2; polynucleotide probes unique to HERV-K(CH), HERV-KII, and HERV-K10 3' pol only regions are provided in Table 3.

5

H.1.3 - HERV-K(CH) fragments plus heterologous sequences

[0117] The disclosure also provides an isolated polynucleotide comprising (a) a segment that is a fragment of the 10 sequence shown in SEQ IDs 7-10 or SEQ IDs 27-39, wherein (i) said fragment is at least 10 nucleotides in length and (ii) corresponds identically in its entirety to a portion of SEQ ID 44 and/or 45; and, optionally, (b) one or more segments flanking the segment defined in (a), wherein the presence of said optional segment(s) causes said polynucleotide to not correspond identically to any portion of a sequence shown in SEQ IDs 7-10 or SEQ IDs 27-39. In some embodiments, the optional flanking segments share less than 40% sequence identity to the nucleic acid sequences shown in SEQ IDs

- 15 7-10, SEQ ID 44 and/or SEQ ID 45. In other embodiments, the optional flanking segments have no contiguous sequence of 10, 12, 15 or 20 nucleotides in common with SEQ IDs 7-10, SEQ ID 44 and/or SEQ ID 45. In yet other embodiments, the optional flanking segment is not present. In further embodiments, a fragment of the polynucleotide sequence is up to at least 30,40,50,60,70,80,90,100,200,300,400, 500, 1000, or 1500 nucleotide in length.
- [0118] The disclosure also provides an isolated polynucleotide having formula 5'-A-B-C-3', wherein: A is a nucleotide 20 sequence consisting of α nucleotides; B is a nucleotide sequence consisting of a fragment of b nucleotides from (i) the nucleotide sequence shown in SEQ IDs 7-10, (ii) the nucleotide sequence shown in any of SEQ IDs 27-39, (iii) the complement of the nucleotide sequence shown in SEQ IDs 7-10, or (iv) the complement of the nucleotide sequence shown in any of SEQ IDs 27-39; C is a nucleotide sequence consisting of c nucleotides; and wherein said polynucleotide is not a fragment of (i) the nucleotide sequence shown in SEQ IDs 7-10, (ii) the nucleotide sequence shown in any of
- 25 SEQ IDs 27-39, (iii) the complement of the nucleotide sequence shown in SEQ IDs 7-10, or (iv) the complement of the nucleotide sequence shown in any of SEQ IDs 27-39. [0119] In this polynucleotide, a+c is at least 1 (e.g. at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 etc.) and b is at least 7 (e.g. at least 8, 9, 10, 11, 12,13, 14, 15, 16,17, 18,19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 etc.). It is preferred
- 30 that the value of *a+b+c* is at least 9 (*e.g.* at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90,100 etc.). It is preferred that the value of a+b+c is at most 200 (e.g. at most 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9). [0120] A and/or C may comprise a promoter sequence (or its complement).

H.14 - Homologous sequences 35

40

45

[0121] Also provided is a polynucleotide having at least s% identity to: (a) SEQ IDs 7-10; (b) a fragment of nucleotides of SEQ IDs 7-10; (c) SEQ IDs 11-13; (b) a fragment of x nucleotides of SEQ IDs 11-13. The value of s is at least 50 (e.g. at least 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 9S, 96, 97, 98, 99, 99.5, 99.9 etc.). The value of x is at least 7 (e.g. 8, 9, 10,11, 12,13, 14,15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 etc.).

[0122] These polynucleotides include naturally-occurring variants (e.g. degenerate variants, allelic variants, etc.), homologs, orthologs, and functional mutants.

[0123] Variants can be identified by hybridization of putative variants with the polynucleotide sequences disclosed in SEQ IDs 14-39 herein, preferably by hybridization under stringent conditions. For example, by using appropriate wash conditions, variants can be identified where the allelic variant exhibits at most about 25-30% base pair (bp) mismatches relative to the selected polynucleotide probe. In general, allelic variants contain 15-25% bp mismatches, and can contain

as little as even 5-15%, or 2-5%, or 1-2% bp mismatches, as well as a single bp mismatch. [0124] The disclosure also encompasses homologs corresponding to any one of the polynucleotide sequences provided herein, where the source of homologous genes can be any mammalian species (e.g. primate species, particularly

- 50 human; rodents, such as rats, etc.). Between mammalian species (e.g. human and primate), homologs generally have substantial sequence similarity (e.g. at least 75% sequence identity, usually at least 90%, more usually at least 95%) between nucleotide sequences. Sequence similarity is calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, domain, etc. A reference sequence will usually be at least about 18 contiguous nt long, more usually at least about 30 nt long, and may extend to the complete
- 55 sequence that is being compared. Algorithms for sequence analysis are known in the art. [0125] A preferred HERV-K(CH) isolate is an isolate sequence which is shown in SEQ IDs 7-10. Another preferred class of HERV-K(CH) isolates are those having a nucleotide sequence identity of at least 90%, preferably at least 95% to the 3' polymerase region shown in SEQ ID 13 which relates to integrase, as measured by the alignment program

GCG Gap (Suite Version 10.1) using the default parameters: open gap = 3 and extend gap = 1. Another preferred class of HERV-K(CH) isolates are those having a nucleotide sequence identity of at least 98%, more preferably at least 99% to the 5' polymerase region shown in SEQ ID 12 which relates to reverse transcriptase, as measured by the alignment program GCG Gap (Suite Version 10.1) using the default parameters: open gap = 3 and extend gap = 1. Another typical

- ⁵ classification of the relationship of retroviruses is based on the amino acid sequence similarities in the reverse transcriptase protein. Thus, an even more preferred class of HERV-K(CH) isolates are those having an amino acid sequence identity of at least 90%, more preferably 95% to the 5' polymerase region encoded by the nucleotide sequence shown in SEQ ID 12, as determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. Thus, these prostate cancer-associated
- ¹⁰ polynucleotide sequences define a class of human endogenous retroviruses, designated herein as HERV-K(CH), whose members comprise variations which, without wanted to be bound by theory, may be due to the presence of polymorphisms or allelic variations.

H.1.5 - HERV-K(H) hybridizable sequences

15

35

55

[0126] Disclosed is an isolated polynucleotide comprising a polynucleotide that selectively hybridizes, relative to a known polynucleotide, to: (a) the nucleotide sequence shown in SEQ IDs 7-10; (b) the nucleotide sequence shown in any of SEQ IDs 27-39; (c) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-10; (d) the complement of the nucleotide sequence shown in SEQ IDs 7-

7-10; (f) a fragment of the nucleotide sequence shown in any of SEQ IDs 27-39; (g) the complement of a fragment of the nucleotide sequence shown in SEQ IDs 7-10; (h) the complement of a fragment of the nucleotide sequence shown in any of SEQ IDs 27-39; (j) a nucleotide sequence shown in SEQ IDs 14-39; or (k) polynucleotides found in ATCC deposits having ATCC accession numbers given in Table 7. The fragment of (e), (f), (g) or (h) is preferably at least x nucleotides in length, wherein x is as defined in H.1.2 above, and is preferably not one of the sequences (i), (ii), (iv),
 (v) or (vi) as defined H.1.2 above.

[0127] Hybridization reactions can be performed under conditions of different "stringency", as described in B.4 above. In some embodiments, the polynucleotide hybridizes under low stringency conditions; in other embodiments it hybridizes under intermediate stringency conditions; in other embodiments, it hybridizes under high stringency conditions.

30 H.1.6 - Deposited HERV-K sequences

[0128] Also disclosed is an isolated polynucleotide comprising: (a) a HERV-K(CH) cDNA insert as deposited at the ATCC and having an ATCC accession number given in Table 7; (b) a HERV-K(CH) sequence as shown in any one of SEQ IDs 14-26; (c) a HERV-K(CH) sequence as shown in any one of SEQ IDs 27-39; or (d) a fragment of (a), (b) or (c). The fragment of (d) is preferably at least x nucleotides in length, wherein x is at least 7 (*e.g.* at least 8, 9, 10, 11, 12, 13,

14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 etc.).

H.1.7 - Preferred HERV-K(CH) sequences

- 40 [0129] Preferred polynucleotides are those having a sequence set forth in any one of the polynucleotide sequences SEQ IDs 7-10 and SEQ IDs 14-39 provided herein; polynucleotides obtained from the biological materials described herein, in particular, polynucleotide sequences present in the isolates deposited with the ATCC and having ATCC accession numbers given in Table 7 or other biological sources (particularly human sources) or by hybridization to the above mentioned sequences under stringent conditions (particularly conditions of high stringency); genes corresponding
- to the provided polynucleotides; variants of the provided polynucleotides and their corresponding genes particularly those variants that retain a biological activity of the encoded gene product (e.g. a biological activity ascribed to a gene product corresponding to the provided polynucleotides as a result of the assignment of the gene product to a protein family(ies) and/or identification of a functional domain present in the gene product). Other polynucleotides and polynucleotide compositions contemplated by and within the scope of the disclosure will be readily apparent to one of ordinary skill in the art when provided with the disclosure here.
- ⁵⁰ skill in the art when provided with the disclosure here.

H.1.8 - General features of polynucleotides

[0130] General features of the polynucleotides described in this section H.1 are the same as those described in section B.4 above.

[0131] The isolated polynucleotides preferably comprise a polynucleotide having a HERV-K(CH) sequence.

[0132] A polynucleotide can encode all or a part of a polypeptide, such as the gag region, 5' pol region or 3' pol region of a human endogenous retrovirus. Double or single stranded fragments can be obtained from the DNA sequence by

chemically synthesizing oligonucleotides in accordance with conventional methods, by restriction enzyme digestion, by PCR amplification, etc.

[0133] Polynucleotides can be cDNAs or genomic DNAs, as well as fragments thereof, particularly fragments that encode a biologically active gene product and/or are useful in the methods disclosed herein (e.g. in diagnosis, as a

- 5 unique identifier of a differentially expressed gene of interest, etc). The term "cDNA" as used herein is intended to include all nucleic acids that share the arrangement of sequence elements found in native mature mRNA species, where sequence elements are exons and 3' and 5' non-coding regions. Normally mRNA species have contiguous exons, with the intervening introns, when present, being removed by nuclear RNA splicing, to create a continuous open reading frame encoding a polypeptide. mRNA species can also exist with both exons and introns, where the introns may be removed
- 10 by alternative splicing. Furthermore it should be noted that different species of mRNAs encoded by the same genomic sequence can exist at varying levels in a cell, and detection of these various levels of mRNA species can be indicative of differential expression of the encoded gene product in the cell. [0134] A genomic sequence of interest comprises the nucleic acid present between the initiation codon and the stop
- codon, as defined in the listed sequences, including all of the introns that are normally present in a native chromosome. 15 It can further include the 3' and 5' untranslated regions found in the mature mRNA. It can further include specific transcriptional and translational regulatory sequences, such as promoters, enhancers, etc., including about 1 kb, but possibly more, of flanking genomic DNA at either the 5' and 3' end of the transcribed region. The genomic DNA can be isolated as a fragment of 100 kbp or smaller; and substantially free of flanking chromosomal sequence. The genomic DNA flanking the coding region, either 3' and 5', or internal regulatory sequences as sometimes found in introns, contains
- 20 sequences required for proper tissue, stage-specific, or disease-state specific expression. [0135] Polynucleotides can be provided as linear molecules or within circular molecules, and can be provided within autonomously replicating molecules (vectors) or within molecules without replication sequences. Expression of the polynucleotides can be regulated by their own or by other regulatory sequences known in the art. The polynucleotides can be introduced into suitable host cells using a variety of techniques available in the art, such as transferrin polycation-
- 25 mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated DNA transfer, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, gene gun, calcium phosphate-mediated transfection, and the like.

[0136] A polynucleotide sequence that is "shown in" or "depicted in" a SEQ ID NO or Figure means that the sequence is present as an identical contiguous sequence in the SEQ ID NO or Figure. The term encompasses portions, or regions of the SEQ ID NO or Figure as well as the entire sequence contained within the SEQ ID NO or Figure.

H.2 - HERV-K(CH) polypeptides

H.2.1- HERV-K(CH) open reading frames

35

30

[0137] Disclosed is an isolated polypeptide: (a) encoded within a HEPV-K(CH) open reading frame; (b) encoded by a polynucleotide shown in SEQ ID 11,12 or 13; or (c) comprising an amino acid sequence as shown in any one of SEQ IDs 46-49, 50-55, 56-57 or 58.

- [0138] Deduced polypeptides encoded by the HERV-K(CH) polynucleotides include the gag translations shown in 40 SEQ IDS 46-49 and the 3' pol translations shown in SEQ IDs 50-55. A polypeptide sequence encoded by the polynucleotide having the sequence shown in SEQ ID 15 is provided in SEQ ID 56; a polypeptide sequence encoded by the polynucleotide having the sequence shown in SEQ ID 14, is shown in SEQ ID 57. A consensus 3' pol polypeptide sequence encoded by the polynucleotides having the sequence shown in SEQ IDs 21-27, inclusive, is provided in SEQ ID 58.
- 45 [0139] The polypeptides encompassed by the present disclosure include those encoded by polynucleotides of the disclosure e.g. SEQ IDs 7-10 and SEQ IDs 14-39, as well as polynucleotides deposited with the ATCC as disclosed herein, as well as nucleic acids that, by virtue of the degeneracy of the genetic code, are not identical in sequence to the disclosed polynucleotides and encode the polypeptides. Thus, the disclosure includes within its scope a polypeptide encoded by a polynucleotide having the sequence of any one of the polynucleotide sequences provided herein, or a

50 variant thereof.

[0140] While the over-expression of the polynucleotides associated with prostate tumor is observed, elevated levels of expression of the polypeptides encoded by these polynucleotides may likely play a role in prostate tumors.

[0141] Typically, in retroviruses, a single large gag polypeptide is synthesized (e.g. a 73 kDa gag protein in HERV-K10) which is subsequently cleaved into multiple functional peptides by a functional protease encoded by the pol or

55 protease region of the genome. Overexpression of sequences corresponding to both gag and pol domains of the HERV-K(CH) suggest such a mechanism. Sequences corresponding to the env and the nuclear RNA transport protein cORF region of the HERV-K(CH) genome may also be overexpressed. The polypeptides encoded by the open reading frames within the over-expressed polynucleotide sequences may play a significant role in the progression of prostate tumors.

[0142] The detection of these polypeptides by antibodies or other reagents that specifically recognize them may aid in the early diagnosis of prostate tumor or any other cancers associated with the overexpression of these HERV-K(CH) sequences.

5 H.2.2 - HERV-K(CH) fragments

[0143] Disclosed is an isolated polypeptide comprising a fragment of: (a) a polypeptide sequence encoded within a HERV-K(CH) open reading frame; (b) a polypeptide sequence encoded by a polynucleotide shown in SEQ ID 11, 12 or 13; or (c) an amino acid sequence as shown in any one of SEQ IDs 46-49, 50-55, 56-57 or 58.

¹⁰ **[0144]** The fragment is preferably at least x amino acids in length, wherein x is at least 5 (*e.g.* at least 6,7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 75, 80, 90, 100, 125, 150, 200, 300, 400, 500 or more *etc.*). The value of x will typically not exceed 1000.

[0145] The fragment may include an epitope *e.g.* an epitope of the amino acid sequence shown in SEQ IDs 56, 57 or 58.

- [0146] SEQ IDs 46-49 provide a translation of the HERV-K(CH) polynucleotides having a sequence shown in SEQ IDs 14, 15, 16 and 40 (the sequence of SEQ ID 40 is from a polynucleotide found in a normal prostate library) corresponding to polynucleotides encoding the gag region. SEQ IDs 50-55 provide a translation of the HERV-K(CH) polynucleotides having a sequence shown in SEQ IDs 21-26, inclusive, corresponding to the 3' region of pol. SEQ IDs 56 & 57 provide translations of the HERV-K(CH) polynucleotide of SEQ ID 15 and SEQ ID 14, respectively. SEQ ID 58 provides a consensus translation of the polynucleotide from the 3' pol region (SEQ IDs 21-26, inclusive). Encompassed with the
- ²⁰ present disclosure are polypeptide fragments, such as, epitopes, of at least 5 amino acids, at least 6 amino acids, at least 8 amino acids, at least 10 amino acids, at least 11 amino acids, at least 12 amino acids, at least 13 amino acids, at least 14 amino acids and at least 15 amino acids of the translations shown in SEQ IDs 46-49 and 50-55. In a preferred embodiment, the HERV-K(CH) epitopes of the amino acid sequence as shown in SEQ IDs 56-58 were determined by the Jameson-Wolf antigenic index
- [0147] The following regions in 3' pol (SEQ ID 58) were determined to be antigenic by Jameson-Wolf algorithm: amino acids: 1-10; 15-35; 45-55; 60-85; 100-115; 125-140; 170-190; 195-215; 230-268. Additional epitope-containing fragments include amino acids 1-8; 2-10; 1-15; 5-15; 7-15; 10-20; 12-20; 15-23; 20-28; 28-35; 15-30; 15-40; 20-30; 45-52; 48-55; 60-68; 60-70; 65-73; 70-78; 75-83; 70-80; 65-75; 68-75; 75-85; 78-85; 65-85; 60-75; 100-108; 103-110; 105-113; 108-115; 125-133; 128-135; 132-140; 170-178; 175-182; 180-187; 182-190; 195-202; 200-208; 205-212; 208-215; 230-237;
- 235-242; 240-247; 245-252; 250-257; 255-262; 260-268; 230-250; 235-255; 240-260; 245-268; 230-245; 235-245; 235-250; 240-255; 245-260; 250-268; 15-55; 170-215; 45-85.
 [0148] The following regions in gag (SEQ ID 56) were determined to be antigenic by Jameson-Wolf algorithm: amino acids: 1-40; 45-60; 80-105; 130-145; 147-183; 186-220; 245-253; 255-288. Additional epitope-containing fragments include amino acids 1-8; 2-10; 1-15; 5-15; 7-15; 10-20; 12-20; 15-23; 20-28; 28-35; 30-37; 33-40; 1-20; 20-40; 1-15;
- ³⁵ 15-30; 15-40; 45-52; 50-57; 55-62; 50-60; 1-60; 80-87; 85-92; 80-90; 90-97; 95-102; 98-105; 85-100; 90-105; 80-100; 85-105; 130-137; 135-142; 140-147; 145-152; 150-157; 155-162; 160-167; 165-172; 170-177; 175-183; 180-187; 185-192; 190-197; 195-202; 200-207; 205-212; 210-217; 213-220; 185-220; 190-220; 195-220; 200-220; 205-220; 255-262; 260-267; 265-272; 270-277; 275-282; 280-288; 245-288; 250-288; 260-288; 265-288; 270-288.
 [0149] The following regions in gag (SEQ ID 57) were determined to be antigenic by Jameson-Wolf algorithm: amino
- acids: 1-40; 80-105; 145-180; 185-225; 240-335. Additional epitope-containing fragments include amino acids 1-8; 2-10; 1-15; 5-15; 7-15; 10-20; 12-20; 15-23; 20-28; 28-35; 30-37; 33-40; 1-20; 20-40; 1-15; 15-30; 15-40; 80-87; 85-92; 80-90; 90-97; 95-102; 98-105; 85-100; 90-105; 80-100; 85-105; 145-152; 150-157; 155-162; 160-167; 165-172; 170-177; 175-182; 180-187; 185-192; 190-197; 195-202; 200-207; 205-212; 210-217; 215-212; 218-225; 145-160; 150-165; 155-170; 160-175; 170-185; 180-225; 185-225; 190-225; 195-225; 200-225; 205-225; 210-225; 215-225; 215-225; 240-247;
- 45 245-252; 250-257; 255-262; 260-267; 265-272; 270-277; 275-282; 280-287; 285-292; 290-297; 295-302; 300-307; 305-312; 310-317; 315-322; 320-327; 325-332; 328-335; 245-285; 250-285; 260-285; 265-285; 270-295; 275-300; 280-305; 285-310; 295-315; 300-320; 305-325; 325-335; 245-335; 250-335; 255-335; 260-335; 270-335; 275-335; 280-335; 285-335; 290-335; 295-335; 305-335; 310-335; 315-335; 320-335.

50 H.2.3 - HERV-K(CH) fragments plus heterologous sequences

[0150] Also disclosed is isolated polypeptide having formula 5'-A-B-C-3', wherein: A is an amino acid sequence consisting of a amino acids; B is an amino acid sequence consisting of a fragment of b amino acids from (i) the amino acid sequence encoded by a polynucleotide shown in SEQ ID 11, 12 or 13; (ii) any one of SEQ IDs 46-49, 50-55, 56-57 or

⁵⁵ 58; C is an amino acid sequence consisting of c amino acids; and wherein said polypeptide is not a fragment of the amino acid sequence defined in (i) or (ii).

[0151] In this polypeptide, a+c is at least 1 (*e.g.* at least 2, 3, 4, 5, 6, 7, 8, 9,10,11, 12,13,14, 15, 16,17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 *etc.*) and b is at least 7 (*e.g.* at least 8, 9, 10,

11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 *etc*.). It is preferred that the value of *a+b+c* is at least 9 (e.g. at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, *100 etc*.). It is preferred that the value of *a+b+c* is at most 200 (*e.g.* at most 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9).

5

10

H.2.4 - Homologous sequences

[0152] Also disclosed is a polypeptide having at least s% identity to: (a) the polypeptide sequences encoded by SEQ IDs 7-45; (b) a fragment of x amino acids of the polypeptide sequences encoded by SEQ IDs 7-45; (c) the polypeptide sequences SEQ IDs 46-58; (d) a fragment of x amino acids of the polypeptide sequences SEQ IDs 46-58. The value of s is at least 35 (e.g. at least 40, 45, 50, 55, 60, 65, 70, 75, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.5, 99.9 *etc.*). The value of x is at least 7 (*e.g.* 8, 9,10,11,12,13, 14, 15, 16,17,18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90,100.

[0153] These polypeptides include naturally-occurring variants (*e.g.* allelic variants, *etc.*), homologs, orthologs, and functional mutants.

[0154] Variants of the naturally-occurring polypeptides, wherein such variants are homologous or substantially similar to the naturally occurring polypeptide, can be of an origin of the same or different species as the naturally occurring polypeptide (*e.g.* human, murine, or some other species that naturally expresses the recited polypeptide, usually a mammalian species). These polypeptide variants are encoded by polynucleotides, and the genetic code can be used

²⁰ to select appropriate codons to construct the corresponding variants.

H.2.5 - Preferred HERV-K(CH) sequences

[0155] Polypeptides, such as those shown in SEQ IDs 46-58, encoded by HERV-K(CH) polynucleotides are differentially expressed in prostate cancer cells. Such polypeptides are referred to herein as "polypeptides associated with prostate cancer" or "HERV-K(CH) polypeptides". The polypeptides can be used to generate antibodies specific for a polypeptide associated with prostate cancer, which antibodies are in turn useful in diagnostic methods, prognostic methods, therametric methods, and the like as discussed in more detail herein. Polypeptides are also useful as targets for therapeutic intervention, as discussed in more detail herein.

30

H.2.6 - General features of polypeptides

[0156] General features of the polypeptides described in this section H.2 are the same as those described in section C.3 above.

- ³⁵ **[0157]** The isolated polypeptides preferably comprise a polypeptide having a HERV-K(CH) sequence.
- **[0158]** Polypeptides, such as polypeptides of the gag regions or polypeptides of the pol regions, encoded by the polynucleotides disclosed herein, such as polynucleotides having the sequences as shown in SEQ IDs 7-10 and SEQ IDs 14-39, and in isolates deposited with the ATCC and having ATCC accession numbers given in Table 7 and/or their corresponding full length genes, can be used to screen peptide libraries to identify binding partners, such as receptors,
- 40 from among the encoded polypeptides. Peptide libraries can be synthesized according to methods known in the art (*e.g.* see refs. 151 & 152).
 101501 In general, the term "polypeptide" as used berein refere to both the full length polypeptide encoded by the

[0159] In general, the term "polypeptide" as used herein refers to both the full length polypeptide encoded by the recited polynucleotide, the polypeptide encoded by the gene represented by the recited polynucleotide, as well as portions or fragments thereof.

⁴⁵ **[0160]** A polypeptide sequence that is "shown in" or "depicted in" a SEQ ID NO or Figure means that the sequence is present as an identical contiguous sequence in the SEQ ID NO or Figure. The term encompasses portions, or regions of the SEQ ID NO or Figure as well as the entire sequence contained within the SEQ ID NO or Figure.

H.3 - Anti-HERV-K(CH) antibodies

50

[0161] The present disclosure also provides isolated antibodies or antigen binding fragments thereof, that bind to a polypeptide disclosed herein. The present disclosure also provides isolated antibodies or antigen binding fragments thereof, that bind to a polypeptide encoded by a polynucleotide disclosure herein. The present disclosure also provides isolated antibodies that bind to a polypeptide, or antigen binding fragment thereof, encoded by a polynucleotide made

by the method comprising the following steps i) immunizing a host animal with a composition comprising said polypeptide, or antigen binding fragment thereof, and ii) collecting cells from said host expressing antibodies against the antigen or antigen binding fragment thereof. Also provided are isolated antibodies that bind to a polypeptide, or antigen binding fragment thereof, encoded by a polynucleotide disclosed herein made by the method comprising the following steps:

providing a cell line producing an antibody, wherein said antibody binds to a polypeptide, or antigen binding fragment thereof, encoded by a polynucleotide disclosed herein and culturing said cell line under conditions wherein said antibodies are produced. In additional embodiments, the antibodies are collected and monoclonal antibodies are produced using the collected host cells or genetic material derived from the collected host cells. In additional embodiments, the antibody

⁵ is a polyclonal antibody. In a further embodiment, the antibody is attached to a solid surface or further comprises a detectable label.

10

[0162] Antibodies, may be isolated antibodies, that bind a polypeptide encoded by a polynucleotide described herein. Antibodies can be provided in a composition comprising the antibody and a buffer and/or a pharmaceutically acceptable excipient. Antibodies specific for a polypeptide associated with cancer are useful in a variety of diagnostic and therapeutic methods, as discussed in detail herein.

- **[0163]** Expression products of a polynucleotide described herein, as well as the corresponding mRNA (particularly mRNAs having distinct secondary and/or tertiary structures), cDNA, or complete gene, or fragments of said expression products can be prepared and used for raising antibodies for experimental, diagnostic, and therapeutic purposes. For polynucleotides to which a corresponding gene has not been assigned, this provides an additional method of identifying
- ¹⁵ the corresponding gene. The polynucleotide or related cDNA is expressed as described above, and antibodies are prepared. These antibodies are specific to an epitope on the polypeptide encoded by the polynucleotide, and can precipitate or bind to the corresponding native polypeptide in a cell or tissue preparation or in a cell-free extract of an *in vitro* expression system.
- [0164] Polyclonal or monoclonal antibodies to the HERV -K(CH) polypeptides or an epitope thereof can be made for use in immunoassays by any of a number of methods know in the art. By epitope reference is made to an antigenic determinant of a polypeptide. The presence of an epitope is demonstrated by the ability of an antibody to bind a polypeptide with specificity. Two antibodies are considered to be directed to the same epitope if they cross block each others binding to the same polypeptide.
- [0165] One approach for preparing antibodies to a polypeptide is the selection and preparation of an amino acid sequence of all or part of the polypeptide, chemically synthesizing the sequence and injecting it into an appropriate animal, typically a rabbit, hamster or a mouse.

[0166] Oligopeptides can be selected as candidates for the production of an antibody to the HERV-K(CH) polypeptide based upon the oligopeptides lying in hydrophilic regions, which are thus likely to be exposed in the mature polypeptide. Additional oligopeptides can be determined using, for example, the Antigenicity Index [30].

- ³⁰ **[0167]** In other embodiments, humanized monoclonal antibodies are provided, wherein the antibodies are specific for HERV-K(CH) polypeptides and do not appreciably bind other HERV polypeptides. The phrase "humanized antibody" refers to an antibody derived from a non-human antibody, typically a mouse monoclonal antibody. Alternatively, a humanized antibody may be derived from a chimeric antibody that retains or substantially retains the antigen-binding properties of the parental, non-human, antibody but which exhibits diminished immunogenicity in humans as compared
- ³⁵ to the parental antibody. The phrase "chimeric antibody," as used herein, refers to an antibody containing sequence derived from two different antibodies (see, *e.g.* ref. 153) which typically originate from different species. Most typically, chimeric antibodies comprise human and murine antibody fragments, generally human constant and mouse variable regions.

[0168] Methods for preparation of the human or primate HERV-K(CH) or an epitope thereof include, but are not limited

- 40 to chemical synthesis, recombinant DNA techniques or isolation from biological samples. Chemical synthesis of a peptide can be performed, for example, by the classical Merrifeld method of solid phase peptide synthesis [154] or the FMOC strategy on a Rapid Automated Multiple Peptide Synthesis system (E.I. du Pont de Nemours Company, Wilmington, DE) [155].
- [0169] Polyclonal antibodies can be prepared by immunizing rabbits or other animals by injecting antigen followed by
 subsequent boosts at appropriate intervals. The animals are bled and sera assayed against purified HERV-K(CH) usually
 by ELISA or by bioassay based upon the ability to block the action of HERV-K(CH). When using avian species, e.g.
 chicken, turkey and the like, the antibody can be isolated from the yolk of the egg. Monoclonal antibodies can be prepared
 after the method of Milstein and Kohler by fusing splenocytes from immunized mice with continuously replicating tumor
 cells such as myeloma or lymphoma cells. [156, 157, 158]. The hybridoma cells so formed are then cloned by limiting
 dilution methods and supernates assayed for antibody production by ELISA, RIA or bioassay.
- **[0170]** The unique ability of antibodies to recognize and specifically bind to target polypeptides provides an approach for treating an overexpression of the polypeptide.

[0171] Specific antibodies, either polyclonal or monoclonal, to the HERV-K(CH) polypeptides can be produced by any suitable method known in the art as discussed above. For example, murine or human monoclonal antibodies can be

⁵⁵ produced by hybridoma technology or, alternatively, the HERV-K(CH) polypeptides, or an immunologically active fragment thereof, or an anti-idiotypic antibody, or fragment thereof can be administered to an animal to elicit the production of antibodies capable of recognizing and binding to the HERV-K(CH) polypeptides. Such antibodies can be from any class of antibodies including, but not limited to IgG, IgA, IgM, IgD, and IgE or in the case of avian species, IgY and from any subclass of antibodies.

40

H.6-HERV-K(CH)-based diagnostic methods

- ⁵ **[0172]** The invention provides methods for diagnosing the presence of prostate cancer in a test sample associated with expression of a polynucleotide in a test cell sample, comprising the steps of: i) detecting a level of expression of at least one polynucleotide of the invention, or a fragment thereof, or at least one polynucleotide found in an isolate selected from the group consisting of ATCC accession numbers given in Table 7, or a fragment thereof; and ii) comparing said level of expression of the polynucleotide in the test sample with a level of expression of polynucleotide in the control cell
- ¹⁰ sample, wherein differential expression of the polynucleotide in the test cell sample relative to the level of polynucleotide expression in the control cell sample is indicative of the presence of cancer in the test cell sample. [0173] In yet other embodiments of the present invention, the detecting is measuring the level of an RNA transcript; measuring the level of a polynucleotide; or measuring by a method including PCR, TMA, bDNA, NAT or Nasba. In further embodiments, the polynucleotide is attached to a solid support.
- ¹⁵ **[0174]** Also disclosed are compositions comprising a test cell sample and an isolated polynucleotide disclosed herein. The present invention further provides methods for detecting prostate cancer associated with expression of a polypeptide in a test cell sample, comprising the steps of: i) detecting a level of expression of at least one polypeptide disclosed herein, or a fragment thereof and ii) comparing said level of expression of the polypeptide in the test sample with a level of expression of polypeptide in the control cell sample, wherein an altered level of expression of the polypeptide in the
- 20 test cell sample relative to the level of expression of the polypeptide in the control cell sample is indicative of the presence of cancer in the test cell sample. The present disclosure also provides methods for detecting prostate cancer associated with the presence of an antibody in a test cell sample, comprising the steps of: i) detecting a level of an antibody and ii) comparing said level of said antibody in the test sample with a level of said antibody in the control cell sample, wherein an altered level of antibody in said test cell sample relative to the level of antibody in the control cell sample is indicative
- of the presence of cancer in the test cell sample.
 [0175] This disclosure also provides methods for detecting prostate cancer associated with elevated levels of HERV-K(CH) polynucleotides, by means of (i) detecting polynucleotides having at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90% at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or at least 100% identity to the polynucleotide shown in SEQ IDs 7-10 or to
- ³⁰ polynucleotides in isolates deposited with the ATCC and having ATCC deposit accession numbers PTA-2561, PTA-2572, PTA-2566, PTA-2571, PTA-2562, PTA-2573, PTA-2560, PTA-2565, PTA-2568, PTA-2564, PTA-2569, PTA-2567, PTA-2559, PTA-2563, PTA-2570, as measured by the alignment program GCG Gap (Suite Version 10.1) using the default parameters: open gap = 3 and extend gap =1 or polynucleotides hybridizing under high stringency conditions to the polynucleotide shown in SEQ IDs 7-10; (ii) detecting polypeptides, or fragments thereof encoded by the sequences
- ³⁵ of (i); and (iii) detecting antibodies specific for one or more of the polypeptides. Furthermore, (iv) detecting particles associated with overexpression of HERV-K(CH) polynucleotides may also be used in the diagnosis of prostate cancer, and monitoring its progression.

[0176] The treatment regimen of a prostate cancer associated with elevated levels of HERV-K(CH) polynucleotides may also monitored by detecting levels of the polynucleotides and polypeptides in order to assess the staging of the cancer and/or efficacy of particular cancer therapies.

- **[0177]** The present invention provides methods of using the polynucleotides described herein for detecting prostate cancer cells, facilitating diagnosis of prostate cancer and the severity of a cancer (*e.g.* tumor grade, tumor burden, and the like) in a subject, facilitating a determination of the prognosis of a subject, and assessing the responsiveness of the subject to therapy (*e.g.* by providing a measure of therapeutic effect through, for example, assessing tumor burden
- ⁴⁵ during or following a chemotherapeutic regimen). Detection can be based on detection of a polynucleotide that is differentially expressed in a prostate cancer cell and/or detection of a polypeptide encoded by a polynucleotide that is differentially expressed in a prostate cancer cell. The detection methods of the invention can be conducted *in vitro* or *in vivo*, on isolated cells, or in whole tissues or a bodily fluid *e.g.* blood, plasma, serum, urine, and the like).
- [0178] The detection methods can be provided as part of a kit. Thus, the disclosure further provides kits for detecting the presence and/or a level of a polynucleotide that is differentially expressed in a cancer cell (*e.g.* by detection of an mRNA encoded by the differentially expressed gene of interest), and/or a polypeptide encoded thereby, in a biological sample. Procedures using these kits can be performed by clinical laboratories, experimental laboratories, medical practitioners, or private individuals. The kits of the disclosure for detecting a polypeptide encoded by a polynucleotide that is differentially expressed in a prostate cancer cell may comprise a moiety that specifically binds the polypeptide, which
- ⁵⁵ may be an antibody that binds the polypeptide or fragment thereof. The kits of the disclosure used for detecting a polynucleotide that is differentially expressed in a prostate cancer cell may comprise a moiety that specifically hybridizes to such a polynucleotide. The kit may optionally provide additional components that are useful in the procedure, including, but not limited to, buffers, developing reagents, labels, reacting surfaces, means for detection, control samples, standards,

instructions, and interpretive information.

[0179] Accordingly, the present disclosure provides kits for detecting prostate cancer comprising at least one of polynucleotides having the sequence as shown in SEQ IDs 7-10, SEQ IDs 14-39, or fragments thereof, or having the sequence found in an isolate deposited with the ATCC and having ATCC accession numbers PTA-2561, PTA-2572, DTA 0552, DTA 05521, DTA 05520, DTA 055200, DTA 05520, DTA 05520, DTA 055200, DTA 05520, DTA 05520, DTA 05520, DT

- ⁵ PTA-2566, PTA-2571, PTA-2562, PTA-2573, PTA-2560, PTA-2565, PTA-2568, PTA-2564, PTA-2569, PTA-2567, PTA-2559, PTA-2563, PTA-2563, PTA-2570 or fragments thereof.
 [0180] In some embodiments, methods are provided for detecting a polypeptide encoded by a gene differentially expressed in a prostate cancer cell. Any of a variety of known methods can be used for detection, including, but not limited to, immunoassay, using antibody that binds the polypeptide, e.g. by enzyme-linked immunosorbent assay (ELISA),
- 10 radioimmunoassay (RIA), and the like; and functional assays for the encoded polypeptide, e.g. binding activity or enzymatic activity.

[0181] As will be readily apparent to the ordinarily skilled artisan upon reading the present specification, the detection methods and other methods described herein can be readily varied. Such variations are within the intended scope of the invention. For example, in the above detection scheme, the probe for use in detection can be immobilized on a solid

¹⁵ support, and the test sample contacted with the immobilized probe. Binding of the test sample to the probe can then be detected in a variety of ways, *e.g.* by detecting a detectable label bound to the test sample to facilitate detected of test sample-immobilized probe complexes.

[0182] The present invention further provides methods for detecting the presence of and/or measuring a level of a polypeptide in a biological sample, which polypeptide is encoded by a polynucleotide that is differentially expressed in a prostate cancer cell, using an antibody specific for the encoded polypeptide. The methods generally comprise: a) contacting the cample with an antibody specific for a polyneptide and by a polynucleotide that is differentially comprise: a)

- contacting the sample with an antibody specific for a polypeptide encoded by a polynucleotide that is differentially expressed in a prostate cancer cell; and b) detecting binding between the antibody and molecules of the sample. [0183] Detection of specific binding of the antibody specific for the encoded prostate cancer-associated polypeptide,
- when compared to a suitable control is an indication that encoded polypeptide is present in the sample. Suitable controls include a sample known not to contain the encoded polypeptide or known not to contain elevated levels of the polypeptide; such as normal prostate tissue, and a sample contacted with an antibody not specific for the encoded polypeptide, e.g. an anti-idiotype antibody. A variety of methods to detect specific antibody-antigen interactions are known in the art and can be used in the method, including, but not limited to, standard immunohistological methods, immunoprecipitation, an enzyme immunoassay, and a radioimmunoassay. In general, the specific antibody will be detectably labeled, either
- ³⁰ directly or indirectly. Direct labels include radioisotopes; enzymes whose products are detectable (e.g. luciferase, ß-galactosidase, and the like); fluorescent labels (e.g. fluorescein isothiocyanate, rhodamine, phycoerythrin, and the like); fluorescence emitting metals, e.g. ¹⁵²Eu, or others of the lanthanide series, attached to the antibody through metal chelating groups such as EDTA; chemiluminescent compounds, e.g. luminol, isoluminol, acridinium salts, and the like; bioluminescent compounds, e.g. luciferin, aequorin (green fluorescent protein), and the like. The antibody may be at-
- ³⁵ tached (coupled) to an insoluble support, such as a polystyrene plate or a bead. Indirect labels include second antibodies specific for antibodies specific for the encoded polypeptide ("first specific antibody"), wherein the second antibody is labeled as described above; and members of specific binding pairs, e.g. biotin-avidin, and the like. The biological sample may be brought into contact with and immobilized on a solid support or carrier, such as nitrocellulose, that is capable of immobilizing cells, cell particles, or soluble proteins. The support may then be washed with suitable buffers, followed by
- 40 contacting with a detectably-labeled first specific antibody. Detection methods are known in the art and will be chosen as appropriate to the signal emitted by the detectable label. Detection is generally accomplished in comparison to suitable controls, and to appropriate standards.

[0184] In some embodiments, the methods are adapted for use *in vivo*, e.g. to locate or identify sites where cancer cells, such as prostate cancer cells, are present.

- ⁴⁵ **[0185]** In some embodiments, methods are provided for detecting a cancer cell by detecting expression in the cell of a transcript that is differentially expressed in a cancer cell. Any of a variety of known methods can be used for detection, including, but not limited to, detection of a transcript by hybridization with a polynucleotide that hybridizes to a polynucleotide that is differentially expressed in a prostate cancer cell; detection of a transcript by a polymerase chain reaction using specific oligonucleotide primers; *in situ* hybridization of a cell using as a probe a polynucleotide that hybridizes to
- ⁵⁰ a gene that is differentially expressed in a prostate cancer cell. The methods can be used to detect and/or measure mRNA levels of a gene that is differentially expressed in a prostate cancer cell. In some embodiments, the methods comprise: a) contacting a sample with a polynucleotide that corresponds to a differentially expressed gene described herein under conditions that allow hybridization; and b) detecting hybridization, if any.
- [0186] Detection of differential hybridization, when compared to a suitable control, is an indication of the presence in the sample of a polynucleotide that is differentially expressed in a cancer cell. Appropriate controls include, for example, a sample which is known not to contain a polynucleotide that is differentially expressed in a cancer cell, and use of a labeled polynucleotide of the same "sense" as the polynucleotide that is differentially expressed in the cancer cell. The cancer cell is a prostate cancer cell. Conditions that allow hybridization are known in the art, and have been described

in more detail above. Detection can also be accomplished by any known method, including, but not limited to, *in situ* hybridization, PCR (polymerase chain reaction), RT-PCR (reverse transcription-PCR), TMA, bDNA, and Nasba and "Northern" or RNA blotting, or combinations of such techniques, using a suitably labeled polynucleotide. A variety of labels and labeling methods for polynucleotides are known in the art and can be used in the assay methods of the invention. Specific hybridization can be determined by comparison to appropriate controls.

- ⁵ invention. Specific hybridization can be determined by comparison to appropriate controls. [0187] Polynucleotide generally comprising at least 10 nt, at least 12nt or at least 15 contiguous nucleotides of a polynucleotide provided herein, such as, for example, those having the sequence as depicted in SEQ IDs 7-10, and 3-28, are used for a variety of purposes, such as probes for detection of and/or measurement of, transcription levels of a polynucleotide that is differentially expressed in a prostate cancer cell. A probe that hybridizes specifically to a polynucleotide.
- ¹⁰ nucleotide disclosed herein should provide a detection signal at least 5-, 10-, or 20-fold higher than the background hybridization provided with other unrelated sequences. It should be noted that "probe" as used herein is meant to refer to a polynucleotide sequence used to detect a differentially expressed gene product in a test sample. As will be readily appreciated by the ordinarily skilled artisan, the probe can be detectably labeled and contacted with, for example, an array comprising immobilized polynucleotides obtained from a test sample (*e.g.* mRNA). Alternatively, the probe can be
- ¹⁵ immobilized on an array and the test sample detectably labeled. These and other variations of the methods of the invention are well within the skill in the art and are within the scope of the invention. [0188] Nucleotide probes are used to detect expression of a gene corresponding to the provided polynucleotide. In Northern blots, mRNA is separated electrophoretically and contacted with a probe. A probe is detected as hybridizing to an mRNA species of a particular size. The amount of hybridization can be quantitated to determine relative amounts
- of expression, for example under a particular condition. Probes are used for in situ hybridization to cells to detect expression. Probes can also be used *in vivo* for diagnostic detection of hybridizing sequences. Probes are typically labeled with a radioactive isotope. Other types of detectable labels can be used such as chromophores, fluorophores, and enzymes. Other examples of nucleotide hybridization assays are described in refs. 185 and 186.
- [0189] PCR is another means for detecting small amounts of target nucleic acids (see, e.g. refs. 187, 188 & 189). Two primer polynucleotides nucleotides that hybridize with the target nucleic acids are used to prime the reaction. The primers can be composed of sequence within or 3' and 5' to the HERV-K(CH) polynucleotides disclosed herein. Alternatively, if the primers are 3' and 5' to these polynucleotides, they need not hybridize to them or the complements. After amplification of the target with a thermostable polymerase, the amplified target nucleic acids can be detected by methods known in the art (e.g. Southern blot). mRNA or cDNA can also be detected by traditional blotting techniques (e.g. Southern blot,
- Northern blot, etc.) described in ref. 8 (e.g. without PCR amplification). In general, mRNA or cDNA generated from mRNA using a polymerase enzyme can be purified and separated using gel electrophoresis, and transferred to a solid support, such as nitrocellulose. The solid support is exposed to a labeled probe, washed to remove any unhybridized probe, and duplexes containing the labeled probe are detected.

[0190] Methods using PCR amplification can be performed on the DNA from a single cell, although it is convenient to

- ³⁵ use at least about 10⁵ cells. The use of the polymerase chain reaction is described in ref. 190, and a review of techniques may be found in pages 14.2 to 14.33 of reference 8. A detectable label may be included in the amplification reaction. Suitable detectable labels include fluorochromes, (*e.g.* fluorescein isothiocyanate (FITC), rhodamine, Texas Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 6-carboxy-X-rhodamine (ROX), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein, 5-carboxyfluorescein (5-FAM), N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA), or 6-car-
- ⁴⁰ boxy-2',4',7',4,7-hexachlorofluorescein (BEX)), radioactive labels, (*e.g.* ³²P, ³⁵S, ³H, *etc.*), and the like. The label may be a two stage system, where the polynucleotides is conjugated to biotin, haptens, *etc.* having a high affinity binding partner, *e.g.* avidin, specific antibodies, *etc.*, where the binding partner is conjugated to a detectable label. The label may be conjugated to one or both of the primers. Alternatively, the pool of nucleotides used in the amplification is labeled, so as to incorporate the label into the amplification product.

⁴⁵ **[0191]** The present invention further relates to methods of detecting/diagnosing a neoplastic or preneoplastic condition in a mammal (for example, a human).

[0192] Examples of conditions that can be detected/diagnosed in accordance with these methods include, but are not limited to prostate cancers. Polynucleotides corresponding to genes that exhibit the appropriate expression pattern can be used to detect prostate cancer in a subject Reference 191 reviews markers of cancer.

- 50 [0193] One detection/diagnostic method comprises: (a) obtaining from a mammal (eg a human) a biological sample, (b) detecting the presence in the sample of a HERV-K(CH) polypeptide and (c) comparing the amount of product present with that in a control sample. In accordance with this method, the presence in the sample of elevated levels of a HERV-K(CH) gene product indicates that the subject has a neoplastic or preneoplastic condition.
- [0194] The compound is preferably a binding protein, *e.g.* an antibody, polyclonal or monoclonal, or antigen binding fragment thereof, which can be labeled with a detectable marker (eg fluorophore, chromophore or isotope, etc). Where appropriate, the compound can be attached to a solid support. Determination of formation of the complex can be effected by contacting the complex with a further compound (eg an antibody) that specifically binds to the first compound (or complex). Like the first compound, the further compound can be attached to a solid support and/or can be labeled with

a detectable marker.

[0195] The identification of elevated levels of HERV-K(CH) polypeptide in accordance with the present invention makes possible the identification of subjects (patients) that are likely to benefit from adjuvant therapy. For example, a biological sample from a post-primary therapy subject (*e.g.* subject having undergone surgery) can be screened for the presence

- ⁵ of circulating HERV-K(CH) polypeptide, the presence of elevated levels of the polypeptide, determined by studies of normal populations, being indicative of residual tumor tissue. Similarly, tissue from the cut site of a surgically removed tumor can be examined (*e.g.* by immunofluorescence), the presence of elevated levels of product (relative to the surrounding tissue) being indicative of incomplete removal of the tumor. The ability to identify such subjects makes it possible to tailor therapy to the needs of the particular subject. Subjects undergoing non-surgical therapy (*e.g.* chemotherapy or
- ¹⁰ radiation therapy) can also be monitored, the presence in samples from such subjects of elevated levels of HERV-K (CH) polypeptide being indicative of the need for continued treatment Staging of the disease (for example, for purposes of optimizing treatment regimens) can also be effected, for example, by prostate biopsy *e.g.* with antibody specific for a HERV-K(CH) polypeptide.
- **[0196]** A kit can be used in the detection of a HERV-K(CH) polypeptide. The kit can comprise a compound that specifically binds a HERV-K(CH) polypeptide, such as, for example, binding proteins including antibodies or binding fragments thereof (e.g. F(ab')₂ fragments) disposed within a container means. The kit can further comprise ancillary reagents, for processing the binding assay.

DEFINITIONS

20

[0197] The term "comprising" means "including" as well as "consisting" *e.g.* a composition "comprising" X may consist exclusively of X or may include something additional *e.g.* X + Y.

[0198] The term "about" in relation to a numerical value x means, for example, $x \pm 10\%$.

[0199] The terms "neoplastic cells", "neoplasia", "tumor", "tumor cells", "cancer" and "cancer cells", (used interchangeably) refer to cells which exhibit relatively autonomous growth, so that they exhibit an aberrant growth phenotype characterized by a significant loss of control of cell proliferation (i.e. de-regulated cell division). Neoplastic cells can be malignant or benign and include prostate cancer derived tissue.

BRIEF DESCRIPTION OF DRAWINGS

30

[0200]

differently.

Figure 1 is a schematic representation of a human endogenous retrovirus with a depiction of the HERV-K(CH) polynucleotides and their position relative to the retrovirus.

35

40

50

Figure 2 is a schematic representation of open reading frames within the HERV-K(HML-2.HOM) (also known as 'ERVK6') genome [1].

Figure 3 shows splicing events described in the prior art [16] for HERV-K mRNAs.

- Figure 4 shows splice sites identified near the 5' and 3' ends of the env ORF. The three reading frames are shaded
- Figure 5 shows northern blot analysis of PCAV transcripts in cancer cell lines. The top arrow on the left shows the position of the genomic mRNA transcript. The next arrow shows the position of the env transcript. The bottom two arrows show the positions of other ORFs. The lanes contain RNA from the following cell lines: (1) Tera 1; (2) DU145; (3) PC3; (4) MDA Pca-2b; (5) LNCaP. Tera 1 is a teratocarcinoma cell line; the others are prostatic carcinoma cell lines.

Figure 6 shows an alignment of env genomic DNA sequences from 27 HERV-K viruses. A consensus sequence (SEQ ID 157) is shown on the bottom line.

Figures 7-9 show alignments of inferred polypeptide sequences for gag (7), pol (8) and env (9) from various HERV-K viruses, together with consensus sequences (SEQ IDs 158-160).

55 MODES FOR CARRYING OUT THE INVENTION

[0201] Certain aspects of the present invention are described in greater detail in the non-limiting examples that follow. The examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description

of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all and only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature, *etc.*) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight,

⁵ molecular weight is weight average molecular weight, temperature is in degrees Celsius, and pressure is at or near atmospheric.

Source of human prostate cell samples and isolation of polynucleotides expressed by them

- 10 [0202] Candidate polynucleotides that may represent genes differentially expressed in cancer were obtained from both publicly-available sources and from cDNA libraries generated from selected cell lines and patient tissues. A normalized cDNA library was prepared from one patient tumor tissue and cloned polynucleotides for spotting on microarrays were isolated from the library. Normal and tumor tissues from 13 patients were processed to generate T7 RNA polymerase transcribed polynucleotides, which were, in turn, assessed for expression in the microarrays. The tissues that served 15 as sources for these libraries and polynucleotides are summarized in Table 4.
- ¹⁵ as sources for these libraries and polynucleotides are summarized in Table 4. [0203] <u>Normalization</u>: The objective of normalization is to generate a cDNA library in which all transcripts expressed in a particular cell type or tissue are equally represented [refs. 192 & 193], and therefore isolation of as few as 30,000 recombinant clones in an optimally normalized library may represent the entire gene expression repertoire of a cell, estimated to number 10,000 per cell. The source materials for generating the normalized prostate libraries were cryop-
- 20 reserved prostate tumor tissue from a patient with Gleason grade 3+3 adenocarcinoma and normal prostate biopsies from a pool of at-risk subjects under medical surveillance. Prostate epithelia were harvested directly from frozen sections of tissue by laser capture microdissection (LCM, Arcturus Engineering Inc., Mountain View, CA), carried out according to methods well known in the art (*e.g.* ref. 194), to provide substantially homogenous cell samples.
 [0204] Total RNA was extracted from LCM-harvested cells using RNeasy™ Protect Kit (Qiagen, Valencia, CA), fol-
- ²⁵ lowing manufacturer's recommended procedures. RNA was quantified using RiboGreen[™] RNA quantification kit (Molecular Probes, Inc. Eugene, OR). One µg of total RNA was reverse transcribed and PCR amplified using SMART[™] PCR cDNA synthesis kit (ClonTech, Palo Alto, CA). The cDNA products were size-selected by agarose gel electrophoresis using standard procedures (ref. 8). The cDNA was extracted using Bio 101 Geneclean® II kit (Qbiogene, Carlsbad, CA). Normalization of the cDNA was carried out using kinetics of hybridization principles: 1.0 µg of cDNA was denatured by
- 30 heat at 100° C for 10 minutes, then incubated at 42°C for 42 hours in the presence of 120 mM NaCl, 10 mM Tris.HCl (pH=8.0), 5 mM EDTA.NA⁺ and 50% formamide. Single-stranded cDNA ("normalized" cDNA) was purified by hydroxya-patite chromatography (#130-0520, BioRad, Hercules, CA) following the manufacturer's recommended procedures, amplified and converted to double-stranded cDNA by three cycles of PCR amplification, and cloned into plasmid vectors using standard procedures (ref 8). All primers/adaptors used in the normalization and cloning process are provided by
- ³⁵ the manufacturer in the SMART[™] PCR cDNA synthesis kit (ClonTech, Palo Alto, CA). Supercompetent cells (XL-2 Blue Ultracompetent Cells, Stratagene, California) were transfected with the normalized cDNA libraries, plated on plated on solid media and grown overnight at 36°C.

[0205] <u>Characterization of normalized libraries</u>: The sequences of 10,000 recombinants per library were analyzed by capillary sequencing using the ABI PRISM 3700 DNA Analyzer (Applied Biosystems, California). To determine the

40 representation of transcripts in a library, BLAST analysis was performed on the clone sequences to assign transcript identity to each isolated clone, i.e. the sequences of the isolated polynucleotides were first masked to eliminate low complexity sequences using the XBLAST masking program (refs. 195, 196 and 197). Generally, masking does not influence the final search results, except to eliminate sequences of relative little interest due to their low complexity, and to eliminate multiple "hits" based on similarity to repetitive regions common to multiple sequences *e.g.* Alu repeats. The

remaining sequences were then used in a BLASTN vs. GenBank search. The sequences were also used as query sequence in a BLASTX vs. NRP (non-redundant proteins) database search.
[0206] Automated sequencing reactions were performed using a Perkin-Elmer PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit containing AmpliTaq DNA Polymerase, FS, according to the manufacturer's directions. The reactions were cycled on a GeneAmp PCR System 9600 as per manufacturer's instructions, except that they were

⁵⁰ annealed at 20° C. or 30° C. for one minute. Sequencing reactions were ethanol precipitated, pellets were resuspended in 8 microliters of loading buffer, 1.5 microliters was loaded on a sequencing gel, and the data was collected by an ABI PRISM 3700 DNA Sequencer. (Applied Biosystems, Foster City, CA). [0207] The number of times a sequence is represented in a library is determined by performing sequence identity

[0207] The number of times a sequence is represented in a library is determined by performing sequence identity analysis on cloned cDNA sequences and assigning transcript identity to each isolated clone. First, each sequence was checked to see if it was a mitochondrial, bacterial or ribosomal contaminant. Such sequences were excluded from the subsequent analysis. Second, sequence artifacts (*e.g.* vector and repetitive elements) were masked and/or removed from each sequence.

[0208] The remaining sequences were compared via BLAST [198] to GenBank and EST databases for gene identifi-

cation and were compared with each other via FastA [199] to calculate the frequency of cDNA appearance in the normalized cDNA library. The sequences were also searched against the GenBank and GeneSeq nucleotide databases using the BLASTN program (BLASTN 1.3MP [198]). Fourth, the sequences were analyzed against a non-redundant protein (NRP) database with the BLASTX program (BLASTX 1.3MP [198]). This protein database is a combination of

- the Swiss-Prot, PIR, and NCBI GenPept protein databases. The BLASTX program was run using the default BLOSUM-62 substitution matrix with the filter parameter: "xnu+seg". The score cutoff utilized was 75.
 [0209] Assembly of overlapping clones into contigs was done using the program Sequencher (Gene Codes Corp.; Ann Arbor, Mich.). The assembled contigs were analyzed using the programs in the GCG package (Genetic Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711) Suite Version 10.1.
- 10 [0210] Summary of polynucleotides described herein: Table 6 provides a summary of polynucleotides isolated as described above and identified as corresponding to a differentially expressed gene (see below). Specifically, Table 6 provides: 1) the HERVK ORF for each clone ID; 2) the clone ID assigned to each sequence; 3) the % patients having the expression ratio of >/= 2X; >/= 2-5X; >/= 5X; and less than 1/2 X; and the Tumor/Normal mRNA Expression Ratio per patient "Pat", eg, patient 93, patient 95, patient 96, etc.
- 15

Detection of elevated levels of cDNA associated with prostate cancer using arrays

[0211] cDNA sequences representing a variety of candidate genes to be screened for differential expression in prostate cancer were assayed by hybridization on polynucleotide arrays. The cDNA sequences included cDNA clones isolated

- from cell lines or tissues as described above. The cDNA sequences analyzed also included polynucleotides comprising sequence overlap with sequences in the Unigene database, and which encode a variety gene products of various origins, functionality, and levels of characterization. cDNAs were spotted onto reflective slides (Amersham) according to methods well known in the art at a density of 9,216 spots per slide representing 4608 sequences (including controls) spotted in duplicate, with approximately 0.8 μl of an approximately 200ng/μl solution of cDNA.
- ²⁵ [0212] PCR products of selected cDNA clones corresponding to the gene products of interest were prepared in a 50% DMSO solution. These PCR products were spotted onto Amersham aluminum microarray slides at a density of 9216 clones per array using a Molecular Dynamics Generation III spotting robot. Clones were spotted in duplicate, for a total of 4608 different sequences per chip.
- [0213] cDNA probes were prepared from total RNA obtained by laser capture microdissection (LCM, Arcturus Engin ³⁰ ering Inc., Mountain View, CA) of tumor tissue samples and normal tissue samples isolated from the patients described above.

[0214] Total RNA was first reverse transcribed into cDNA using a primer containing a T7 RNA polymerase promoter, followed by second strand DNA synthesis. cDNA was then transcribed *in vitro* to produce antisense RNA using the T7 promoter-mediated expression (*e.g.* ref. 200), and the antisense RNA was then converted into cDNA. The second set

- of cDNAs were again transcribed *in vitro*, using the T7 promoter, to provide antisense RNA. This antisense RNA was then fluorescently labeled, or the RNA was again converted into cDNA, allowing for third round of T7-mediated amplification to produce more antisense RNA. Thus the procedure provided for two or three rounds of *in vitro* transcription to produce the final RNA used for fluorescent labeling. Probes were labeled by making fluorescently labeled cDNA from the RNA starting material. Fluorescently-labeled cDNAs prepared from the tumor RNA sample were compared to fluo-
- 40 rescently labeled cDNAs prepared from normal cell RNA sample. For example, the cDNA probes from the normal cells were labeled with Cy3 fluorescent dye (green) and cDNA probes prepared from the tumor cells were labeled with Cy5 fluorescent dye (red).

[0215] The differential expression assay was performed by mixing equal amounts of probes from tumor cells and normal cells of the same patient. The arrays were pre-hybridized by incubation for about 2 hrs at 60°C in 5X SSC/0.2%

- ⁴⁵ SDS/1 mM EDTA, and then washed three times in water and twice in isopropanol. Following pre-hybridization of the array, the probe mixture was then hybridized to the array under conditions of high stringency (overnight at 42°C in 50% formamide, 5X SSC, and 0.2% SDS. After hybridization, the array was washed at 55°C three times as follows: 1) first wash in 1X SSC/0.2% SDS; 2) second wash in 0.1X SSC/0.2% SDS; and 3) third wash in 0.1X SSC.
- [0216] The arrays were then scanned for green and red fluorescence using a Molecular Dynamics Generation III dual color laser-scanner/detector. The images were processed using BioDiscovery Autogene software, and the data from each scan set normalized. The experiment was repeated, this time labeling the two probes with the opposite color in order to perform the assay in both "color directions." Each experiment was sometimes repeated with two more slides (one in each color direction). The data from each scan was normalized, and the level fluorescence for each sequence on the array expressed as a ratio of the geometric mean of 8 replicate spots/genes from the four arrays or 4 replicate spots/gene from 2 arrays or some other permutation.
- ⁵⁵ spots/gene from 2 arrays or some other permutation.
 [0217] Table 6 summarizes the results for gene products differentially expressed in the prostate tumor samples relative to normal cells. The ratio of differential expression is expressed as the normalized hybridization signal associated with the tumor probe divided by the normalized hybridization signal with the normal probe; thus, a ratio greater than 1 indicates

that the gene product is increased in expression in cancerous cells relative to normal cells, while a ratio of less than 1 indicates the opposite. The results from each patient are identified by "Pat" with the corresponding patient identification number. "Concordance" indicates the % of patients in which differential expression of the selected gene product in tumor cells was at least a two-fold different from normal cells.

5 [0218] In at least 79% of prostate patients assayed, 8 out of 10 genes, whose expression was elevated by at least 500%, were represented in HERV-K(CH) sequences.

[0219] Table 6 provides those gene products that were differentially expressed and were classified as gag, 5'-pol (reverse transcriptase) and 3'-pol (integrase) related sequences. It may be possible to examine the function of these gene products in development of cancer and metastasis through use of small molecule inhibitors known to affect the activity of such enzymes.

10

Analysis of the Prostate Cancer Associated Sequences

[0220] In order to determine whether there was homology to any known sequences, the PCR products of 16 different 15 clones from one prostate tumor patient were sequenced. PCR products from these and other clones from the same library were spotted on DNA microarrays. RNA from 13 prostate tumor patients were assayed on the microarrays and then the full inserts of some of the 16 clones were sequenced (Table 6).

[0221] The 16 isolates were initially determined in a first pass sequencing reaction to have the sequences as shown in SEQ IDs 27-39, inclusive. The isolate from the normal prostate tissue was initially determined in a first pass sequencing

- 20 reaction to have the sequence as shown in SEQ ID 41. A first pass sequencing reaction refers to a high-throughput process, where PCR reactions generate the sequencing template then sequencing is performed with one of the PCR primers, in a single direction. A search of public databases revealed that these 16 isolates have some degree of identity to regions of the human endogenous retrovirus HERV-K(II) sequence disclosed in Genbank accession number AB047240 and shown in SEQ ID 44, and also to HERV-K(10), but are nonetheless unique.
- 25 [0222] The isolates were subjected to a second round of nucleic acid sequencing and were found to have the sequences as shown in SEQ IDs 14-26, inclusive. The isolate from the normal prostate tissue was subjected to a second round of nucleic acid sequencing and found to have the sequence as shown in SEQ ID 40. This second round of sequencing is a customized process, where sequencing is performed on purified dsDNA template in a DNA vector. Sequencing is done from both ends of the template, forward and reverse, with primers designed from the flanking regions of the vector, 30
- and new primers are synthesized for every additional reaction needed to span the entire insert. [0223] The Genbank disclosure of HERV-K(II) provides only an incomplete characterization of its genetic features and no association with any disease. The Genbank disclosure characterizes HERV-KII as having a gag gene located at nucleotide 2113-4116 and an env gene located at nucleotide 7437-8174. Detailed analysis of the reported HERV-K(II) sequence indicates that the HERV-K(II) genome includes regions related to gag, protease, 5'-end of pol (reverse tran-
- 35 scriptase) and 3'-end of pol (integrase) domains of a retrovirus. Specifically, the location of the protease gene is from about nucleotide 3917 to about 4920 and the location of the polymerase domain is from about nucleotide 4797 to about 7468.

[0224] Composite HERV-K(CH) polynucleotide sequences are shown in SEQ IDs 7, 8, 9 and 10 and Figure 1 provides a schematic illustration of a human endogenous retrovirus and the HERV-K(CH) species within the schematic illustration.

40 SEQ ID 7 is a composite sequence of the polynucleotides SEQ IDs 14-16, inclusive, and has a consensus sequence as shown in SEQ ID 11. This region corresponds to the gag region of a human endogenous retrovirus. SEQ IDs 8 and 9 are composites sequence of the polynucleotides having a sequence as shown in SEQ IDs 17-20, inclusive, and has a consensus sequence as shown in SEQ ID 12. This region corresponds to the 5' pol region of a human endogenous retrovirus. SEQ ID 10 is a composite sequence of the polynucleotides having a sequence as shown in SEQ IDs 21-26,

45 inclusive, and has a consensus sequence as shown in SEQ ID 13. This region corresponds to the 3' pol region of a human endogenous retrovirus [0225] Homology to HERV-K(II) gag region varied from 87% to 99%. Homology to HERV-K(II) 5'-pol (reverse transcriptase) region varied from 87% to 97%. Homology to HERV-K(II) 3'-pol (integrase) region was approximately 89%. When compared to the human endogenous provirus HERV-K10, the homology of the gag region clones was approxi-

- 50 mately 79%, the 5'-pol region between 81 % and 89% and the 3'-pol region was approximately 89%. Table 5 illustrates the homology of the sequences of the individual clones with the corresponding HERV-K(II) and HERV-K(10) regions. Because the presence of polyA stretches in the HERV-K(CH) sequences (and deposited isolates) may be an artifact of cloning, the % identity shown in Table 5 was determined with alignments performed with polynucleotides excluding the terminal polyA stretch.
- 55 [0226] Consensus polynucleotide sequences SEQ IDs 11-13 were generated with Multiple Sequence Alignment (MSA), a web implementation of the GCG Pileup and Pretty programs. The program uses a clustering algorithm similar to the Clustal program described in reference 201. The default values for the alignments and consensus extraction were 8 for gap open and 2 for gap extension. The poling plurality or minimum number of like sequences specified to assign a

residue to the consensus sequence was 2.

[0227] The polynucleotide sequences shown in SEQ IDs 14-16, inclusive, were used for the consensus polynucleotide sequence shown in SEQ ID 11. The polynucleotide sequences shown in SEQ IDs 17-20, inclusive, were used for the consensus polynucleotide sequence shown in SEQ ID 12. The polynucleotide sequences shown in SEQ IDs 21-26, inclusive, were used for the consensus polynucleotide shown in SEQ ID 13. The "N" represents where there is no

- qualifying minimum representative base. i.e. at least two sequences with the same base at that site. [0228] Northern blotting of prostate cancer cell lines using nucleotides 243-end of SEQ ID 150 labeled as a probe indicates that they express PCAV transcripts of several sizes, corresponding to both full-length viral genomic sequences and to sub-genomic spliced transcripts (Figure 5). Expression of such transcripts have also been observed in teratocar-10 cinoma cell lines [15], as shown in lane 1 of figure 14.

Investigation of other human endogenous retroviruses

[0229] HERV-K(CH) is a member of the HML-2 subgroup of the HERV-K family. HERV-K(H) and HERV-K(10) are 15 also members of this sub-group.

[0230] The same microarray techniques as described above were used to study the expression of members of the HERV-K family in the HML-2 and HML-6 subgroups in prostate tumor tissue. The expression of HERV-H viruses was also studied.

[0231] The results in table 9 show that HERV-H is not up-regulated in prostate tumors. The HML-6 subgroup of HERV-20 K is also not up-regulated. The only endogenous retroviruses that are up-regulated in prostate tumors are in the HML-2 subgroup.

Investigation of tumors other than prostate tumors

25 [0232] HML-2 endogenous retroviruses are up-regulated in prostate tumors. Tumor samples taken from patients with breast and colon cancer were investigated for up-regulation of HML-2 and HML-6 HERV-K viruses using the microarray techniques described above.

[0233] The results in table 10 show that the HML-2 viruses are up-regulated in tissue from prostate tumors, but not from colon or breast tumors. HML-6 expression is not up-regulated in any of the tumors.

30

35

5

Detection of HERV-K(CH) sequences in human prostate cancer cells and tissues.

[0234] DNA from prostate cancer tissue and other human cancer tissues, human colon, normal human tissues including non-cancerous prostate, and from other human cell lines are extracted following the procedure of ref. 202. The DNA is re-suspended in a solution containing 0.05 M Tris HCI buffer, pH 7.8, and 0.1 mM EDTA, and the amount of DNA recovered is determined by microfluorometry using Hoechst 33258 dye [ref. 203].

[0235] Polymerase chain reaction (PCR) is performed using Tag polymerase following the conditions recommended by the manufacturer (Perkin Elmer Cetus) with regard to buffer, Mg²⁺, and nucleotide concentrations. Thermocycling is performed in a DNA cycler by denaturation at 94° C. for 3 min. followed by either 35 or 50 cycles of 94° C. for 1.5 min.,

- 40 50° C. for 2 min. and 72° C. for 3 min. The ability of the PCR to amplify the selected regions of the HERV-K(CH) gene is tested by using a cloned HERV-K(CH) polynucleotide(s) as a positive template(s). Optimal Mg²⁺, primer concentrations and requirements for the different cycling temperatures are determined with these templates. The master mix recommended by the manufacturer is used. To detect possible contamination of the master mix components, reactions without template are routinely tested.
- 45 [0236] Southern blotting and hybridization are performed as described in reference 204, using the cloned sequences labeled by the random primer procedure [205]. Prehybridization and hybridization are performed in a solution containing 6xSSPE, 5% Denhardt's, 0.5% SDS, 50% formamide, 100 μg/ml denaturated salmon testis DNA, incubated for 18 hrs at 42° C., followed by washings with 2xSSC and 0.5% SDS at room temperature and at 37° C. and finally in 0.1xSSC with 0.5% SDS at 68° C. for 30 min (ref. 8). For paraffin-embedded tissue sections the conditions described in ref. 206 50 are followed using primers designed to detect a 250 bp sequence.

Expression of cloned polynucleotides in host cells.

[0237] To study the polypeptide products of HERV-K(CH) cDNA, restriction fragments from the HERV-K(CH) cDNA 55 are cloned into the expression vector pMT2 (pages 16.17-16.22 of ref. 8) and transfected into COS cells grown in DMEM supplemented with 10% FCS. Transfections are performed employing calcium phosphate techniques (pages 16.32-16.40 of ref. 8) and cell lysates are prepared forty-eight hours after transfection from both transfected and untransfected COS cells. Lysates are subjected to analysis by immunoblotting using anti-peptide antibody.

[0238] In immunoblotting experiments, preparation of cell lysates and electrophoresis are performed according to standard procedures. Protein concentration is determined using BioRad protein assay solutions. After semi-dry electrophoretic transfer to nitro-cellulose, the membranes are blocked in 500 mM NaCl, 20 mM Tris, pH 7.5, 0.05% Tween-20 (TTBS) with 5% dry milk. After washing in TTBS and incubation with secondary antibodies (Amersham), enhanced chemiluminescence (ECL) protocols (Amersham) are performed as described by the manufacturer to facilitate detection.

5

Generation of antibodies against polypeptides.

[0239] Polypeptides, unique to HERV-K(CH) are synthesized or isolated from bacterial or other (*e.g.* yeast, baculovirus)
 expression systems and conjugated to rabbit serum albumin (RSA) with m-maleimido benzoic acid N-hydroxysuccinimide ester (MBS) (Pierce, Rockford, III.). Immunization protocols with these peptides are performed according to standard methods. Initially, a pre-bleed of the rabbits is performed prior to immunization. The first immunization includes Freund's complete adjuvant and 500 µg conjugated peptide or 100 µg purified peptide. All subsequent immunizations, performed four weeks after the previous injection, include Freund's incomplete adjuvant with the same amount of protein. Bleeds
 are conducted seven to ten days after the immunizations.

[0240] For affinity purification of the antibodies, the corresponding HERV-K(CH) polypeptide is conjugated to RSA with MBS, and coupled to CNBr-activated Sepharose (Pharmacia, Sweden). Antiserum is diluted 10-fold in 10 mM Tris-HCl, pH 7.5, and incubated overnight with the affinity matrix. After washing, bound antibodies are eluted from the resin with 100 mM glycine, pH 2.5.

20

25

ELISA assay for Detecting HERV-K(CH) Gag and/or Pol related sequences.

[0241] To test blood samples for antibodies that bind specifically to recombinantly produced HERV-K(CH) antigens, the following procedure is employed. After the recombinant HERV-K(CH) pol or gag or env related polypeptides are purified, the recombinant polypeptide is diluted in PBS to a concentration of 5 μg/ml (500 ng/100 μl). 100 microliters of the diluted antigen solution is added to each well of a 96-well Immulon 1 plate (Dynatech Laboratories, Chantilly, Va.), and the plate is then incubated for 1 hour at room temperature, or overnight at 4° C., and washed 3 times with 0.05% Tween 20 in PBS. Blocking to reduce nonspecific binding of antibodies is accomplished by adding to each well 200 μl of a 1% solution of bovine serum albumin in PBS/Tween 20 and incubation for 1 hour. After aspiration of the blocking

- ³⁰ solution, 100 μl of the primary antibody solution (anticoagulated whole blood, plasma, or serum), diluted in the range of 1/16 to 1/2048 in blocking solution, is added and incubated for 1 hour at room temperature or overnight at 4° C. The wells are then washed 3 times, and 100μl goat anti-human IgG antibody conjugated to horseradish peroxidase (organon Teknika, Durham, N.C.), diluted 1/500 or 1/1000 in PBS/Tween 20, 100 μl of o-phenylenediamine dihydrochloride (OPD, Sigma) solution is added to each well and incubated for 5-15 minutes. The OPD solution is prepared by dissolving a 5
- ³⁵ mg OPD tablet in 50 ml 1% methanol in H₂O and adding 50 μ l 30% H₂O₂ immediately before use. The reaction is stopped by adding 25 1 of 4M H₂SO₄ Absorbance are read at 490 nm in a microplate reader (Bio-Rad).

Preparation of vaccines.

- 40 [0242] The present invention also relates to a method of stimulating an immune response against cells that express HERV-K(CH) polypeptides in a patient using HERV-K(CH) gag, and/or pol polypeptides of the invention that acts as an antigen produced by or associated with a malignant cell. This aspect of the invention provides a method of stimulating an immune response in a human against prostate cells or cells that express a HERV-K(CH) pol or gag polynucleotides and polypeptides. The method comprises the step of administering to a human an immunogenic amount of a polypeptide
- ⁴⁵ comprising: (a) the amino acid sequence of a human endogenous retrovirus HERV-K(CH) polypeptide or (b) a mutein or variant of a polypeptide comprising the amino acid sequence of a human endogenous retrovirus HERV-K(CH) polypeptide.

Generation of transgenic animals expressing polypeptides as a means for testing therapeutics.

50

[0243] HERV-K(CH) nucleic acids are used to generate genetically modified non-human animals, or site specific gene modifications thereof, in cell lines, for the study of function or regulation of prostate tumor-related genes, or to create animal models of diseases, including prostate cancer. The term "transgenic" is intended to encompass genetically modified animals having an exogenous HERV-K(CH) gene(s) that is stably transmitted in the host cells where the gene (s) may be altered in sequence to produce a modified polypeptide, or having an exogenous HERV-K(CH) LTR promoter

⁵⁵ (s) may be altered in sequence to produce a modified polypeptide, or having an exogenous HERV-K(CH) LTR promoter operably linked to a reporter gene. Transgenic animals may be made through a nucleic acid construct randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, YACs, and the like. Of interest are transgenic mammals, *e.g.* cows, pigs, goats, horses, *etc.*, and particularly rodents, *e.g.* rats, mice, *etc.*

[0244] The modified cells or animals are useful in the study of HERV-K(CH) gene function and regulation. For example, a series of small deletions and/or substitutions may be made in the HERV-K(CH) gene to determine the role of different domains in prostate tumorigenesis. Specific constructs of interest include, but are not limited to, anti-sense constructs to block HERV-K(CH) gene expression, expression of dominant negative HERV-K(CH) gene mutations, and over-

- ⁵ expression of a HERV-K(CH) gene. Expression of a HERV-K(CH) gene or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development is provided. In addition, by providing expression of polypeptides derived from HERV-K(CH) in cells in which it is otherwise not normally produced, changes in cellular behavior can be induced.
- [0245] DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and negative selection are included. For various techniques for transfecting mammalian cells, see ref. 207.

[0246] For embryonic stem (ES) cells, an ES cell line is employed, or embryonic cells is obtained freshly from a host, e.g. mouse, rat, guinea pig, etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in the presence of appropriate growth factors, such as leukemia inhibiting factor (LIF). When ES cells are transformed, they may be used

- ¹⁵ to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the
- ²⁰ blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting chimeric animals screened for cells bearing the construct. By providing for a different phenotype of the blastocyst and the ES cells, chimeric progeny can be readily detected. [0247] The chimeric animals are screened for the presence of the modified gene and males and females having the
- modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs are maintained as allogeneic or congenic grafts or transplants, or in in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animals, domestic animals, etc. The transgenic animals are used in functional studies, drug screening, etc., e.g. to determine the effect of a candidate drug on prostate cancer, to test potential therapeutics or treatment regimens, etc.

30 Diagnostic Imaging Using HERV-K(CH) Specific Antibodies

[0248] The present invention encompasses the use of antibodies to HERV-K(CH) polypeptides to accurately stage prostate cancer patients at initial presentation and for early detection of metastatic spread of prostate cancer. Radioimmunoscintigraphy using monoclonal antibodies specific for HERV-K(CH) gag or HERV-K(CH) pol or portions thereof or other HERV-K(CH) polypeptides can provide an additional tumor-specific diagnostic test. The monoclonal antibodies of

the instant invention are used for histopathological diagnosis of prostate carcinomas. **[0249]** Subcutaneous human xenografts of prostate cancer cells in nude mice is used to test whether a technetium-99m (^{99m}Tc)-labeled monoclonal antibody of the invention can successfully image the xenografted prostate cancer by external gamma scintography as described for seminoma cells in ref. 208. Each monoclonal antibody specific for a

- 40 HERV-K(CH) polypeptide is purified from ascitic fluid of BALB/c mice bearing hybridoma tumors by affinity chromatography on polypeptide A-Sepharose. Purified antibodies, including control monoclonal antibodies such as an avidinspecific monoclonal antibody [209] are labeled with ^{99m}Tc following reduction, using the methods of refs. 210 and 211. Nude mice bearing human prostate cancer cells are injected intraperitoneally with 200-500 µCi of ^{99m}Tc-labeled antibody. Twenty-four hours after injection, images of the mice are obtained using a Siemens ZLC3700 gamma camera equipped
- ⁴⁵ with a 6 mm pinhole collimator set approximately 8 cm from the animal. To determine monoclonal antibody biodistribution following imaging, the normal organs and tumors are removed, weighed, and the radioactivity of the tissues and a sample of the injectate are measured. Additionally, HERV-K(CH) -specific antibodies conjugated to antitumor compounds are used as prostate cancer-specific chemotherapy.

50 <u>DEPOSITS</u>

35

[0250] The materials listed in Table 7 were deposited with the American Type Culture Collection.

[0251] The foregoing description of preferred embodiments of the invention has been presented by way of illustration and example for purposes of clarity and understanding. It is not intended to be exhaustive or to limit the invention to the precise forms disclosed. It is intended that the scope of the invention be defined by the appended claims and their equivalents.

	<u> </u>	ADLE I - GAG	protease (5) probes	s, isolate spec	
	Isolate	Nucleotides	SEQ ID	Isolate	Nucleotides	SEQ ID
5	K(CH)	1224-1238	161		1490-1510	188
	KII	2098-2114	162		1502-1520	189
		874-890	163		1522-1538	190
10		894-908	164		1561-1576	191
10		910-927	165		1586-1605	192
		927-944	166		1620-1635	193
		989-1004	167		1653-1669	194
15		1019-1036	168		1698-1723	195
		1046-1063	169		1722-1743	196
		1063-1078	170		1748-1762	197
22		1084-1103	171		1773-1788	198
20		1131-1145	172		1820-1834	199
		1148-1163	173		1872-1887	200
		1164-1185	174	K10	1917-1935	201
25	K10	1206-1223	175		1940-1955	202
		1216-1235	176		1955-1969	203
		1243-1260	177		1973-1995	204
30		1258-2375	178		2008-2042	205
30		1277-1295	179		2049-2064	206
		1300-1329	180		2076-2093	207
		1347-1361	181		2097-2113	208
35		1367-1382	182		2122-2139	209
		1392-1410	183		2148-2118	210
		1412-1428	184		2176-2196	211
40		1426-1442	185		2198-2212	212
עד		1445-1461	186		2219-2235	213
		1463-1477	187		2246-2261	214
	<u> </u>			•		

TABLE 1 - GAG protease (5') probes, isolate specific

Isolate	Nucleotides	SEQ ID	Isolate	Nucleotides	SEQ ID
	170-188	215		11-38	113
	205-221	216		37-54	114
	253-268	217		70-90	115
	316-336	218		226-243	116
K(CH) consensus	401-417	219		249-264	117
	490-504	220		308-324	118
	538-552	221		327-342	119
K(CH)	872-886	222		381-397	120
	109-125	223	K10	440-454	121
	1374-1388	224		541-557	122
	1402-1416	225	K10	678-698	123
	140-159	110		722-741	124
KII	410-426	111	-	753-767	125
	1127-1141	112		771-785	126
				854-869	127
				872-890	128
				1195-1209	129
				1308-1323	130
				1335-1349	131
				1349-1365	132

TABLE 2 - Protease (3'seg) Polymerase (5'seq) Probes

Isolate	Nucleotides	SEQ ID
	3-17	133
	25-39	134
	82-104	135
	136-151	136
K(CH) consensus	154-169	137
	189-203	138
	322-337	139
	461-475	140
	630-645	141
	712-727	142
	757-771	143
	818-833	144
KII	1636-1651	145

TABLE 3 - 3' POL probes only

	HERVK ORF	Chiron Clone ID	Source of Clone
5	gag	035JN002.E02	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
	gag	035JN013.H09	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
10	gag	035JN023.F12	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
	gag	037XN001.D10	Normal Prostate Tissue, Pooled from 10 individuals
15	po15'	035JN001.F06	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
	po15'	035JN003.E06	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
20	po15'	035JN013.C11	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
	po15'	035JN013.F03	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
25			
	pro13'	035JN003.G09	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
30	po13'	035JN010.A09	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
-	po13'	035JN015.F06	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
35	po13'	035JN020.B12	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
55	po13'	035JN020.D07	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
	po13'	035JN022.G09	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
40	po13'	035JN015.H02	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3
45	po13'	035JN016.H02	Prostate Cancer Tissue, Patient 101, Gleason Grade 3+3

TABLE 4 - ORFS and sources of initial isolates/clones from prostate cDNA libraries

TABLE 5 - Identity of HERV-K(CH) polynucleotides with HERV-K(II) and HERV-K(10)

			• • • • • •
Clone ID	Region	% Identity HERV-K(II)	% Identity HERV-K(10)
035JN003.G09	3'-pol	89.423	89.423
035JN010.A09	3'-pol	89.663	89.663
035JN015.F06	3'-pol	89.423	89.423
035JN020.B12	3'-pol	89.303	89.303
035JN020.D07	3'-pol	89.614	89.614

50

	(co	ontinued)	
Clone ID	Region	% Identity HERV-K(II)	% Identity HERV-K(10)
035JN022.G09	3'-pol	89.354	89.354
035JN002.E02	gag	99.524	79.881
035JN013.H09	gag	99.017	79.975
035JN023.F12	gag	98.849	79.335
035XN001.D10	gag	87.383	79.947
035JN001.F06	5'-pol	97.211	88.844
035JN003.E06	5'-pol	97.450	86.723
035JN013.C11	5'-pol	97.156	85.444
035JN013.F03	5'-pol	87.962	81.521

TABLE 6

		DNA	V microarra	ay result	DNA microarray results: 13 patients tumor vs. normal prostate, expression of HERV-K RNA	nts tum	Or vs. r	Jormal	orostate	expr€	ssion c	of HER/	/-K RN/					
		Perc	Percent Patient with Expression Ratio	ent with Exp Ratio	ression					u Expre	Tu Expression Ratio mor/Normal mRNA	atio mo	r/Norma	I mRNA				
HERVK ORF	Chiron Clone ID	>=2x	>=2-5x	, ×=2×	<=halfx	Pat 93	Pat 95	Pat 96	Pat 97	Pat 151	Pat 155	Pat 231	Pat 232	Pat 251	Pat 282	Pat 286	Pat 294	Pat 351
gag	035JN002.E02	57.1	42.9	7.1	0.0	4.8	3.0	2.1	1.0	2.3	2.5	1.9	1.7	6.9	1.5	0.6	2.6	2.9
gag	035JN013.H09	78.6	78.6	50.0	0.0	9.3	4.5	5.2	1.4	5.5	13.8	4.2	3.5	31.2	4.5	1.0	12.1	8.6
gag	035JN023.F12	78.6	78.6	57.1	0.0	9.1	4.1	5.1	1.6	5.5	17.0	4.5	3.2	28.2	5.2	1.0	12.7	7.3
qaq	037XN001.D10	64.3	64.3	14.3	0.0	5.4	3.4	2.5	1.5	3.6	4.6	2.9	1.8	10.0	1.7	1.0	3.5	4.3
									<u> </u>	·								
pol5prime	035JN001.F06	42.9	21.4	7.1	0.0	2.0	2.6	1.8	9.5	2.7	1.8	2.0	1.8	7.8	1.2	1.0	1.9	2.3
pol5prime	035JN003.E06	42.9	21.4	7.1	0.0	2.1	2.6	1.8	1.4	2.6	1.9	2.0	1.7	7.7	1.2	1.0	1.8	2.1
pol5prime	035JN013.C11	85.7	78.6	57.1	0.0	6.9	5.6	6.9	2.0	7.4	24.0	4.8	4.3	37.4	4.4	1.0	13.1	8.8
pol5prime	035JN013.F03	85.7	71.4	21.4	0.0	4.6	3.4	3.7	2.2	4.6	8.4	4.1	3.4	21.8	2.3	1.0	5.0	5.8
									[
pol3prime	035JN003.G09	71.4	57.1	7.1	0.0	4.1	3.3	3.3	1.6	4.9	3.3	2.2	3.5	14.9	1.5	1.0	2.5	3.9
pol3prime	035JN010.A09	85.7	78.6	71.4	0.0	8.0	4.4	12.6	2.1	12.4	55.9	5.1	9.5	70.0	5.8	1.0	26.3	9.7
pol3prime	035JN015.F06	85.7	78.6	71.4	0.0	7.6	4.0	12.8	2.2	11.9	53.4	5.1	8.0	69.7	5.9	1.0	25.3	9.1
pol3prime	035JN020.B12	85.7	78.6	64.3	0.0	7.0	4.0	10.5	2.2	11.9	34.9	5.0	6.8	44.5	5.2	1.0	15.2	8.1
pol3orime	035JN020.D07	85.7	78.6	57.1	0.0	6.0	3.2	8.7	2.0	13.7	22.9	4.6	8.6	58.2	3.8	1.0	15.0	7.6
pol3prime	035JN022.G09	78.6	78.6	57.1	0.0	6.6	4.2	6.6	2.0	8.8	12.7	4.5	5.3	28.0	2.6	1.0	5.9	7.8
pol3prime	035JN015.H02	85.7	78.6	57.1	0.0	7.9	4.2	9.0	2.1	10.7	35.3	4.7	7.5	49.5	4.8	1.0	18.2	8.7
pol3prime	035JN018.H02	71.4	71.4	14.3	0.0	3.8	3.0	3.4	1.9	4.3	5.0	3.0	3.1	14.1	1.7	1.0	2.6	5.0

EP 1 415 005 B1

		<u></u>
CMCC = Chiron	an Type Culture Collectior Master Culture Collection le 10th April 2000	
Cell Line	CMCC Accession No.	ATCC Accession No.
035JN003G09	5400	PTA 2561
035JN010A09	5401	PTA 2572
035JN015F06	5402	PTA 2566
035JN015H02	5403	PTA 2571
035JN020B12	5405	PTA 2562
035JN020D07	5406	PTA 2573
035JN022G09	5413	PTA 2560
035JN002E02	5404	PTA 2565
035JN013H09	5408	PTA 2568
035JN023F12	5409	PTA 2564
035XN001D10	5410	PTA 2569
035JN001F06	5411	PTA 2567
035JN003E06	5412	PTA 2559
035JN013C11	5407	PTA 2563
035JN013F03	5415	PTA 2570

TABLE 7 - DEPOSITS

TABLE 8 - Sequence listing

	SEQ ID	DESCRIPTION
35	1	U5 region of herv-k(hml-2.hom) [GenBank AF074086]
	2	U3 region of herv-k(hml-2.hom)
	3	R region of herv-k(hml-2.hom)
40	4	RU5 region of herv-k(hml-2.hom)
40	5	U3R region of herv-k(hml-2.hom)
	6	Non-coding region between U5 and first 5' splice site of herv-k(hml-2.hom)
	7	Composite of three HERV-K(CH) polynucleotides [SEQ IDs 14-16] positioned in the gag region.
45	8 & 9	Composite of four HERV-K(CH) polynucleotides [SEQ IDs 17-20] positioned in the 5' pol region
	10	Composite of six HERV-K(CH) polynucleotides [SEQ IDs 21-26] positioned in the 3' pol region
	11	Consensus sequence of HERV-K(CH) gag region
50	12	Consensus sequence of HERV-K(CH) 5' pol region
00	13	Consensus sequence of HERV-K(CH) 3' pol region
	14	Sequence for clone 035JN002.E02.
	15	Sequence for clone 035JN023.F12.
55	16	Sequence for clone 035JN013.H09.
	17	Sequence for clone 035JN013.C11

(continued)

SEQ ID	DESCRIPTION
18	Sequence for clone 035JN003.E06.
19	Sequence for clone 35JN001.F06.
20	Sequence for clone 035JN013.F03.
21	Sequence for clone 035JN020.D07.
22	Sequence for clone 035JN015.F06.
23	Sequence for clone 035JN003.G09.
24	Sequence for clone 035JN020.B12.
25	Sequence for clone 035JN022.G09.
26	Sequence for clone 035JN010.A09.
27	Sequence for clone 0353N002.E02.
28	Sequence for clone 035JN023.F12.
29	Sequence for clone 035JN013.H09.
30	Sequence for clone 035JN013.C11.
31	Sequence for clone 035JN003.E06.
32	Sequence for clone 035JN001.F06.
33	Sequence for clone 035JN013.F03.
34	Sequence for clone 035JN020.D07.
35	Sequence for clone 035JN015.F06.
36	Sequence for clone 035JN003.G09.
37	Sequence for clone 035JN020.B12.
38	Sequence for clone 035JN022.G09.
39	Sequence for clone 035JN010.A09.
40	Sequence for clone 037XN001.D10 and isolated from normal prostate tissue.
41	Sequence for clone 037XN001.D10 and isolated from normal prostate tissue.
42	EST polynucleotide sequence shown in GenBank accession number Q60732.
43	EST polynucleotide sequence SEQ ID 407 of WO 00/04149
44	Polynucleotide sequence for HERV-KII
45	Polynucleotide sequence for HERV-K10
46-49	Amino acid translations of SEQ IDs 11, 14, 15, 16
50-55	Amino acid translations of SEQ IDs 21-26 (note PSFGK motifs)
56-57	Amino acid translations of SEQ IDs 27 & 28
58	Consensus polypeptide sequence inferred from SEQ IDs 21-26
59-82	Polynucleotide probes not in SEQ IDs 42-45
83 & 84	Polynucleotide probes shared with SEQ IDs 42-45
85	HERV-K108 gag CDS
86	HERV-K108 prt CDS
87	HERV-K108 pol CDS
88	HERV-K108 env CDS

(continued)

	SEQ ID	DESCRIPTION
	89	HERV-K108 cORF 5' CDS
	90	HERV-K108 cORF 3' CDS
	91	HERV-K(C7) gag CDS
	92	HERV-K(C7) gag amino acid sequence
)	93	HERV-K(C7) pol CDS
	94	HERV-K(C7) pol amino acid sequence
	95	HERV-K(C7) env CDS
	96	HERV-K(C7) env amino acid sequence
i	97	BERV-K(II) gag CDS
	98	HERV-K(H) gag amino acid sequence
	99	HERV-K(II) prt CDS
)	100	HERV-K(II) pol CDS
	101	BERV-K(II) env CDS
	102	HERV-K10 gag CDS
	103	HERV-K10 gag(i)
i	104	HERV-K10 gag(ii)
	105	HERV-K10 prt CDS
	106	HERV-K10 prt amino acid sequence
)	107	HERV-K10 pol/env CDS
	108	HERV-K10 pol/env amino acid sequence
	109	CORF amino acid sequence
	110-132	Table 2 probes (cont ^d at SEQ IDs 215-225)
i	133-145	Table 3 probes
	146	HML-2.HOM ('ERVK6') gag amino acid sequence
	147	HML-2.HOM ('ERVK6') prt amino acid sequence
)	148	HML-2.HOM ('ERVK6') pol amino acid sequence
	149	HML-2.HOM ('ERVK6') env amino acid sequence
	150	LTR of herv-k(hml-2.hom)
	151-154	HML-2 LTR sequences
i	155 & 156	herv-k(hml-2.hom) RU5 region (5' and 3' regions, respectively)
	157	Env consensus nucleic acid sequence (Figure 6)
	158	Gag consensus sequence (Figure 7)
)	159	Pol consensus sequence (Figure 8)
	160	Env consensus sequence (Figure 9)
	161-214	Table 1 probes
	215-225	Table 2 probes (cont ^d from SEQ IDs 110-132)
í.	L	

TABLE 9 - Expression of HERV-H and HERV-K in prostate tumors

The "Result" column gives the % of patient samples which showed up-regulation of the GenBank sequence given in first c olur nn in tu r tie . Alativa ta 41. a tu

b.

	the first column in tumor tis	ssue relative to non-tumor tiss	sue.	
5	GenBank ID	HERV	HML Subgroup	Result
	AB047240	К	HML-2	65
10	AF164611	К	HML-2	63
	AF164612	К	HML-2	63
	AF079797	К	HML-6	3
	BC005351	Н	-	0
	XM_054932	Н	-	0

15

TABLE 10 - Expression of HERV-K viruses in colon and breast tumors

The "Result" columns give the % of patient samples which showed up-regulation of the GenBank sequence given in the first column in tumor tissue relative to non-tumor tissue.

20				С.											
20	GenBank ID	HERV	HML Subgroup		Result							Result			
				Prostate	Breast	Colon									
	AB047240	К	HML-2	65	0	2									
25	AF079797	К	HML-6	3	6	0									
	AF164611	К	HML-2	63	0	2									
	AF164612	К	HML-2	63	6	2									

30

35

45

		External of	-5	191	φ φ	, th	ġ.	ιġ	uș lu	ç iç	Ş	Ş	φ	2	2	9	9	ę	9	9	\$	9	φ.		<u>а</u> гч	\$
5		Dpen Cap. B	20	-20	20	50	-20	-50	50	38	8	-20	8	50		197	នុ	-20	-20	-20	-20	-20	-20	-201	20	-20
		Target	20899	37193	143951	101616	172033	26303	65910	30777	129165	182103	123823	110860	15/03/0	100383	187384	147637	55040	183705	33146	164470	21821	138737	47546	6758
10		Target	8	29800	- 1	<u> </u>	164603	18873	62776	23642	122036	174979	116705	103675	141741	03282	179318	140614	48116	176629	25289	146733	13951	131375	415/1	
		Oubry and	-	7428	7428		7428	7428	7428	1	7428	-	-	7428	7428	8	-	-	236	=	7428	7425	-		5981 7428	
15	· .	Ouery Start O	8	-	7428	7428	-	-	╡	7428	-	7428	7428	- 00	1460	7428	7427	7024	7428	7428	-	=†	7428	7428	1773	6704
	:	Pertent	1 1	88	98 98	140	95	36	96	95	94	94	66	88	20	16	8	06	88	8	2	8	8	8	8/	75
20	י י י	Percent P	86	88	98 86	97	96	32	96	8	8	8	8	88	3 4	26	82	98	16	8	79	E	8	82	16	82
		miation 's	7334	7334	7329	2221	7112	7112	7058	7072	7040	7049	6913	6910	E70E	6818	6723	6734	6630	8,8	6276	6166	6184	221-9	5859 5651	2600
25	Family	Matches, 151		7334	7328	2221	7112	7112	7058	7072	7040	7049	6913	6810	670F	6618	6723	6734	6830	6420	6275	6166	6184	617	5859	E600
	ERV-K	Pacora 14	_	3.96-47	5.3E-47	1.3E-45	6.9E-44	5.9E-44	3.4E-44	2.15-44	6.1E-44	4.4E-44	3.4E-42	6E-42	6.3E-40	3E41	2.1E-36	2.6E-42	2.6E-40	1.3E-35	4.1E-31	9.6E-30	1.4E-31	1.5E-31	1.4E-36 A 2E-36	4.5E-28
30	p of H	electron of	1 1	72570	12379	1	1 1		68942		1 1			1	10220	1	1	1	1 1		1	1	1	- 1	57370	11
	subgrou	timer:	102399	102399	102399	102399	102399	102399	102308	10Z399	102389	97618	102399	85887	RECZUI	102309	102399	B134B	102399	102399	102399	102399	102399	102399	102399	102399
35	BLE 11 - HML-2 subgroup of HERV-K Family		119 /chrom=2 Homo	1d=250001 /chrom=8	end=500002 /chrom=12 3180 /and=7873180	310 /chrom=8	536 /end≈861270	403 /chrom=11 [tomo	(end=2266320	10~230000 /clij011-3 278 /end=1248561	/end=656779 /chrom=3	d=649838 (chrom=19	4760 /end=1824759	/end=985557 /chrom=1	Jand-separably (chume)	Jand=5756084	001 /chroni#19	648 /chram=3	nd=214350 /chrom=3	271841 /chrom=3	002 /end=455242	3 /end=33 18083	/and=1000003 /chrom=8	/end=500001 /chrom=8	5 /end=1250005 and=151365 /chimm=7	1081 /chrom=5 Homo
40			/contig_orient=none /start=1 /end#160119 /chrom=2 Homo		/config_orient=forward/start=260002/end=500002/chrom /config_orient=connelement/start=7633180/and=7873180	/control orient=name /start=1/and=166310 /chrom=8	/config_orient=complement /start=455636 /end=8612/0	/contig_orient=none /start=1 /end=167403 /chrom=11 [tomo	/contig_oriant=forward/start=2016320	/contig_orient=complement/start=1/stiu=20000/tdi90it=2 (contigorient=complement/start=999278 /srd=1248551	/contig_orient=forward /start=405779/				control drient-forward (stath-3075000	/config offent=forward /start=5505084	/contro orient⇒none /start=1 /end=250	Icontig orient=rone /start=1 /end=163	/contig_orient=complement /start=1 /end=214350 /chrom=3		/contig_orient=complement /start=250	contig_orient=forward /start=3068083	/contig_orient=forward (start=750003	contig_orient=forward /start=250001	/contig_ortent=forward /start=1000005 /end=1250005 /contig_ortent=commisment /start=1/and=151365 /chmm=7	contig onient=none /start=1 /end=103
45			/contig_o	/contig_o		Т	/contig_o	/contig_o	/contig_c	(contto o	/contig_c	/contig_c	(contig. c	(contig c	T	T	T	Γ	Π		1	1	1	T	T	Π
50		and the state of the second	7428 NT 02228351.2	7428 NT 00738651.3	7428 NT 00950952.3	7428 NT 02390151.2	7428 NT 02582053 2	7426NT 024249S1.2	7428 NT 01161959.5	742BINT 00485855.3	7428 NT 00578653.3	7428 NT 026140S3.3	7428 NT 00933458.3	7428 NT 00440654.3	7420/N1 01119254.3	7428 NT 011512233 3	7428INT 01963851.3	7428 NT 02241161.3	7428NT_005632S1.3	7428 NT 022504S1.3	7428 NT 02339762.3	7428 NT 011620S13 5	٤Ì	7428 NT 01948352.3	7428NT 02403355.3	7428NT 02332351.2
55			¥	N	N4 NA	N4	N4	NA	Z .	2 2	N4	ž	ž	×		- PN		NA	N4	N4	NA	NA	N4	Ň	N N	N4

REFERENCES

[0252]

5 1 Mayer et al. (1999) Nat. Genet. 21 (3), 257-258 (1999) 2 Farrell (1998) RNA Methodologies (Academic Press; ISBN 0-12-249695-7). 3 Yang et al. (1999) Proc Natl Acad Sci USA 96(23):13404-8 4 Robbins et al. (1997) Clin Lab Sci 10(5):265-71. 5 Ylikoski et al. (1999) Clin Chem 45(9):1397-407 10 6 Ylikoski et al. (2001) Biotechniques 30:832-840 7 Shirahata & Pegg (1986) J. Biol. Chem. 261(29):13833-7. 8 Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual. NY, Cold Spring Harbor Laboratory 9 Short protocols in molecular biology (4th edition, 1999) Ausubel et al. eds. ISBN 0-471-32938-X. 10 US patent 5,707,829 15 11 Current Protocols in Molecular Biology (F.M. Ausubel et al., eds., 1987) Supplement 30. 12 EP-B-0509612 13 EP-B-0505012 14 Berkhout et al. (1999) J. Virol. 73:2365-2375. 15 Löwer et al. (1995) J. Virol. 69:141-149. 20 16 Magin et al. (1999) J. Virol. 73:9496-9507. 17 Magin-Lachmann (2001) J. Virol. 75(21):10359-71. 18 Hashido et al. (1992) Biochem. Biophys. Res. Comm. 187:1241-1248. 19 Geysen et al. (1984) PNAS USA 81:3998-4002. 20 Carter (1994) Methods Mol Biol 36:207-23. 25 21 Jameson, BA et al., 1988, CABIOS 4(1):181-186. 22 Raddrizzani & Hammer (2000) Brief Bioinform 1(2):179-89. 23 De Lalla et al. (1999) J. Immunol. 163:1725-29. 24 Brusic et al. (1998) Bioinformatics 14(2):121-30 25 Meister et al. (1995) Vaccine 13(6):581-91. 30 26 Roberts et al. (1996) AIDS Res Hum Retroviruses 12(7):593-610. 27 Maksyutov & Zagrebelnaya (1993) Comput Appl Biosci 9(3):291-7. 28 Feller & de la Cruz (1991) Nature 349(6311):720-1. 29 Hopp (1993) Peptide Research 6:183-190. 30 Welling et al. (1985) FEBS Lett. 188:215-218. 35 31 Davenport et al. (1995) Immunogenetics 42:392-297. 32 Smith and Waterman, Adv. Appl. Math. (1981) 2: 482-489. 33 Go et al, Int. J. Peptide Protein Res. (1980) 15:211 34 Querol et al., Prot. Eng. (1996) 9:265 35 Olsen and Thomsen, J. Gen. Microbiol. (1991) 137:579 40 36 Clarke et al., Biochemistry (1993) 32:4322 37 Wakarchuk et al., Protein Eng. (1994) 7:1379 38 Toma et al., Biochemistry (1991) 30:97 39 Haezerbrouck et al., Protein Eng. (1993) 6:643 40 Masul et al., Appl. Env. Microbiol. (1994) 60:3579 45 41 US patent 4,959,314 42 Breedveld (2000) Lancet 355(9205):735-740. 43 Gorman & Clark (1990) Semin. Immuno/. 2:457-466 44 Jones et al., Nature 321:522-525 (1986) 45 Morrison et al., Proc. Natl. Acad. Sci, US.A., 81:6851-6855 (1984) 50 46 Morrison and Oi, Adv. Immunol., 44:65-92 (1988) 47 Verhoeyer et al., Science 239:1534-1536 (1988) 48 Padlan, Molec. Immun. 28:489-498 (1991) 49 Padlan, Molec. Immunol. 31(3):169-217 (1994). 50 Kettleborough, C.A. et al., Protein Eng. 4(7):773-83 (1991). 55 51 Chothia et al., J. Mol. Biol. 196:901-917 (1987) 52 Kabat et al., U.S. Dept. of Health and Human Services NIH Publication No. 91-3242 (1991) 53 US patent 5,530,101. 54 US patent 5,585,089.

	55 WO 98/24893
	56 WO 91/10741
	57 WO 96/30498
	58 WO 94/02602
5	59 US Patent 5,939,598.
	60 WO 96/33735
	61 WO 93/14778
	62 Findeis et al., Trends Biotechnol. (1993) 11:202
	63 Chiou et al. (1994) Gene Therapeutics: Methods And Applications Of Direct Gene Transfer. ed. Wolff
10	64 Wu et al., J. Biol. Chem. (1988) 263:621
	65 Wu et al., J. Biol. Chem. (1994) 269:542
	66 Zenke et al., Proc. Natl. Acad. Sci. (USA) (1990) 87:3655
	67 Wu et al., J. Biol. Chem. (1991) 266:338
	68 Jolly, Cancer Gene Therapy (1994) 1:51
15	69 Kimura, Human Gene Therapy (1994) 5:845
10	70 Connelly, Human Gene Therapy (1994) 1:185
	71 Kaplitt, Nature Genetics (1994) 6:148
	72 WO 90/07936
20	73 WO 94/03622
20	74 WO 93/25698
	75 WO 93/25234
	76 US patent 5,219,740
	77 WO 93/11230
25	78 WO 93/10218
25	79 US patent 4,777,127
	80 GB Patent No. 2,200,651
	81 EP-A- 0 345 242
	82 WO 91/02805
20	83 WO 94/12649
30	84 WO 93/03769
	85 WO 93/19191
	86 WO 94/28938
	87 WO 95/11984
	88 WO 95/00655
35	89 Curiel, Hum. Gene Ther. (1992) 3:147
	90 Wu, J. Biol. Chem. (1989) 264:16985
	91 US patent 5,814,482
	92 WO 95/07994
	93 WO 96/17072
40	94 WO 95/30763
	95 WO 97/42338
	96 WO 90/11092
	97 US patent 5,580,859
	98 US patent 5,422,120
45	99 WO 95/13796
	100 WO 94/23697
	101 WO 91/14445
	102 EP 0524968
	103 Philip, Mol. Cell Biol. (1994) 14:2411
50	104 Woffendin, Proc. Natl. Acad. Sci. (1994) 91:11581
	105 US patent 5,206,152
	106 WO 92/11033
	107 US patent 5,149,655
	108 US patent 5,206,152
55	109 WO 92/11033
	110 WO90/14837
	111 Vaccine Design - the subunit and adjuvant approach (1995) ed. Powell & Newman
	112 WO00/07621

	113 GB-2220221
	114 EP-A-0689454
	115 EP-A-0835318
	116 EP-A-0735898
5	117 EP-A-0761231
	118 WO99/52549
	119 WO01/21207
	120 WO01/21152
	121 WO00/62800
10	122 WO00/23105
	123 WO99/11241
	124 WO98/57659
	125 WO93/13202.
	126 McSharry (1999) Antiviral Res 43(1):1-21.
15	127 Kuhelj et al. (2001) J Biol Chem 276(20):16674-82.
	128 Schommer et al. (1996) J Gen Virol 77:375-379.
	129 Magin et al. (2000) Virology 274:11-16.
	130 Boese et al. (2001) FEBS Lett 493(2-3):117-21.
	131 Larsson, E., et al., Current Topics in Microbiology and Immunology 148:115 (1989)
20	132 Mariani-Costantini, et al., J. Virol. 63:4982 (1989) and Shih, et al., Virology 182:495 (1991)
	133 Tönjes et al. (1996) J. AIDS Hum. Retrovir. 13(Suppl 1):S261-S267.
	134 Barbulescu et al., Curr. Biol. 9:861 (1999)
	135 Ono, et al., J. Virol. 58:937 (1986)
	136 Löwer et al., Proc. Natl. Acad. Sci USA 90:4480 (1993)
25	137 Ono et al., (1986) J. Virol. 60:589
	138 Boller, et al., Virol. 196:349 (1993)
	139 Yang et al., Proc. Natl. Acad. Sci USA 96:13404 (1999)
	140 Mueller- Lantzsch et al., AIDS Research and Human Retroviruses 9:343-350 (1993)
	141 Herbst et al., Amer. J. Pathol. 149:1727 (1996)
30	142 US patent 5,858,723
	143 Löwer et al., Proc. Natl. Acad. Sci USA 93:5177 (1996)
	144 Löwer et al., Virology 192:501 (1993)
	145 Genbank accession number AB047240
	146 Andersson et al. (1999) J. Gen. Virol. 80:255-260.
35	147 Zsíros et al. (1998) J. Gen. Virol. 79:61-70.
	148 Tönjes et al. (1999) J. Virol. 73:9187-9195.
	149 Johnston et al. (2001) Ann Neurol 50(4):434-42.
	150 Medstrand et al. (1998) J Virol 72(12):9782-7.
	151 US patent 5,010,175
40	152 International patent application WO 91/17823.
	153 US Patent 4,816,567.
	154 Merrifeld, J. Am. Chem. Soc. 85:2149, 1963
	155 Caprino and Han, J. Org. Chem. 37:3404, 1972
	156 Milstein and Kohler, Nature 256:495-497, 1975
45	157 Gulfre and Milstein, Methods in Enzymology: Immunochemical Techniques 73:1-46
	158 Langone and Banatis eds., Academic Press, 1981
	159 Altschul et al. Nucleic Acids Res. (1997) 25:3389-3402
	160 Brutlag et al. Comp. Chem. (1993) 17:203
	161 Schena et al. (1996) Proc Natl Acad Sci U S A. 93(20):10614-9
50	162 Schena et al. (1995) Science 270(5235):467-70
	163 Shalon et al. (1996) Genome Res. 6(7):639-45
	164 US patent 5,807,522
	165 European patent application 0799897
	166 WO 97/29212
55	167 WO 97/27317
	168 European patent application 0785280
	169 WO 97/02357
	170 US patent 5,593,839

	171 US patent 5,578,832
	172 European patent application 0728520
	173 US patent 5,599,695
	174 European patent application 0721016.
5	175 US patent 5,556,752
	176 WO 95/22058
	177 US patent 5,631,734
	178 Pappalarado et al., Sem. Radiation Oncol. (1998) 8:217
	179 Ramsay Nature Biotechnol. (1998) 16:40
10	180 US patent 5,134,854
	181 US patent 5,445,934
	182 WO 95/35505
	183 US patent 5,631,734
	184 US patent 5,800,992
15	185 WO92/02526.
	186 US patent 5,124,246.
	187 Mullis et al., Meth. Enzymol. (1987) 155:335
	188 US patent 4,683,195
	189 US patent 4,683,202
20	190 Saiki et al. (1985) Science 239:487
	191 Hanahan et al. Cell 100:57-70 (2000)
	192 Weissman SM Mol Biol. Med. 4(3),133-143 (1987
	193 Patanjali, et al. Proc. Natl. Acad. Sci. USA 88 (1991)
	194 Simone et al. Am J Pathol. 156(2):445-52 (2000)
25	195 Claverie (1996) Meth. Enzymol. 266:212-227.
	196 Automated DNA Sequencing and Analysis Techniques Adams et al., eds., Chap. 36, p. 267 Academic
	Press, San Diego, 1994
	197 Claverie et al. Comput. Chem. (1993) 17:191
	198 Altschul et. al, J. Mol. Biol., 215:403-410, 1990
30	199 Pearson & Lipman, PNAS, 85:2444, 1988
	200 Luo et al. (1999) Nature Med 5:117-122
	201 Higgins & Sharp CABIOS 5; 151-153 (1989)
	202 Delli Bovi et al. (1986, Cancer Res. 46:6333-6338)
	203 Cesarone, C. et al., Anal Biochem 100:188-197 (1979)
35	204 Southern, E. M., J. Mol. Biol. 98:503-517 (1975)
	205 Feinberg, A. P., et al., 1983, Anal. Biochem. 132:6-13
	206 Wright and Manos (1990, in "PCR Protocols", Innis et al., eds., Academic Press, pp. 153-158)
	207 Keown et al., Methods in Enzymology 185:527-537 (1990)
	208 Marks, et al., Brit. J. Urol. 75:225 (1995)
40	209 Skea, et al., J. Immunol. 151:3557 (1993)
	210 Mather, et al., J. Nucl. Med. 31:692 (1990)
	211 Zhang et al., Nucl. Med. Biol. 19:607 (1992)

SEQUENCE LISTING

⁴⁵ **[0253]**

50

SEQ ID 1:

SEQ ID 2:

5

10

SEQ ID 3:

¹⁵ SEO ID 4:

²⁰ SEO ID 5:

SEQ ID 6:

50

55

 ${\tt tctggtgcccaacgtggaggcttttctctagggtgaggctcgtcgtggtcattgaggacaagtcgagagatcccgagtaggtcatcaggcctaggccaagtcgacgagagatcccgagtaggtcagccttacg$

SEQ ID 7:

GGGAAGAGACTCAAGTAGGAGCGCCTGCCCGAGCTGAGACTAGATGTGAACCTTTCACCATGAAAATGTTAAAAGATATAAAGGAAG 35 GAGTTAAACAATATGGRTCCAACTCCCCTTATATAAGAACAKTATTAGATTCCATTGCYCATGGAAATAGACTTACTCCTTATGACT GGGAAATTTTGGCCAAATCTTCCCCTTTCATCCTCTCAGTATCTACAGTTTAAAACCTGGTGGATTGATGGRGTACARGAACAGGTAC GAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAACAGGTCCAAAATTGGAGCACCATTAACCAAC AATCAGTGATGCAGAATGAGGCTATTGAACAAGTAAGGGCTATTTGCCTCAGGGCCTGGGGAAAAATTCAGGACCCAGGAACAGCTT TCCCTATTAATTCAATTAGACAAGGCTCTAAAGAGCCATATCCTGACTTTGTGGCAAGATTACAAGATGCTGCTCAAAAGTCTATTA CAGATGACAATGCCCGAAAAGTTATTGTAGAATTAATGGCCTATGAAAATGCAAAATGCCAGAATGTCAGTCGGCCATAAAGCCATTAA AAGGAAAAGTTCCAGCAGGAGTTGATGTAATTACMGAATATGTGAAGGCTTGTGATGGGATTGGAGGAGCTATGCATAAGGCAATGC 40 ${\tt TAATGGCTCAAGCAATGAGGGGGGCTCACTCTAGGAGGACAAGTTAGAACATTTGGGAAAAAATGTTATAATTGTGGTCAAATCGGTC$ ATCKGAAAAAGGAGTTGCCCAGGCTTAAAYAARCAGAATATAATAAATCAAGCTATTACAGCAAAAAATAAAAAGCCATCTGGCCTGT GTCCAAAATGTGGAAAAGCAAAACATTGGGCCAATCAATGTCATTCTAAATTTGATAAAGATGGGCAACCATTGTCTGGAAACAGGA TTTATGTTCCACCCAAATGGTCTTTTTACTCCCTGGAAAGCCCCCACAAAAGATTCCTAGAGGGGTATATGGCCCGCTGCCAGAAGG 45 GAGGGTAGGCCTTTGAGGGAGATCAAGTCTAAATTTGAAGGGAGTCCAAATTCATACTGGGGTAATTTACTCAGATTATAAAGGGGG AATTCAGTTAGTGATCAGCTCCACTGTTCCCCCGGAGTGCCAATCCAGGTGATAGAATTGCTCAATTACTGCTTTTGCCTTATGCA SEO ID 8:

SEO ID 9:

5

25

30

45

55

ACAACAATGGCATGCAGAGATTACTATCCCAGCCTCCCTATACAGCCCCCAGGAATCAAAAAATCATGACTAAAATGGGATAGCTCCC TAAAAAGGGACTAGGAAAGAAAGAAGTCCCAATTGAGGCTGAAAAAAATYAAAAAAGGAAAAGGAATAGGGCATCCTTTTTAGGAGCG 10 AAARCTGGAGGCYTTACAYTTATTAGCAAAGAAACMATTAGAAAAAGGACATTGAGCCTTCATTTTCGCCTTGGAATTCTGTTTGTR ATTCAGAAAAAATCCGGCAGATGGCGTATGCTAACTGAGCCATTAATGCCGTAATTCAACCCATGGGGGGCTCTCCCACCCCGGTTGC CCTCTCCAGCCATGGTCCCCTTTAATTATAATTGATCTGAAGGATTGCTTTTTTACCATTCCTCTGGCAAAACAGGATTTTGARAAA TTTGCTTTYACCACCACCCAGCCTAAATAATAAAGAACCAGCCACCAGGTTTCAGTGGAAAGTATTGCCTCAGGGAATGCTTAATAGTT CAACTATTTGTCAGCTCAAGCTCTGCAACCAGTTAGAGACAAGTTTTCAGACTGTTACATCGTTCACTATGTTGATATTTTGTGTGGC 15 ACACATTAAAAGCATTAAATGAGTTTCAAAAGTTGCTAGGAGATACTAATTGGATTTGGAGATATTAATTGGATTTGGCCAACTCTA GGCATTCCTACTTATGCCATGTCAAATTTGTWCTCTTTCTTAAGAGGGGACTCGGAATTAAATAGTGAAAGAACGTTAACTCCAGAG GCAACTAAAGAAATTAAATTAAATTGAAGAAAAAAATTCGGTCAGCACAAGTAAATAGAATAGATCACTTGGCCCCACTCCAAATTTTG AAGACTTTTACATTGTACTTGGATCAAATGGCTACATTAATTGGTCAGGGAAGATTATGAATAACATTGTGTGGAAATGACCCA 20 GATAAAATCACTGTTCCTTTCAACAAGCAACAGGTTAGACAAGCCTTTATCAATTCTGGTGCATGGCAGATTGGTCTTGCCGATTTT

GTGGGAATTATTGACAATCGTTACCACA

SEQ ID 10:

SEQ ID 11:

GGGAAGAGACTCAAGTAGGAGCGCCTGCCCGAGCTGAGACTAGATGTGAACCTTTCACCATGAAAATGTTAAAAGATATAAAGGAAG GAGTTAAACAATATGGATCCAACTCCCCTTATATAAGAACAGTATTAGATTCCATTGCTCATGGAAATAGACTTACTCCTTATGACT GGGAAATTTTGGCCAAATCTTCCCTTTCATCCTCTCAGTATCTACAGTTTAAAACCTGGTGGATTGATGGAGGACAAGAACAGGTAC GnAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAACAGGTCCAAATTGGAGCACCATTAACCAA

35 GnAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAACAGGTCCAAATTGGAGCACCATTAACCAA CAATCAGTGATGCAGAATGAGGCTATTGAACAAGTAAGGGCTATTGCCTCAGGGCCTGGGGAAAAATTCAGGACCCAGGAACAGCT TTCCCTATTAATTCAATTAGACAAGGCTCTAAAGAGCCATATCCTGACTTTGTGGCCAAGATTACAAGATGCTGCTCCAAAAGTCTATT ACAGATGACAATGCCCGAAAAGTTATTGTAGAATAATGGCCTATGAAAATGCAAATCCAGGAGCAGCATTACAAGGCCATTA AAAGGAAAAGTTCCAGGCAGGAGTTGATGTAATTACAGAATATGTGAAGACTTGTGGAAAAATGCCAATGCCAATAGCCAATG CTAATGGCTCAAGCAATGAGGGGGCCTCACTCTAGGAGGACAAGTTAGGAGACAATGTGGAGGAGCTATGCAAAGGCCAATG 40 CATCTGAAAAGGAGTTGCCCAGGCTTGTAAATAAACAGAATATAATAAATCAAGCTATTA

SEQ ID 12:

⁵⁰ SEO ID 13:

CATGCACGTACCTTCATTTGGAAAATTGTCATTTGTCCATGTGACAGTTGATACTTATTCACATTTCATATGGGCAACCTGCCAGAC

AGGAGAAAGTACTTCCCATGTTAAGAGACATTTATTATCTTGTTTTTCCTGTCATGGGAGTTCCAGAAAAAGTTAAAACAGACAATGG GCCAGGTTACTGTAGTAAAGCAGTTCAAAAAATTCTTTAAATCAGTGGAAAATTACACATACAATAGGAATTCTCTATAATTCCCAAGG ACAGGCCATAATTGAAAGAACTAATAGAACACTCAAAGCTCAATTGGTTA

SEQ ID 14:

SEQ ID 15:

GGGAAGAGACTCAAGTAGGAGCGCCCGCCCGAGCTGAGACTAGATGTGAACCTTTCACCATGAAAATGTTAAAAGATATAAAGGAAG GAGTTAAACAATATGGGTCCAACTCCCCTTATATAAGAACATTATTAGATTCCATTGCTCATGGAAATAGACTTACTCCTTATGACT 15 **GGGAAATTTTTGGCCAAATCTTCCCTTTCATCCTCCAGTATCTACAGTTTAAAACCTGGTGGATTGATGGAGTACAAGAACAGGTAC** GAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAACAGGTCCAAATTGGAGCACCATTAACCAAC AATCAGTGATGCAGAATGAGGCTATTGAACAAGTAAGGGCTATTTGCCTCAGGGCCAGGAAAAATTCAGGACCCAGGAACAGCTT TCCCTATTAATTCAATTAGACAAGGCTCTAAAGAGCCATATCCTGACTTTGTGGCAAGATTACAAGATGCTGCTCAAAAGTCTATTA CAGATGACAATGCCCGAAAAAGTTATTGTAGAATTAATGGCCTATGAAAATGCAAATCCAGAATGTCAGTCGGCCATAAAAGCCATTAA AAGGAAAAGTTCCAGCAGGAGTTGATGTAATTACAGAATATGTGAAGGCTTGTGATGGGATTGGAGGAGCTATGCATAAGGCAATGC 20 TAATGGCTCAAGCAATGAGGGGGGCTCACTCTAGGAGGACAAGTTAGAACATTTGGGAAAAAATGTTATAATTGTGGTCAAATCGGTC TACAGAAAATACCACCACTTCAGGGAGTCAGCCAATTACAACAATCCAACAGCTGTCCCGCGCCACAGCAGCAGCAGCAGCAGCAGTAGA TTTATGTTCCACCCAAATGGTCTTTTTACTCCCTGGAAAGCCCCCCACAAAAGATTCCTAGAGGGGTATATGGCCCGCTGCCAGAAGG GAGGGTAGGCCTTTGAGGGAGATCAAGTCTAAATTTGAAGGGAGTCCAAATTCATACTGGGGTAATTTACTCAGATTATAAAGGGGG 25 AATTCAGTTAGTGATCAGCTCCACTGTTCCCCGGAGTGCCAATCCAGGTGATAGAATTGCTCAATTACTGCTTTTGCCTTATGCAAA Адалалалалалалалалалалалалал

SEQ ID 16:

SEO ID 17:

45

30

35

40

Ŧ**U**

50

SEQ ID 18:

AAA

55 SEQ ID 19:

CATTAGAAAAAGGACATTGAGCCTTCATTTTCGCCTTGGAATTCTGTTTGTAATTCAGAAAAAATCCGGCAGATGGCGTATGCTAAC TGAGCCATTAATGCCGTAATTCAACCCATGGGGGGCTCTCCCCACCCCGGTTGCCCTCTCCAGCCATGGTCCCCCTTTAATTATAATTGA

SEQ ID 20:

5

10

15

30

35

40

50

АААААААА

SEQ ID 22:

SEO ID 23:

⁴⁰ SEQ ID 24:

55 SEQ ID 25:

TGATCAGTTAAACCCGCTGTTTAATTTGTTACAACAAAATGTAAGAAAAAGAAATTTCCCATTTTATATTACTCATATTCGAGCACA ${\tt CCAGTGTCAGATTCTACACCTGGCCACTCAGGAGGCAAGAGTTAATCCCAGAGGTCTATGTCCTAATGTGTTATGGCAAATGGATGT$ ${\tt CATGCACGTACCTTCATTTGGAAAAATTGTCATTTGTCCACGTGACAGTTGATACTTATTCACATTTCATATGGGCAACCTGCCAGAC$ GCCAGGTTACTGTAGTAAAGCAGTTCAAAAATTCTTAAATCAGTGGAAAATTACACAATAGGAATTCTCTATAATTCCCAAGG

GCCGCCTCGGCC

SEO ID 26:

5

20

30

45

10 TCAGTCTATTAACATTGTATCAGATTCTGCATATGTAGTACAGGCTACAAAGGATATTGAGAGAGCCCTAATCAAATACATTATGGA TGATCAGTTAAACCCGGCTGTTTAATTTGTTACAACAAAATGTAAGAAAAAGAAATTTCCCATTTTATATTACTCATATTCGAGCACA CACTAATTTACCAGGGCCTTTAACTAAAGCAAATGAACAAGCTGACTTGCTAGTATCATCTGCATTCATGGAAGCACAAGAACTTCA ${\tt CCAGTGTCAGAATTCTACACCTGGCCACTCAGGAGGCAAGAGTTAATCCCAGAGGTCTATGTCCTAATGTGTTATGGCAAATGGATGT$ CATGCACGTACCTTCATTTGGAAAATTGTCATTTGTCCATGTGACAGTTGATACTTATTCACATTTCATATGGGCAACCTGCCAGAC

15 AGGAGAAAGTACTTCCCATGTTAAGAGACATTTATTATCTTGTTTTTCCTGTCATGGGAGTTCCAGAAAAAGTTAAAACAGACAATGG GCCAGGTTACTGTAGTAAAGCAGTTCAAAAATTCTTAAATCAGTGGAAAATTACACATACAATAGGAATTCTCTATAATTCCCAAGG SEO ID 27:

ACCGGCCTTACGGCCGGGGAAGAGNTCAAGTAGGAGCGCCTGCCCGAGCTGAGACTAGATGTGAACCTTTCACCATGAAAATGTTAA AAGATATAAAGGAAGGAGTTAAACAATATGGATCCAACTCCCCTTATATAAGAACAGTATTAGATTCCATTGCTCATGGAAATAGAC TTACTCCTTATGACTGGGAAATTTTGGCCAAATCTTCCCTTTCATCCTCTCAGTATCTACAGTTTAAAAACCTGGTGGATTGATGGAG TACAGGAACAGGTACGAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAACAGGTCCAAATTGGA GCACCATTAACCAACAATCAGTGATGCAGAATGAGGCTATTGAACAAGTAAGGGCTATTTGCCTCAGGGCCTGGGGAAAAATTCAGG ACCCAGGAACAGCTTTCCCTATTAATTCAATTAGACAAGGCTCTAAAGAGCCATATCCTGACTTTGTGGCAAGATTACAAGATGCTG CTCAAAAGTCTATTACAGATGACAATGCCCGAAAAGTTATTGTAGAATTAATGGCCTATGAAAATGCAAATCCAGAATGTCAGTCGG CCATAAAGCCATTAAAAGGAAAAGTTCCAGCAGGAGTTGATGTAATTACAGAATATGTGAAGGCTTGTGATGGGATTGGAGGAGCTA

25 TGCNTAAGGCAATGCTAATGGCTCAAGCAATGAGGGGGGCTCACTCTAGGAGGACAAGTTAGAACATTTGGGAAAAAATGTTT SEO ID 28:

TTACGGCCTTACGGCCGGGGAAGNNTNTCAAGTAGGAGCGCCTGCCCGAGCTGAGACTAGATGTGAACCTTTCACCATGAAAATGTT AAAAGATATAAAGGAAGGAGTTAAACAATATGGGTCCAACTCCCCTTATATAAGAACATTATTAGATTCCATTGCTCATGGAAATAG **ACTTACTCCTTATGACTGGGAAATTTTGGCCAAATCTTCCCTTTCATCCTCTCAGTATCTACAGTTTAAAACCTGGTGGATTGATGG** AGTACAAGAACAGGTACGAAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAACAGGTCCAAATTG GAGCACCATTAACCAACAATCAGTGATGCAGAATGAGGCTATTGAACAAGTAAGGGCTATTTGCCTCAGGGCCTGGGGAAAAATTCA GGACCCAGGAACAGCTTTCCCTATTAATTCAATTAGACAAGGCTCTAAAGAGCCATATCCTGACTTTGTGGCAAGATTACAAGATGC TGCTCAAAAGTCTATTACAGATGACAATGCCCCGAAAAGTTATTGTAGAATTAATGGCCTATGAAAATGCAAATCCAGAATGTCAGTC GGCCATAAAGCCATTAAAAGGAAAAGTTCCAGCAGGAGTTGATGTAATTACAGAATATN

SEQ ID 29:

CGGCCTTACGGCCCGGGGAGANNTCAAGTAGGAGCGCCTGCCCGAGCTGAGACTAGATGTGAACCTTTCACCATGAAAATGTTAAAA 35 GATATAAAGGAAGGAGTTAAACAATATGGATCCAACTCCCCTTATATAAGAACAGTATTAGATTCCATTGCCCATGGAAATAGACTT ACTCCTTATGACTGGGAAATTTTGGCCAAATCTTCCCTTTCATCCTCTCAGTATCTACAGTTTAAAACCTGGTGGATTGATGGGGGTA CAAGAACAGGTACGAAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAACAGGTCCAAATTGGAG ${\tt CACCATTAACCAACAATCAGTGATGCAGAATGAGGGCTATTGAACAAGTAAGGGCTATTTGCCCTCAGGGCCTGGGGAAAAATTCAGGA$ CCCAGGAACAGCTTTCCCTATTAATTCAATTAGACAAGGCTCTAAAGAGCCATATCCTGACTTTGTGGCAAGATTACAAGATGCTGC TCAAAAGTCTATTACAGATGACAATGCCCGAAAAGTTATTGTAGAATTAATGGCCTATGAAAATGCAAAATCCAGAATGTCAGTCGGC 40

CATAAAGCCATTAAAAGGAAAAGTTCCAGCAGGAGTTGATGTAATTACCG

SEQ ID 30:

NCCGGCCTTACGGCCGGGNCAACAATGGCATGCAGAGNTTACTATCCCAGCCTCCCTATACAGCCCCAGGAATCAAAAAATCATGAC CCGNTTCCAAAACAAAAGCTGGAGGCCTTACACTTATTANCAAAGAANCCATTANAAAAAGGACATTGAGCCTTCATTTTCGCCTTG GAATTCTGTTTGTGATTCAAAAAAAATCCGGCANATGGCGTATGCTAACTGANCCATTAATGCCGTAATTCAACCCATGGGGGGCTCT CCCACCCCGGTTGCCCTNTCCAGCCATGGTCCCCTTTAATTAATTGATCTGAAGGATTGCTTTTTTACCATTCCTCTGGCAAAAC

AGGATTTTGAAAAATTTGCTTTTACCACACCAGCCTAAATAATAAANAACCANCCACCAGGTTTCAGTGGAAAGTATTGCCTCAGGG AATGCTTAATAGTTCAACTATTNGTCAGCTCAAGCTCTGCAACCAGTTAGAGACN SEQ ID 31:

- NCCTGGCCTTACGGCCGGGGCTGAAAAAAATCAAAAAAGAAAAGGAATAGGGCATCCTTTTTAGGAGCGGTCACTGTAGAGCCTCCA 50 AAACCCATTCCATTAACTTGGGGGGAAAAAAAAAAACAACTGTATGGTAAATCAGCAGCGCTTCCAAAAACAAAAACTGGAGGCTTTACAT TTATTAGCAAAGAAACAATTAGAAAAAGGACATTGAGCCTTCATTTTCGCCTTGGAATTCTGTTATTCAGAAAAAAATCCCGGCA GATGGCGTATAATGCCGTAATTCAACCCATGGGGGGCTCTCCCACCCGGTTGCCCTCTCCAGCCATGGTCCCCTTTAATTATAATTG ATCTGAAGGATTGCTTTTTACCATTCCTCTGGCAAAACAGGATTTTGAGAAAATTTGCCTTTTACCACACCAGCCTAAATAATAAAAGA ACCAGCCACCAGGTTTCAGTGGAAAGTATTGCCTCAGGGAATGCTTAATAGTTCAACTATTTGTCAGCTCAAGCTCTGCAACCAGTT 55
- ACATTTCTGCAGACAGAGGTTGCCAACGCGGGACTGACAATAACATCTGATAAGATTCAAACCTCTACTCCGTTACTTGGGA

SEQ ID 32:

NNNNNCNCGGGCATTAGAAAAAGGACATTGAGCCTTCATTTTCGCCTTGGAATTCTGTTTGTAATTCAGAAAAAATCCGGCAGATGG CGTATGCTAACTGAGCCATTAATGCCGTAATTCAACCCATGGGGGGCTCTCCCACCCCGGTTGCCCCTCTCCAGCCATGGTCCCCTTTA ATTATAATTGATCTGAAGGATTGCTTTTTTACCATTCCTCTGGCAAAACAGGAATTTTGAAAAAATTTGCTTTTACCACCACCAGCCTAA ATAATAAAGAACCAGCCACCAGGTTTCAGTGGAAAGTATTGCCTCANGGAATGCTTAATAGTTCAACTATTTGTCAGCTCAAAGCTC TGCACCCAGNTAGAGACAAGTTTCAGACTGGTTCATCGTCCTATGTGATAATTTGGTGCTGCAGAACGAGAGAAAAATTTGGCT GTTCACATTTTTGCAGACAGGTTGCCCAACGCGGGACTGACAATAACATCTGATAAGATTAAACCTCTTACTCCTTCCGTACTTGGG AATCCAGGTGGAGAAAGGAAAATTAACCCCCCNNAAAATTGAATTANGAAAAGACCCNTTAAAGCCTTAAATGAGTTCAAAAGTTG CTAGGAGAAACTAATTGGATTTGGAGANATTAATTGGATTTGGCAACTNTAGGCATTCCTACTTATGCCN

SEQ ID 33:

5

20

30

45

- 10
 TCCGGCCTTACGGCCGGGGNTCTTTACCCTGTATAAACATCTTTCTCTCTCCCAGTATTTCTAAGCATGTGACAATGAATATGCAAAGG AAGCGCAGCAGTCCACCAGGTGTGGGATATGTGTGGGCACAATTCAAGACAATGATTAAACCTCCACTTGATGTTGCCAAAAGAGATTT TGAAAAATTTGCTTTCACCACCAGCCTAAATAATAAAGAACCAGCCACCAGGTTTCAGTGGAAAGTATTGCCTCAGGGAATGCTT AATAGTTCAACTATTTGTCAGCCCAGGCCTAAATAATAAAGAACCAGCCACCAGGTTTCAGTGGAAAGTATTGCCTCAGGGAATGCTT TGTGTGCTGCAGAAACGAGAGACAAATTAATTGACCGTTACAACTTTCTGCAGACGAGGGTTGCCAACGCGGGACTGACAATAACAT CTGATAAGATTCAAGCACTCACCCTTTCCGTTACTTGGGAATGCAGGTAGAGGAAAGGAAAATTAAACCACAAAAAAATAGAAATA

 15
 AGAAAAGACACATTAAATGAGTTTCAAAACATCTTTCGGAGATACTAATTGGAGTATTAGGATTTGGCC
 - AACTCTAGGCATTCCTACTTATGCCATGTCAAATTTGTTCTCTTTCT

SEQ ID 34:

25 SEQ ID 35:

SEQ ID 36:

AATCAAATACATTATGGATGATCAGTTAAACCCGGCCGTTTAAATTTGTTACAACAAAATGTAAGAAAAAGAAATTTCCCATTTTATAT TACTCATATTCGAGCACAACATCAGTTAAACCCGGCCTTTAACTAAAGCAAATGGAACAAGCTGACTTGCTAGTATCATCTGCATTCAT GGAGGCACAAGAACTTCATGCCCTGAGCTCATGTAAATGCAATAGGATTAGAAAATAGATTTGATATCACATGGAAACAGACAAAAAA TATTGTACAACATTGCACCCAGTGTCAGATTCTACACCTGGCCACTCAGGAGGCAAGAGTTAATCCCAGAGGTCTATGTCCTAATGT GTTATGGCAAATGGATGTCATGCACGTACCTTCATTTGGAAAATGTCATTGTCCATGTGACACGACGATCATTCACATTTCACATTTGATATCACATTGTCCCAGATGTCATGCCACGTCAATGCGACACTGCCAGAGAGTTACTTCACATTTCACATTTGCACATTTGGGAGATGTCCAGAAAA ATGGGCAACCTGCCCAGACAGGAGAAAGTACTTCCCCATGTTAGAGAGACATTTATCTTGTTTTCCTGGTATGGGAGTTCCAGAAAA AGTTAAAACAGACAATGGGCCANGTTACTGTAGTAGTAAGTCAGTTCAAAATCAGTGGAAAATTACAGTGGAAAATTACACACTN

⁴⁰ SEQ ID 37:

SEQ ID 38:

SEQ ID 39:

TTAACAAGATTTTAATCAGTCTATTAACATTGTATCAGATTCTGCATATGTAGTACAGGCTACAAAGGATATTGAGAGAGCCCCTAAT TCATATTCGAGCACACACTAATTTACCAGGGCCTTTAACTAAAGCAAATGAACAAGCTGACTTGCTAGTATCATCGCATTCATGGA ${\tt TGTACAACATTGCACCCAGTGTCAGATTCTACACCTGGCCACTCAGGAGGCAAGAGTTAATCCCAGAGGTCTATGTCCTAATGTGTT$ ATGGCAAATGGATGTCATGCACGTACCTTCATTTGGAAAATTGTCATTTGTCCATGTGACAGTTGATACTTATTCACATTTCATATG GGCAACCTGCCAGACAGGAGAAAGTACTTCCCATGTTAAGAGACATTTATTATCTTGTTTTCCTGTCATGGGAGTTCCAGAAAAAGT TAAAACAGACAATGGGCCAGGTTACTGGAGTAAAGCAGTTCAAAAATTCTTAAATCAGTGG

SEO ID 40:

5

20

AAGGCAGTCAAGCAGGAGTTAAACAATATGGACCTAACTCTCCTTATATTAGAATATTATTAAATTCCATTGCTCATGGAAATAGAC 10 TTATTTCTTATGATTGGGAAATTCTGGCTATATCTTCCCTTTCACCCTCTCAGTATCTCCAGTTTAAAAACCTGGTGGATTGATGGGG TACAAGAACAGGTACGAAAAAAAACAGGCTACTAATCCTGTTGCTTATATAGATGAAGACCAATTGCTAGGAAGAGGGTCCAAACTGGG ACACTATTAACCAACAATCAGTAATGAAAAATGAGGCTATTGAACAACTATAAGGGCTATTTGCCTCAGGGCCTGGGAAAACATTCAG GACCCAGGAACCTCATGCCCTTCTTTTAGTTCAATCAGACAAGGCTCTAAAGAGCCATATCCAGACTTTGTGGCAAGGTTGCAAGAT GCAGCTCAAAAATCCATTGCAGGTAACGCCCGAAAAGTTATTGTAGAAATAATGGCTTATCAAAAACGCAAATTCAGAGTGTCAATCA GCCATAAAGCCATTAAGAGGAAATGTTTCAGCAGGAGTTGATGTAATTACAGAATATGTGAAGGCTTGTGATGGGGATTGGAGGAGCT 15

ATGCATAAGGCAATGCCATTGGCTCAAGCAATTACAGGGGTTGCTATAGGAGGACAAGTTAAAACATTTGGGGGAAAATGTTATAAT **SEO ID 41:**

NCCGGCCTTACGGCCGGGAAAGGCAGTCAAGCAGGAGTTAAACAATATGGACCTAACTCTCCTTATATTAGAATATTAAAATTCC ATTGCTCATGGAAATAGACTTATTTCTTATGGATTGGGAAATTCTGGCTATATCTTCCCCTTTCACCCTCTCAGTATCTCCCAGTTTAAA ACCTGGTGGATTGATGGGGTACAAGAACAGGTACCGAAAAAATCAGGCTACTAATCCTGTTGCTTATATAGATGAAGACCAATTGCT AGGAAGAGGTCCAAACTGGGACACTATTAACCAACAATCAGTAATGAAAATGAGGCTATTGAACAACTATAAGGGCTATTTGCCTCA GGGCCTGGGAAAACATTCAGGACCCAGGAACCTCATGCCCTTCTTTAGTTCAATCAGACAAGGCTCTAAAGAGCCATATCCAGACT TTGTGGCAAGGTTGCAAGATGCAGCTCAAAAATCCATTGCAGGTAACGCCCGAAAAGTTATTGTAGAAATAATGGCTTATCAAAACG CAAATTCAGAGTGTCAATCAGCCATAAAGCCATTAAGAGGAAATGTTTCAGCAGGAGTTGATGTAATTACAGAATATGTGAAGGCTT GTGATGGGATTGGAGGAGCTATGCATAAGGCAATGCCATTGGCTCAAGCAATTACAGGGGT

SEO ID 42:

25 AAAGGCAGTCAAGCAGGAGTTAAACAATATGGACCTAACTCTCCTTATATGAGAACATTATTAAATTCCATTGCTCATGGAAATAGA CTTATTCTTATGATTGGGAAATTCTGGCTAAATCTTCCCCTTTCACCCTCTCAGTATCTCCAGTTTAAAACCTGGTGGATTGATGGG GTACAAGAACAGGTACGAAAAAATCAGGCTACTAATCCTGTTGCTTATATAGATGAAGACCAATTGCTAGGAAGAGGTCCAAACTGG GACACTATTAACCAACAATCAGTAATGAAAAATGAGGCTATTGAACAACTATAAGGGCTATTTGCCTCAGGGGCCTGGGAAAACATTC AGGACCCAGGGAACCTCATGCCCTTCTTTAGGTTCAATCAGACAAGGT

SEQ ID 43: 30

GCTGACTTGCTAGTATCATCGCATTCATTGAAGCACAAGAACTTCATGCCTTGACTCATGTAAATGCAATAGGATTAAAAAATAAA TTTGATATCACATGGAAACAGACAAAAAAATATTGTACAACATTGCACCCAGTGTCAGATTCTACACCTGGCCACTCAGGAAGCAAGA GTTAATCCCAGAGGTCTATGTCCTAATGTGTTATGGCAAATGGATGTCATGCACGTACCTTCATTTGGAAAATTGTCATTTGTCCAT GTGACAGTTGATACTTATTCACATTTCATATGGGCAACCTGCCAGACAGGAGAAAGTCTTCCCATGTTAAAAGACATTTATTATCTT **GTTTTCCTGTCATGGGAGTTCCAGAAAAAGTTAAAACAGACAATGGGCCAGGTTCTGTAGTAAAG**

SEQ ID 44: 35

35	•
30	GCCAAGGTGGGAGGATTGCTTGAGCACAGGAGTTTGAGGCTGAAGTGAGCTATGATCGCACCACTGCAATCAAT
	TCAGTCAACCCTGCCAGGAGCTATGGAACAATTATTGTTTGT
	TACCTTACTCAGTCACCATATGGGGGGCTGCCCCAGAGAGGTCATGACCTCAAGTGAGGAAGTACTCAGCAGCTGAGCCAGGCCCTAC
	TGATAGCTGGAGGATGCTGCCCCATGCTGCCCACTGTGAGGCAGCAAGCCCTTGCTTG
	TGTCTACCACCCCTAGAAATGGTGCCTAGAGTGAGTGAGT
	CCAGCTACTCAGGAGACTGTGGCAGGAGAATGACTTAAACCAGGGAGTTGGAGGTGAGGTGAGGTCACACACA
40	AGACTGGGTGACAGAGTGAGACTCCCATCTCAAAAAAAAA
	TATTTAAAGGATAGGTGAATGGAGGAAAATAATCAATTGAAGGAGGCTGAGCAGATGAGGTCAAAGAAGATAGAGATCCATAACAGT
	AACCTCATAGAAGCTTATGGAAGCATTTTGACAGTGCTAAAAGCCACATAAAGTTCAAGTAAGACAGTTTCAGAAATGTATAAACAT
	GAATGCCTTTGCAGTGACTTAAGTGTGATTCTGGTGTTTCCTTCTAAAAATACTGCCTTCTCAGGTGTGGGAAGGATTCTATCTTTT
	TAGGCTTTACCACCATAGTTCTCTGCAGGCTTGCAATCCTGAATCAGGCTTGACTTCAGAAAGTGCTTTAAAAGGGAGGCTGGGCGC
	GGTGGCTCATGCCTGTAATCCCAGCACTCTGAGAGGGCTGAGGTTGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTGTGT
45	AGAAAGAAGTAGACATAGGAGACTCCATTTTGTTCTGTACTAAGAAAAATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTAC
	CCCCAACCCGTGCTCTCTGAAACAGGTGCTGTGTCAAACTCAGGGTTAAATGGATTAAGGGTTGTGCAAGATGTGCTTTGTTAAAC
	AAATGCTTGAAGGCAGGATGGTCCTTAAGAGTCATCACCACTCCCTAATCTCAAGTACCCAGGGACACAAAACACTGCGGAAGGCCGC
	AGAGACCTCTGCCTAGGAAAGCAAGGTATTGTCCAAGGTTTCTCCCCATGTGATAGTCTGAAATATGGCCTCGTGGGAAGGGAAAGA
	CCTGACCGTCCCCCAGCCTGACACCCGTAAAGGGTCTGTGCTGAGGAGGATTAGTGTAAGAGGAAGGCATGCCTCTTGCAGTTGAGA
	CAAGAGGAAGGCATCTGTCTCCTGCCCGTCCCTGGGCAATGGAATGTCTCGGTATAAAACCCGATTGATT
50	AGATAGGAAGAAAACGCCTTAGGGCTGGAGGTGTGGGACAAGCCGGCAGCAATACTGCTTTGTAAAGCATTGAGATGTTTATGTGTA
	TGCATATCTAAAAGCACAGCACTTGATTCTTTACCTTGTCTGTGATGCAAAGACCTTTGTTCACGTGTTTGTCTGCTGACCCTCTCC
	CCACTATTGTCTTGTGACCATGACACATCCCCCCTCTCAGAGAAACACCCCACGAATGATCAATAATACTAAGGGAACTCAGAGACGG
	CGCGGATCCTCCATATGCTGAACGCTGGTTCCCTGGGTCCCCTTATTTCTTTC
	TCGTTCCACCTTACGAGAAACACCCACAGGTGTGGAGGGGGAACCCACCC
	GTGAAGGTACGCTCGAGCGTGGTCATTGAGGACAAGTTGACGAGAGATCCCGAGTACATCTACAGTCAGCCTTGCGGTAAGTTTGTG
55	CGCTCGGAAGAAGCTAGGGTGATAATGGGGCCAAACTAAAAGTAAAAGTAAAATATGCCTCTTATCTCAGCTTTATTAAAAATT
	CTTTTAAAAAGAGGGGGAGTTAGAGTATCTACAAAAAATCTAATCAAGCTATTTCAAATAATAGAACAATTTTGCCCATGGTTTCCA
	GAACAAGGAACTTTAGATCTAAAAGATTGGAAAAGAATTGGCGAGGAACTAAAACAAGCAGGTAGAAAGGGTAATATCATTCCACTT

	ACAGTATGGAATGATTGGGCCATTATTAAAGCAGCTTTAGAACCATTTCAAACAAA
5	GGTCCAGAATTAGTGGGGCCATCAGAGTCTAAACCACGAGGGCCAAGTCCTCTTCCAGCAGGTCAGGTGCCCGTAACATTACAACCT CAAACGCAGGTTAAAGAAAATAAGACCCAACCGCCAGTAGCTTATCAATACTGGCCGCCGGCTGAACTTCAGTATCTGCCACCCCCA GAAAGTCAGTATGGATATCCAGGAATGCCCCCCAGCACTACAGGGCAGGGCGCCCATATCCTCAGCCGCCCACTGTGAGACTTAATCCT
	ACAGCATCACGTAGTGGACAAGGTGGTACACTGCACGCAGTCATTGATGAAGCCAGAAAACAGGGAGATCTTGAGGCATGGCGGTTC CTGGTAATTTTACAACTGGTACAGGCCGGGGAAGAGACTCAAGTAGGAGCGCCTGCCCGAGCTGAGACTAGATGTGAACCTTTCACC ATGAAAATGTTAAAAGATATAAAGGAAGGAGTTAAACAATATGGATCCAACTCCCCTTATATAAGAACATTATTAGATTCCATTGCT CATGGAAATAGACTTACTCCTTATGACTGGGAAAGTTTGGCCAAATCTTCCCTTTCATCCTCTCAGTATCTACAGTTTAAAAACCTGG
10	TGGATTGATGGAGTACAAGAACAGGTACGAAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAAACA GGTCCAAATTGGAGCACCATTAACCAACAATCAGTGATGCAGAATGAGGCTATTGAACAAGTAAGGGCTATTGCCTCAGGGCCTGG GGAAAAATTCAGGACCCAGGAACAGCTTTCCCTATTAATTCAATTAGACAAGGGCTCTAAAGAGCCATATCCTGACTTTGTGGCAAGA
	TTACAAGATGCTGCTCAAAAGTCTATTACAGATGACAATGCCCGAAAAGTTATTGTAGAATTAATGGCCTATGAAAAATGCAAAATGCAAATCCA GAATGTCAGTCGGCCATAAAGCCATTAAAAGGAAAAGTTCCAGCAGGAGTTGATGTAATTACAGAATATGTGAAGGCTTGTGATGGG ATTGGAGGAGCTATGCATAAGGCAATGCTAATGGCTCAAGCAATGAGGGGGGCTCACTCTAGGAGGACAAGTTAGAACATTTGGGAAA
15	AAATGTTATAATTGTGGTCAAATCGGTCATCTGAAAAGGAGTTGCCCAGTCTTAAATAAA
	CCTCAGGGTTTTCAAGGACAACAACCCCTACAGAAAATACCACCACTTCAGGGAGTCAGCCAATTACAACAATCCAACAGCTGTCCC GCGCCACAGCAGGCAGCGCCACAGTAGATTTATGTTCCACCCAAATGGTCTCTTTACTCCCTGGAGAGCCCCCACAAAAGATTCCTA GAGGGGTATATGGCCCGCTGCCAGAAGGGAGGGTAGGCCTTATTTTAGGGAGATCAAGTCTAAATTGAAGGGAGTCCAAATTCATA
20	CTGGGGTAATTTATTCAGATTATAAAGGGGGAATTCAGTTAGTGATCAGCTCCACTGTTCCCTGGAGTGCCAATCCAGGTGATAGAA TTGCTCAATTACTGCTTTTGCCTTATGTTAAATTGGGGAAAACAAAACGGAAAGAACAGGAGGGTTTGGAAGTACCAACCCTGCAG GAAAAGCCACTTATTGGGCTAATCAGGTCTCAGAGGATAGACCCCGTGTACAGTCACTATTCAGGGAAAGAGTTTGAAGGATTAGT
	GGATACCCAGGCTGATGTTTCTATCATCGGCATAGGCACCGCCTCAGAAGTGTATCAAAGTGCCATGATTTTACATTGTCTAGGATC TGATAATCAAGAAAGTACGGTTCAGCCTATGATCACTTCTATTCCAATCAAT
25	CACTGTAGAGCCTCCCAAAACCCATTCCATTAATTGGGGGGGAAAAAAAA
	CCCGGTTGCCCTCTCCAGCCATGGTCCCCTTTAATTATAATTGATCTGAAGGATTGCTTTTTTACCATTCCTCTGGCAAAAGAGGAT TTTGAAAAATTTGCTTTTACTATACCAGCCTAAATAATAAAGAACCAGCCACCAGGTTTCAGTGGAAAGTATTGCCTCAGGGAATGC TTAATAATTCAACTATTTGTCAGACTTTCATAGCTCAAGCTCTGCAACCAGTTAGAGACAAGTTTTCAGACTGTTATATCGTTCATT
30	ATGTTGATATTTTGTGTGCTGCAGAAACGAGAGACAAATTAATT
	GAATGTTACCTCCAGAGGCAACTAAAGAAATTAAATTAA
35	GTGGAAATGACCCAGATAAAATCACTGTTCCTTTCAACAAGCAACAAGTTAGACAAGCCTTTATCAGTTCTGGTGCATGGCAGATTG GTCTTGCTAATTTTCTGGGAATTATTGATAATCATTACCCAAAAAAATCTTCCCAGTTCTTAAAATTGACTACTTGGATTCTAC CTAAAATTACCAGACGTGAACCTTTAGAAAATGCTCTAACAGTATTTACTGATGGTTCCAGCAATGGAAAAGCGGCTTACACAGGGC CGAAAGAACGAGTAATCAAAAACTCCGTATCAATCAGCTCAAAGAGCAGAGTTGGTTG
40	AACCTATCAATATTATATCAGATTCTGCATATGTAGTACAGGCTACAAGGGATGTTGAGACAGCTCTAATTAAATATAGCACGGACG ATCATTTAAACCAGCTATTCAATTTATTACAACAAACTGTAAGAAAAAGAAATTTCCCATTTTATATTACTCATATTCGAGCACACA CTAATTTACCAGGGCCTTTGACTAAAGCAAATGAACAAGCTGACTTACTGGTATCATCTGCATTCATAAAAGCACAAGAACTTCTTG
	CTTTGACTCATGTAAATGCAGCAGGATTAAAAAAACAAATTTGATGTCACATGGAAACAGGCAAAAGATATTGTACAACATTGCACCC AGTGTCAAGTCTTACACCTGTCCACTCAAGAGGCAGGAGTTAATCCCAGAGGTCTGTGTCCTAATGCGTTATGGCAAATGGATGG
45	CAGGATATTGTAGTAAAGCTTTCCAAAAATTCTTAAGTCAGTGGAAAATTTCACAACAGGAAATTCCTTATAATCCCCAAGGAC AGGCCATAGTTGAAAGAACTAATAGAACACTCAAAACTCAATTAGTAAACAAAAAGAAGAGGGGGAGACAGTAAGGAGGTACCACTC CTCAGATGCAACTTAATCTAGCACTCTATACTTTAAACTTTTTTAAACATTTATAGAAATCAGACTACTACTACTACGCAAACAACAAC
	TTACTGGTAAAAAGCACAGCCCACATGAAGGAAAACTAATTTGGTGGAAAGATAATAAAAATAAGACATGGGAAATAGGGAAGGTGA TAACGTGGGGGAGAGGTTTTGCTTGTGTTTCACCAGGAGAAAATCAGCTTCCTGTTTGGATACCCACTAGACATTTGAAGTTCTACA ATGAACCCATCGGAGATGCAAAGAAAAGGGCCTCCACAGAGATGGTAACCCCAGTCACATGGATGATAATCCTATAGAAGTATATG
50	TTAATGATAGTGTATGGGTACCTGGCCCCACAGATGATCGCTGCCCTGCCAAACCTGAGGAAGAAGGGATGATGATAAATATTTCCA TTGTGTATCGTTATCCTCCTATTTGCCTAGGGAGAGACCCAGGATGTTTAATGCCTGCAGTCCAAAATTGGTTGG
	ACCACAATTGCTCAGGACAAACTCAGTCGTGTCCAAGTGCACAAGTGAGTCAGCGACTGATAGCGACTTAACAGAAAGTCTAGACA AACATAAGCATAAAAAAATTACAGTCTTTCTACCCTTGGGAATGGGGAGAAAAAGGAATCTCTACCCCCAAGACCAGAAATAAAGTC CTGTTTCTGGTCCTGAACATCCAGAATTATGGAGGCTTTGGCCTGACACCACATTAGAATTTGGTCTGGAAATCAAACTTTAGAAAC
55	AAGAGATCGTAAGCCATTTTATACTATCGACCTAAATTCCAGTCTAACGGTTCCTTTACAAAGTTGCGTAAAGCCCTCTTATATGCT AGTTGTAGGAAATATAGTTATTAAACCAGACTCCCAAACTATAACCTGTGAAAATTGTAGATTGTTTACTTGCATTGATTCAACTTT TAATTGGCGGCACCGTATTCTGCTGGTGAGAGAGAGAGAG

	ATCCATCCATATTTTGACTGAAGTATTAAAAGACATTTTTAAATAGATCCAAAAGATTCATTTTTACCTTAATTGCAGTGATTATGGG
	ATTAATTGCAGTCACAGCTACGGCTGCTGCGGCAGGAGTTGCATTGCACTCTTCTGTTCAGTCGGTAAACTTTGTTAATGATTGGCA
	AAAGAATTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAAATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCAT
	TTGGATGGGAGACAGACTCATGAGCTTAGAACATTGTTTCCAGTTACAGTGTGACTGGAATACGTCAGATTTTTGTATTACACCCCCA
5	AATTTATAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAGAAGATAATCTCACTTTAGACATTTCCAA
5	ATTAAAATAACAAATTTTCGAAGCATCAAAAGCCCATTTAAATTTGATGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
	CGCAAATCTTAACCCTGTCACTTGGGTTAAGACCATCGGAAGTACTATGATTATAAATCTCATATTAATCCTTGTGTGCCTGTTTTG
	TCTGTTGTTGTTGCCGGGGGGGGGGCCACGGGCCACGGGCCATGACGGGCCATGACGGGCCATGACGGCCGGTTTTGTC
	GAAAAGAAAAGGGGGAAATGTGGGGGAAAAGCAAGAGAGAG
	CATTTTGTTCTGTACTAAGAAAAATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCCAACCCCGTGCTCTCTGAAACA
10	GGTGCTGTGTCAAACTCAGGGTTAAATGGATTAAGGGTTGTGCAAGATGTGCTTTGTTAAACAAATGCTTGAAGGCAGCATGCTCCT TAAGAGTCATCACCACTCCCCTAATCTCAAGTACCCAGGGACACAAAAACTGCGGAAGGCCGCAGGGACCTCTGCCTAGGAAAGCCAG
10	GTATTGTCCAAGGTTTCTCCCCATGTGATAGTCTGAAATATGGCCTCATGGGAAGGCCGCAGGGACCTCTGCCTAGGAAAGCCCGC GTATTGTCCCAAGGTTTCTCCCCCATGTGATAGTCTGAAATATGGCCTCATGGGAAGGGAAGGCCGCAGGGACCCTGACCGTCCCCCAGCCCGACACC
	CGTAAAGGGTCTGTGCTGAGGAGGATTAGTATAAGAGGGAAGGCATTCCTCTTGCAGTTGAGACAAGACGAAGGGAAGGCATCTGTCTCCTGC
	CCGTCCCTGGGCAATGGAATGTCTCGGTATAAAACCCGATTGTACGTTCCATCTACTGAGATAGGAAGAAAACGCCTTAGGGCTGGA
	GGTGGGACATGCAGGCAGCAATACTGCTTTGTAAAGCATTGAGATGTTTATGTGTATGCATATCTAAAAGCACAGCACTTGATTCTT
	TACCTTGTCTATGATGCAAAGACCTTTGTTCACCTGTTTGTCTGCTGACCCTCTCCCCACTATTGTCTTGTGACCATGACACATCCC
15	CTCTCAGAGAAACACCCCACGAATGATCAATAAATACTAAGGGAACTCAGAGACGCGCGCG
	CCTGGGTCCCCTTATTCTTTCTCTATACTTTGTCTCTGTGTCTTTTTCTTTTCCAAGTCTCTCATTCCACCTTAAGAGAAACACTC
	ACAGGTGTGGAGGGGCAACCCATCCCTTCAGAGGTGGGTG
	AACCCCATCTCTACTAAAAATACAAAATTAGCCAGGTGTGGTGGCAGGTGTCTGTAGTCCCAGCTACTTGGGAGGCTGACGAGAATC
	GCTTGAACCTGGGAGGGGGGGGGTTTCAGTGAGCCGAGATTGCACCACTGCAGCCTGGGGGGACAGAGTGAAACTCTGTCTCAA
	AAAAACAACAAAAAACCCCCACCTATAGACAGGACTAGCTACATAAATAA
20	TTTTTCAAAGACGTAGAAGGCCGGGTGCGGTGCGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGC
	CAGGAGTTCGAGACAGCCTGACCAATATGGTGAAACCCCCATCTCTACTAAAAATACAAAAATTAGCTGGGTGTGGTAGCGGGCGCCT
	GTAGTCCCAGCTACTCAGGAGGCTGAGGCAGAAGAATTACTTGAACCCAGGAGACGGAGGTTGCAGTGAGCTGAGATCGTGCCACTG
	CACTCTCCAGCCTCCTCGGTGACAGAGCGAGACTCTGTCTCAAAAAAAA
	ATATAAGGCACTTTCCTTTATTCTGCAATCTGTCTCTCCACTTTTCATAGTATTTTTCATTGTTATTTAACATCATGTTTTGTCA
	GGTGAGGACATTTACTCAGCCAGTGCAGCACTCACTGGTATCCAGGGGCCATAGGTGATTTGACGCACCCACATGGCCCACCAGCTG TTGAGTTCCACCTCCAGCCAGCCACTGGACCAACATGCAGTGCCCTGGCTGG
25	GTCCTCCTGGCCAAACCCACAGGGACAGGTAAACCCCCCTTGTATGTGTTTTGTACTTGGATCTGGGTGGG
	SEQ ID 45:
	•
	TGTGGGGAAAAGCAAGAGAGATCAAATTGTTACTGTGTCTGTGTAGAAAGAA
	GGTTGAATGGATTAAGGGCGGTGCAGGATGTGCTTTGTTAAACAGATGCTTGAAGGCAGCATGCTCCTTAAGAGTCATCACCACTCC
	CTAATCTCAAGTACCCAGGGACACAAAAACTGCGGAAGGCCGCAGGGACCTCTGCCTAGGAAAGCCAGGTATTGTCCAAGGTTTCTC
30	CCCATGTGATAGTCTGAAATATGGCCTCGTGGGAAGGGAAGGCCTGACCGTCCCCCAGCCCGACACCTGTAAAGGGTCTGTGCTGA
	GCAGGATTAGTAAAAGAGGAAGGAATGCCTCTTGCAGTTGAGACAAGAGGAAGGCATCTGTCTCCTGCCTG
	TGTCTCGGTATAAAACCCGATTGTATGCTCCATCTACTGAGATAGGGAAAAACCGCCTTAGGGCTGGAGGTGGGACCTGCGGGCAGC
	AATACTGCTTTGTAAAGCATTGAGATGTTTATGTGTATGCATATCCAAAAGCACAGCACTTAATCCTTTACATTGTCTATGATGCCA
	AGACCTTTGTTCACGTGTTTGTCTGCTGACCCTCTCCCCCACAATTGTCTTGTGACCCCTGACACATCCCCCCTCTTTGAGAAACACCCCA
35	CAGATGATCAATAAATACTAAGGGAACTCAGAGGCTGGCGGGATCCTCCATATGCTGAACGCTGGTTCCCCGGGTCCCCCTTATTTCT
35	TTCTCTATACTTTGTCTCTGTGTCTTTTTCCTTTTCCAAATCTCGTCCCACCTTACGAGAAACACCCCACAGGTGTGTAGGGGGCAAC
	CCACCCCTACATCTGGTGCCCAACGTGGAGGCTTTTCTCTAGGGTGAAGGTACGCTCGAGCGTAATCATTGAGGACAAGTCGACGAG
	AGATCCCGAGTACATCTACAGTCAGCCTTACGGTAAGCTTGCGCGCTCGGAAGAAGCTAGGGTGATAATGGGGCAAACTAAAAGTAA
	AATTAAAAGTAAATATGCCTCTTATCTCAGCTTTATTAAAAATTCTTTTTAAAAAGAGGGGGAGTTAAAGTATCTACAAAAAATCTAAT
	CAAGCTATTTCAAATAATAGAACAATTTTGCCCATGGTTTCCAGAACAAGGAACTTCAGATCTAAAAGATTGGAAAAGAATTGGTAA GGAACTAAAAACAAGCAGGTAGGAAGGGTAATATCATTCCACTTACAGTATGGAATGATTGGGCCATTATTAAAGCAGCTTTAGAACC
40	ATTTCAAAACAAGCAGGTAGGAAGGGTAATATCATTCCACTTACAGTATGGAATGATTGGGCCATTATTAAAGCAGCTTTAGAACC
	CCAGAAAGAAAACCGAAAGTTTACATTGCGAATATGTAGCAGAGCCGGTAATGGCTCAGTCAACGCAAAATGTTGACTATAATCAATT
	ACAGGAGGTGATATATCCTGAAACGTTAAAATTAGAAGGAAAAGGTCCAGAATTAATGGGGCCATCAGAGTCTAAACCACGAGGCAC
	AAGTCCTCTTCCASCAGGTCAGGTGCTCGTAAGATTACAACCTCAAAGCAGGTTAAAGAAATAAGACCCAACCGCAAGTAGCCTA
	TCAATACTGCCGCTGGCTGAACTTCAGTATCGGCCACCCCCAGAAGTCAGTATGGATATCCAGGAATGCCCCCAGCACCACAGGGC
	AGGGCGCCATACCATCAGCCGCCCACTAGGAGACTTAATCCTATGGCACCACCTAGTAGACAGGGTAGTGAATTACATGAAATTATT
45	GATAAATCAAGAAAGGAAGGAGATACTGAGGCATGGCAATTCCCAGTAACGTTAGAACCGATGCCACCTGGAGAAGGAGCCCAAGAG
	GGAGAGCCTCCCACAGTTGAGGCCAGATACAAGTCTTTTTCGATAAAAATGCTAAAAGATATGAAAGAGGGAGTAAAACAGTATGGA
	GGAGAGCCTCCCACAGTTGAGGCCCAGATACAAGTCTTTTTCGATAAAAATGCTAAAAGATATGAAAGAGGGGGGGTAAAACAGTATGGA CCCAACTCCCCTTATATGAGGGACATTATTAGATTCCATTGCTTATGGACATAGACTCCTTATGATGGGAGATTCTGGCAAAA
	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGAATAGGTCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT
	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC
50	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGAATAGGTCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT GAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCTTGGGGAAAAAATCCAAGACCAGGAAGTACCTGCCCCTCATTTAATACA GTAAGACAAGGTTCAAAAGAGCCCTACCCTGATTTTGTGGCCAAGGCTCCAAGATGTTGCTCAAAAGTCCAATGCCAAAAAGCC
50	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGAATAGGTCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT GAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCTTGGGGAAAAAATCCAAGACCAGGAAGTACCTGCCCCTCATTTAATACA GTAAGACAAAGGTTCAAAAGAGCCCTACCCTGATTTTGTGGCCAAGGCTCCAAGATGTTGCTCAAAAGTCCAATGCCAATGAAAAAGCC GGTAAGGTCATAGTGGAGTTGATGGCATATGAAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAAGGGAAAAGGTCCT
50	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGAATAGGTCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT GAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCTTGGGGAAAAAATCCAAGACCCAGGAAGTACCTGCCCCTCATTTAATACA GTAAGACAAGGTTCAAAAGAGCCCTACCCTGATTTTGTGGGCAAGGCTCCAAGATGTTGCTCAAAAGTCAATTGCCGATGAAAAAGCC GGTAAGGTCATAGTGGAGTTGATGGCATATGAAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAAGGGAAAAGGTCCT GCAGGATCAGATGTAATCTCAGAATATGTAAAAGCCTGTGATGGAATCGGAGGAGCTATGCCATTAAAGCCATTAAAGGAAAGGTCCAAG
50	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGAATAGGTCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT GAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCTTGGGAAAAAATCCAAGACCCAGGAAGTACCTGCCCCTCATTTAATACA GTAAGACAAAGGTTCAAAAGAGCCCTACCCTGATTTTGTGGGCAAGGCTCCAAGATGTTGCTCAAAAGTCAATTGCCGATGAAAAAGCC GGTAAGGTCATAGTGGAGTTGATGGCCATATGAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAAGGCAATGGAAAAAGCC GCCAGGATCAGATGTGAATCTCCAGAATATGTAAAAGCCTGTGGAGGAGCTATGCCATTAAAGCCATTAAAAGGAAAGGTTCCT GCAGGATCAGATGTAATCTCCAGAATATGTAAAAGCCTGTGGAGGAGCTATGCGAGGAGCTATGCTAAAGCTATGGCTCAAGACA ATAACAGGAGTTGTTTTAGGAGGACAAGTTAGAAACATTTGGGAGAGAAATGTTATAATTGTGGTCAAATTGGTCACTTAAAAAAAA
50	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGACTAGGGCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT GAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCTTGGGAAAAAATCCAAGACCCAGGAAGTACCTGCCCCTCATTTAATACA GTAAGACAAAGGTTCAAAAGAGCCCTACCCTGATTTTGTGGCCAAGGCTCCAAGATGTTGCTCAAAAGTCAATTGCCGATGAAAAAGCC GGTAAGGTCATAGTGGAGTTGATGGCCATATGAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAAGGCAATGGAAAAAGCC GGTAAGGTCATAGTGGAGTTGATGGCCATATGAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAAGGAAAGGTTCCT GCAGGATCAGATGTAATCTCCAGAATATGTAAAAGCCTGTGATGGAAACGGCAGGAGCTATGCAAAAGCCTATGGCTCAAGACA ATAACAGGAGTTGTTTTAGGAGGACAAGTTAGAACATTTGGGAGAAAATGTTATAATTGTGGTCAAATTGGTCACTTAAAAAAAA
	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGACTAGGGCTCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT GAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCTTGGGGAAAAAATCCAAGACCAGGAAGTACCTGCCCTCATTTAATACA GTAAGACAAAGGTTCAAAAGAGCCCTACCCTGATTTTGTGGGCAAGGCTCCAAGATGTTGCTCAAAAGTCAATTGCCGATGAAAAAGCC GGTAAGGTCATAGTGGAGTTGATGGCCATATGAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAGCCATTAAAAGGAAAGGTTCCT GCAGGATCAGATGTAATCTCCAGAATATGTAAAAGCCTGTGATGGAGAGCCATTGCGAGAGGACCATTAAAGGAAAGGTTCCT GCAGGATCAGATGTTATTTAGGAGGACAAGTTAGAACACCTGTGATGGAGAGGAGCTATGCGTCAAATTGGTCCAATTGGCTCAAGAA ATAACAGGAGTTGTTTTAGGAGGACAAGTTAGAACACTATTGGAGGAAAATGTTATAATTGTGGTCAAATTGGTCACTTAAAAAAAGAAT TGCCCAGTCTTAAACAAACAGAATATAACTATTCAAGCAACTACAACAGGTAGAGAGACCACCTGACTTATGTCCCAAGATGTAAAAA GGAAAACATTGGGCTAGTCAATGTCGTTCTAAATTTGATAAAATGGGCAACCATTGCGGGAAAACGACCAAGGGCCACCTCAG
50 55	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGACTAGGGCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT GAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCTTGGGGAAAAAATCCAAGACCAGGAAGTACCTGCCCCTCATTTAATACA GTAAGACAAAGGTTCAAAAGAGCCCTACCCTGATTTTGTGGCAAGGCTCCAAGATGTTGCTCAAAAGTCAATTGCCGATGAAAAAGCC GGTAAGGTCATAGTGGAGTTGATGGCCATATGAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAGCCATTGAAGGAAAGGTTCCT GCAGGATCAGATGTAATCTCCAGAATATGTAAAAGCCTGTGATGGAAACGGCGAGCTATGCATAAGCCATTAAAAGGAAAGGTTCCT GCAGGATCAGATGTTATTTAGGAGGACAAGTTAGAACACCTGTGATGGAGAGGAGCTATGCGTCAAATTGGTCACTTAAAAAAGAAT TGCCCAGTCTTAAACAAACAGAATATAACTATTCAAGCAACTATGGAGAAGAGGCCACCTGACTTATGGTCCAAGATGTAAAAAA GGAAAACATTGGGCTAGTCAATGTCGTTCTAAATTTGATAAAATGGGCAACCATTGCGGGAAACGAGCAAAGGGCCACCTCAG GCCCCACAACAAACTGGGGCATTCCCAATTCGGCCATTTGTTCCTCAGGGTTTTCAGGGACAACGAGCCACCCCACTGTCCCAAGGGCCACCTCAG
	CCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAAA TCGTCTCTCACCCTCTCAATTTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCC AATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGACTAGGGCTCAAAATTGGAGTACTATTAGTCAACAAGCATTAATGCAAAAT GAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCTTGGGGAAAAAATCCAAGACCAGGAAGTACCTGCCCTCATTTAATACA GTAAGACAAAGGTTCAAAAGAGCCCTACCCTGATTTTGTGGCAAGGCTCCAAGATGTTGCTCAAAAGTCAATTGCCGATGAAAAAGCC GGTAAGGTCATAGTGGAGTTGATGGCCATATGAAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAAGGCATTGAAGAAGGTTCCT GCAGGATCAGATGTAATCTCCAGAATATGTAAAAGCCTGTGATGGAGAGCCATTGCCAAAAGCTATGCCTAAAAGGAAAGGTTCCT GCAGGATCAGATGTTATTTAGGAGGACAAGTTAGAACACCTGTGATGGAGAGGAGCTATGCATAAAGCCATTAAAAGGAAAGGTCCAA ATAACAGGAGTTGTTTTAGGAGGACAAGTTAGAACACTATTGGAGGAAAATGTTATAATTGTGGTCAAATTGGTCACTTAAAAAAAGAAT TGCCCAGTCTTAAACAAACAGAATATAACTATTCAAGCAACTACAACAGGTAGAGAGCCACCTGACCTAAGGGCAACGAGCCACGAGCCACGAGCCACGAGCCACGAGCCACGAGCCACGAGCCACGAGCCACGAGCCACGAGCCACGAGCCACGAGCCACGGCAAAGGGGCCACCCCTGACTTATGTCCAAGAGGGCCACCCCCGGAAAACGAGCCAAGGGGCCACCCTCAGGCAAAGGGCCACCACGGGAAAACGACCACGGGAAAACGACCACGGGCAACGAGCAAAGGGCCACCA

	GAAGATCAAGTCTAAATCTAAAAGGAGTTCAAATTCATACTAGTGTGGTTGATTCAGACTATAAAGGCGAAATTCAATTGGTTATTA GCTCTTCAATTCCTTGGAGTGCCAGTCCAAGAGACAGGATTGCTCAATTATTACTCCTGCCATATATTAAGGGTGGAAATAGTGAAA
	TAAAAAGAATAGGAGGGCTTGTAAGCACTGATCCAACAGGAAAGGCTGCATATTGGGCAAGTCAGGTCTCAGAGAACAGACCTGTGT GTAAGGCCATTATTCAAGGAAAACAGTTTGAAGGGTTGGTAGACACTGGAGCAGATGTCTCTATTATTGCTTTAAATCAGTGGCCAA
5	AAAACTGGCCTAAACAAAAGGCTGTTACAGGACTTGTCGGCATAGGCACAGCCTCAGAAGTGTATCAAAGTATGGAGATTTTACATT GCTTAGGGCCAGATAATCAAGAAAGTACTGTTCAGCCAATGATTACTTCAATCCTCTTAATCTGTGGGGGTCGAGATTTATTACAAC
	AATGGGGTGCGGAAATCACCATGCCCGCTCCATTATATAGCCCCCACGAGTCAAAAAATCATGACCAAGATGGGATATATACCAGGAA AGGGACTAGGGAAAAATGAAGATGGCATTAAAGTTCCAGTTGAGGCTAAAATAAAT
	AGGGGCGGTCACTGTAGAGCCTCCTAAACCCATACCACTAACTTGGAAAAACAGAAAAACCGGTGTGGGTAAATCAGTGGCCGCTACC AAAACAAAAACTGGAGGCTTTACATTTATTAGCAAATGAACAGTTAGAAAAGGGTCACATTGAGCCTTCGTTCTCACCTTGGAATTC
10	TCCTGTGTTTGTAATTCAGAAGAAATCAGGCAAATGGCATACGTTAACTGACTTAAGGGCTGTAAACGCCGTAATTCAACCCCATGGG GCCTCTCCAACCCGGGTTGCCCTCTCCGGCCATGATCCCCAAAAGATTGGCCTTTAATTAA
	CATCCCTCTGGCAGAGCAGGATTGTGAAAAATTTGCCTTTACTATACCAGCCATAAATAA
	${\tt TTCAGACTGTTATATTCATTCATTATTGATGATGATATTTTATGTGCTGCAGAAACGAAAGATAAATTAATT$
15	AGAAAATAGAAAAATTAAGCCACAAAAAAATAGAAATAAGAAAAGACACACATTAAAAACACTAAATGATTTTCAAAAAATTACTAGGAGA TATTAATTGGATTCGGCCAACTCTAGGCATTCCTACTTATGCCATGTCAAAATTTGTTCTCTATCTTAAGAGGGAGACTCAGACTTAAA
10	TAGTCAAAGAATATTAACCCCAGAGGCAACAAAAGAAATTAAATTAGTGGAAGAAAAAATTCAGTCAG
	GTCATTCCTTCCTCACAGTACAGTTAAGACTTTTACATTGTACTTGGATCAAATAGCTACATTAATCGGTCAGACAAGATTACGAATA AACAAAATTATGTGGAAATGACCCAGACAAAATAGTTGTCCCTTTAACCAAGGAACAAGTTAGACAAGCCTTTATCAAATGCGGTGG
00	ATGGCAGATTGGTCTTGCTAATTTTGTGGGGACTTATTGATAATCATTACCCAAAAACAAAGCCTTTACCAATTCTGGTGC ATGGCAGATTGGTCTTGCTAATTTTGTGGGGACTTATTGATAATCATTACCCAAAAAACAAAGATCTTCCAGTTCTTAAAATTGACTAC TTGGATTCTACCTAAAATTACCAGACGTGAACCTTTAGAAAATGCTCTAACAGTATTTACTGATGGTTCCAGCAATGGAAAAGCAGC
20	TTACAGTATICTACCTARATTACCAGAGTAATCAAAACTCCATATCAATCGGCTCAAAGAGACGAGTTGGTTG
	TAGCATGGATGATCAGTTAAAACCAGCTATTCAATTTATTACCAACAAACGAAAACGAAAAAGAAAATTTCCCATTTTATATTACTTATAT TCGAGCACACCTAATTTACCAGGGCCTTTGACTAAAGCAAATGAACAAGCTGACTTACTGGTATCATCTGCACTCATCAAAAAAAA
	AGAACTTCATGCTTTGACTCATGTAAATGCAGGGAGTTAAAAAACAAAC
25	AATGGATGTCACGCATGTACCTTCATTTGGAAGATTATCATATGTTCATGTGATACGGTTGATACCTTATTCACATTTCATATGGGCAAC TTGCCAAACAGGAGAAAGTACTTCCCATGTTAAAAAAACATTTATTGTCTTGTTTTGCTGTAATGGGGAGTTCCAGAAAAAATCAAAAC
	TGACAATGGACCAGGATATTGTAGTAAAGCTTTCCAAAAATTCTTAAGTCAGTGGAAAATTTCACATACAACAGGAATTCCTTATAA TTCCCCAAGGACAGGCCATAGTTGAAAGAACTAATAGAACACTCAAAAACTCAATTAGTTAAACAAAAAGAAGGGGGGAGACAGTAAGGA
	GTGTACCACTCCTCAGATGCAACTTAATCTAGCACTCTATACTTTAAATTTTTTAAACATTTATAGAAATCAGACTACTACTACTTCTGC AGAACAACATCTTACTGGTAAAAAGAACAGCCCCACATGAAGGAAAACTAATTTGGTGGAAAGATAATAAAAATAAGACATGGGAAAT
30	AGGGAAGGTGATAACGTGGGGGGAGAGGTTTTGCTTGTGTTTCACCAGGAGAAAATCAGCTTCCTGTTTGGTTACCCACTAGACATTT GAAGTTCTACAATGAACCCATCGGAGATGCAAAGAAAAGGGCCTCCACGGAGATGGTAACACCAGTCACATGGATGG
	$\begin{tabular}{lllllllllllllllllllllllllllllllllll$
	AGAAGTACCTACTGTCAGTCCCATCAGTAGATTCACTTATCACATGGTAAGCGGGATGTCACTCAGGCCACGGGTAAATTATTTACA AGACTTTTCTTATCAAAGATCATTAAAATTTAGACCTAAAGGGAAACCTTGCCCCAAGGAAATTCCCAAAGAATCAAAAAATACAGA
35	AGTTTTAGTTTGGGAAGAATGTGTGGCCAATAGTGCGGTGATATTATAAAACAATGAATTTGGAACTATTATAGATTGGGCACCTCG AGGTCAATTCTACCACAATTGCTCAGGACAAACTCAGTCGTGTCCAAGTGCACAAGTGAGTCCAGCTGTTGATAGCGACTTAACAGA
	AAGTTTAGACAAAACATAAGCATAAAAAAATTGCAGTCTTTCTACCCTTGGGAATGGGGGAGAAAAAGGAATCTCTACCCCCAAGACCAAA AATAGTAAGTCCTGTTTCTGGTCCTGAACATCCAGAATTATGGAGGCTTACTGTGGCCTCACACCACATTAGAATTTGGTCTGGAAA
	TCAAACTTTAGAAACAAGAGATTGTAAGCCATTTTATACTGTCGACCTAAATTCCAGTCTAACAGTTCCTTTACAAAGTTGCGTAAA GCCCCCTTATATGCTAGTTGTAGGAAATATAGTTATTAAACCAGACTCCCAGACTATAACCTGTGAAAATTGTAGATTGCTTACTTG
40	CATTGATTCAACTTTTAATTGGCAACACCGTATTCTGCTGGTGAGAGCAAGAGAGGGGCGTGTGGATCCCTGTGTCCATGGACCGACC
	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGCTGCTGTAGCAGGAGTTGCATTGCACTCTTCTGTTCAGTCAG
	TAGACAAACTGTCATTTGGATGGGAGACAGACTCATGAGCTTAGAACATCGTTTCCAGTTACAATGTGACTGGAATACGTCAGATTT TTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAGAGAAGATAATCTCAC
45	TTTAGACATTTCCAAATTAAAAGAACAAATTTTCGAAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGG AGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTTAAGACCATTGGAAGTACATCGATTATAAATCTCATATTAATCCT
	TGTGTGCCTGTTTTGTCTGTTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGAGAG
50	AGACATAGGAGACTCCATTTTGTTATGTGCTAAGAAAAATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC GTGCTCTCTGAAACATGTGCTGTGTCAACTCAGGGTTGAATGGATTAAGGGCGGTGCAGGATGTGCTTTGTTAAACAGATGCTTGAA GGCAGCATGCTCCTTAAGAGTCATCACCACTCCCTAATCTCAAGTACCCAGGGACACAAAAACTGCAGAAGGCCGCAGGGACCTCTG
50	CCTAGGAAAGCCAGGTATTGTCCAAGGTTTCTCCCCCATGTGATAGTCCGAAGGAACACAAAAAACTGGCAAGGGAAAGGCCGCAGGGAACCTCTG CCCAGGCCGACACCTGTAAAGGGTCTGTGCTGAGGAGGAGGAATAGTAAAAGAGGAAGGA
	CATCTGTCTCCTGCCTGTCCCTGGGCAATGGAATGTCTCGGTATAAAAACCCGATTGTATGCTCCCATCTACTGAGAAAGGAAAGG GCCTTAGGGCTGGAGGTGGGACCTGCGGGCAGCAATACTGCTTTGTAAAGCATTGAGATGTTTATGTGTATGCATATCCAAAAAGCAA
55	AGCACTTAATCCTTTACATTGTCTATGATGCCAAGACCTTTGTTCACGTGTTTGTCTGCTGACCCTCTCCCCCACAATTGTCTTGTGA CCCTGACACATCCCCCCTCTTTGAGAAACACCCCACAGATGATCAATAAATA
55	CTGAACGCTGGTTCCCCGGGTCCCCTTATTTCTTTCTCTATACTTTGTCTCTGTGTCTTTTCTTTTCCAAATCTCTCGTCCCACCT TACGAGAAACACCCACAGGTGTGTAGGGGGCAACCCACCC

SEQ ID 46:

ETOVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTVLDSIAHGNRLTPYDWEILAKSSLSSSQYLQFKTWWIDGVQEQVRK KSGY*AHC*YRRPIVRNRSKLEHH*PTISDAE*GY*TSKGYLPQGLGKNSGPRNSFPY*FN*TRL*RAIS*LCGKITRCCSKVYYR *QCPKSYCRINGL*KCKSRMSVGHKAIKRKSSSRS*CNYRICEGL*WDWRSYA*GNANGSSNEGAHSRRTS*NIWEKML*LWSNRSS EKELPRLKOAKKKKKKKK

SEO ID 47:

5

15

20

25

30

35

55

EETQVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTVLDSIAHGNRLTPYDWEILAKSSLSSSQYLQFKTWWIDGVQEQVR KNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEAIEQVRAICLRAWGKIQDPGTAFPINSIRQGSKEPYPDFVARLQDAAQKSIT DDNARKVIVELMAYENANPECQSAIKPLKGKVPAGVDVITEYVKACDGIGGAMHKAMLMAQAMRGLTLGGQVRTFGKKCYNCGQIGH LKRSCPGLNKONIINOAITEKKKKKKKK

SEO ID 48: 10

EETQVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTLLDSIAHGNRLTPYDWEILAKSSLSSSQYLQFKTWWIDGVQEQVR KNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEAIEQVRAICLRAWGKIQDPGTAFPINSIRQGSKEPYPDFVARLQDAAQKSIT DDNARKVIVELMAYENANPECQSAIKPLKGKVPAGVDVITEYVKACDGIGGAMHKAMLMAQAMRGLTLGGQVRTFGKKCYNCGQIGH RKRSCPGLNKONIINQAITAKNKKPSGLCPKCGKAKHWANQCHSKFDKDGQPLSGNRKRGQPQAPQQTGAFPVKLFVPQGFQGQQPL OKIPPLQGVSQLQQSNSCPAPQQAAPQ*IYVPPKWSFYSLESPHKRFLEGYMARCQKGG*AFEGDQV*I*RESKFILG*FTQIIKGE FS**SAPLFPGVPIQVIELLNYCFCLMQKKKKKKKK

SEQ ID 49:

GSQAGVKQYGPNSPYIRILLNSIAHGNRLISYDWEILAISSLSPSQYLQFKTWWIDGVQEQVRKNQATNPVAYIDEDQLLGRGPNWD TINQQSVMKMRLLNNYKGYLPQGLGKHSGPRNLMPFF*FNQTRL*RAISRLCGKVARCSSKIHCR*RPKSYCRNNGLSKRKFRVSIS HKAIKRKCFSRS*CNYRICEGL*WDWRSYA*GNAIGSSNYRGCYRRTS*NIWGKML*LWSNRSSKKELPELKLPPKKKKKKKKK

SEO ID 50:

QKNESSKLSIT*LKEQSWLPSLQC*QDFNQSINIVSDSAYVVQATKDIERALIKYIMDDQLNPLFNLLQQNVRKRNFPFYITHIRAH TNLPGPLTKANEQADLLVSSAFMEAQELHALTHVNAIGLKNRFDITWKQTKNIVQHCTQCQILHLATQEARVNPRGLCPNVLWQMDV MHV**PSFGK**LSFVHVTVDTYSHFIWATCQTGESTSHVKRHLLSCFPVMGVPEKVKTDNGPGYCSKAVQKFLNQWKITHTIGILYNSQG OAIIERTNRTLKAOLVKOKKKKKKKKK

SEO ID 51:

QKNESSKLSIT*LKEQSWLPSLQC*QDFNQSINIVSDSAYVVQATKDIERALIKYIMDDQLNPLFNLLQQNVRKRNFPFYITHIRAH TNLPGPLTKANEQADLLVSSAFMEAQELHALTHVNAIGLKNKFDITWKQTKNIVQHCTQCQILHLATQEARVNPRGLCPNVLWQMDV MHV**PSFGK**LSFVHVTVDTYSHFIWATCQTGESTSHVKRHLLSCFPVMGVPEKVKTDNGPGYCSKAVQKFLNQWKITHTIGILYNSQG QAIIERTNRTLKAQLVKQKEKKKKK

SEQ ID 52:

 $\label{eq:construction} QKNESSKLSITRLKEQSWLPSLQC*QDFNQSINIVSDSAYVVQATKDIERALIKYIMDDQLNPLFNLLQQNVRKRNFPFYITHIRAH$ TNLPGPLTKANEQADLLVSSAFMEAQELHALTHVNAIGLKNKFDITWKQTKNIVQHCAQCQILHLATQEVRVNPRGLCPNVLWQMDV MHVPSFGKLSFVHVTVDTYSHFIWATCQTGESTSHVKRHLLSCFPVMGVPEKVKTDNGPGYCSKAVQKFLNQWKITHTIGILYNSQG QAIIERTNRTLKAQLVKQKKKKKKKK

SEO ID 53:

QKNESSKLSIT*LKEQSWLPSLQC*QDFNQSINIVSDSAYVVQATKDIERALIKYIMDDQLNPLFNLLQQNVRK*NFPFYITHIRAH TNLPGPLTKANEQADLLVSSAFMEAQELHALTHVNAIGLKNKFDITWKQTKNIVQHCTQCQILHLATQEARVNPRGLCPNVLWQMDV MHVPSFGKLSFVHVTVDTYSHFIWATCOTGESTSHVKRHLLFCFPVMGVPEKVKTDNGPGYCSKAVQEFLNQWKITHTIGILYNSQG QAIIERTNRTLKAQLVKQKKKKKKKKK

SEO ID 54:

 $\label{eq:construction} QKNESSKLSIT*LKEQSWLPSLQC*QDFNQSINIVSDSAYVVQATKDIERALIKYIMDDQLNPLFNLLQQNVRKRNFPFYITHIRAH$ TNLPGPLTKANEQADLLVSSAFMEAQELHALTHVNAIGLKNKFDITWKQTKNIVQHCTQCQILHLATQEARVNPRGLCPNVLWQMDV MHVPSFGKLSFVHVTVDTYSHFIWATCQTGESTSHVKRHLLSCFPVMGVPEKKKKKKKKK

SEQ ID 55: 40

QKNESSKLSIT*LKEQSWLPSLQC*QDFNQSINIVSDSAYVVQATKDIERALIKYIMDDQLNPLFNLLQQNVRKRNFPFYITHIRAH TNLPGPLTKANEQADLLVSSAFIEAQELHALTHVNAIGLKNKFDITWKQTKNIVQHCTQCQILHLATQEARVNPRGLCPNVLWQMDV MHVPSFGKLSFVHVTVDTYSHFIWATCQTGESTSHVKRHLLSCFPVMGVPEKVKTDNGPGYCSKAVQKFLNQWKITHT IGILYNSQGQAIIERTNRTLKAQLVKQKKKKKKKKTCRPPR

SEO ID 56:

EETQVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTLLDSIAHGNRLTPYDWEILAKSSLSSSQYLQFKTWWIDGVQEQVR 45 KNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEAIEQVRAICLRAWGKIQDPGTAFPINSIRQGSKEPYPDFVARLQDAAQKSIT DDNARKVIVELMAYENANPECQSAIKPLKGKVPAGVDVITEYVKACDGIGGAMHKAMLMAQAMRGLTLGGQVRTFGKKCYNCGQIGH RKRSCPGLNKQNIINQAITAKNKKPSGLCPKCGKAKHWANQCHSKFDKDGQPLSGNRKRGQPQAPQQTGAFPVKLFVPQGFQGQQPL QKIPPLQGVSQLQQSNSCPAPQQAAPQ

SEQ ID 57:

EETQVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTVLDSIAHGNRLTPYDWEILAKSSLSSSQYLQFKTWWIDGVQEQVR 50 KNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEAIEQVRAICLRAWGKIQDPGTAFPINSIRQGSKEPYPDFVARLQDAAQKSIT DDNARKVIVELMAYENANPECQSAIKPLKGKVPAGVDVITEYVKACDGIGGAMHKAMLMAQAMRGLTLGGQVRTFGKKCYNCGOIGH LKRSCPGLNKQNIINQAITEKKKKKKK

SEQ ID 58:

QDFNQSINIVSDSAYVVQATKDIERALIKYIMDDOLNPLFNLLQQNVRKRNFPFYITHIRAHTNLPGPLTKANEQADLLVSSAFMEA QELHALTHVNAIGLKNKFDITWKQTKNIVQHCTQCQILHLATQEARVNPRGLCPNVLWQMDVMHVPSFGKLSFVHVTVDTYSHFIWA TCQTGESTSHVKRHLLSCFPVMGVPEKVKTDNGPGYCSKAVQKFLNQWKITHTIGILYNSQGQAIIERTNRTLKAQLVKQKKKKKKK KTCRPPR

SEQ ID 59: TAGGCCTTTGAGGGA SEQ ID 60: CATTAGAAAAAGGACATTG 5 SEQ ID 61: TTGGAATTCTGTTTGTA SEQ ID 62: TAACTGAGCCATTAAT SEQ ID 63: 10 AGCCATGGTCCCCTTTAATTA SEQ ID 64: TTTTACCACACCAGCCT SEQ ID 65: TTGTCAGCTCAAGCT 15 **SEQ ID 66:** TACATCGTTCACTAT SEQ ID 67: TTAAAAGCATTAAAT **SEQ ID 68:** 20 AGAAGTCCCAATTGAGG SEQ ID 69: GGTCTTGCCGATTTT SEQ ID 70: ACAATCGTTACCACA 25 SEQ ID 71: AAAAGAATGAGTCAT SEQ ID 72: CAGTATCACTTGACT SEQ ID 73: 30 TTTTAATCAGTCTATTAACATTG SEQ ID 74: AAAGGATATTGAGAGA SEQ ID 75: CCTAATCAAATACATT 35 SEQ ID 76: CGCTGTTTAATTTGT SEQ ID 77: TGCATTCATGGAAGCA SEQ ID 78: 40 ACTCAGGAGGCAAGA SEQ ID 79: TTAAGAGACATTTATT SEQ ID 80: TAAAGCAGTTCAAAAA 45 SEQ ID 81: AATAGGAATTCTCTA SEQ ID 82: AAAGCTCAATTGGTTA SEQ ID 83: 50 TAGGAGGACAAGTTAGAACATTTGG SEQ ID 84: AAAATGTTATAATTGTGGTCAAAT

5	SEQ ID 85: ATGGGGCAAACTAAAAGTAAAATTAAAAGTAAATATGCCTCTTATCTCAGCTTTATTAAAATTCTTTTTAAAAAGAGGGGGGAGTTAAA GTATCTACAAAAAATCTAATCAAGCTATTTCAAATAATAGAACAATTTTGCCCATGGTTTCCAGAACAAGGAACTTTAGATCTAAAA GATTGGAAAAGAATTGGTAAGGAACTAAAACAAGCAGGTAGGAAGGGTAATATCATTCCACTTACAGTATGGAATGATTGGGCCATT ATTAAAGCAGCTTTAGAACCATTTCAAACAGAAGAAGAAGATAGCGTTTCAGTTTCTGATGCCCCTGGAAGCTGTATAATAGATTGTAAT GAAAACACAAGGAAAAAATCCCAGAAAGAAACGGAAGGATAGCGTTTACATTGCGAATATGTAGCAGAGCCGGTAATGGCTCAGTCAACGCAA AATGTTGACTATAATCAATTACAGGAGGGTGATATATCCTGAAACGTTAAAATTAGAAGGACAGGTCCAGAATTAGTGGGGCCATCA GAGTCTAAACCACGAGGCCACAAGTCCTCTTCCAGCAGGTCGGGTCAGGTGCCTGTAACATTACAACCTCAAAAGCAGGTTAAAAGAAAATAG ACCCAACCGCCAGTAGCCTATCAATACTGGCCTCCGGCTGAACTTCCAGTATCGGCCACCCCCAGAAAGTCAGTATGGATATCCAGGA
10	ATGCCCCCAGCACCACAGGGCAGGGCGCCATACCCTCAGCCGCCCACTAGGAGACTTAATCCTACGGCACCACCTAGTAGACAGGGT
15	
20	
25	
30	
35	
40	
43 50	
55	

CCTGGAGAAGGAGCCCCAAGAGGGAGAGCCTCCCACAGTTGAGGCCAGATACAAGTCTTTTTCGATAAAAAAAGCTAAAAAGATATGAAA GAGGGAGTAAAACAGTATGGACCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTCATGGACATAGACTCATTCCTTAT GATTGGGAGATTCTGGCAAAATCGTCTCTCTCACCCTCTCAATTTTACAATTTAAGACTTGGTGGATTGATGGGGTACAAGAACAG GTCCGAAGAAATAGGGCTGCCAATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGAATAGGTCAAAATTGGAGTACTATTAGT CAACAAGCATTAATGCAAAATGAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCCTGGGAAAAAATCCAAGACCCAGGAAGT ACCTGCCCCTCATTTAATACAGTAAGACAAGGTTCAAAAGAGCCCTATCCTGATTTTGTGGCAAGGCTCCAAGATGTTGCTCAAAAG TCAATTGCTGATGAAAAAGCCCGTAAGGTCATAGTGGAGTTGATGGCATATGAAAACGCCCAATCCTGAGTGTCAATCAGCCATTAAG CCATTAAAAGGAAAGGTTCCTGCAGGATCAGATGTAATCTCAGAATATGTAAAAGCCTGTGATGGAATCGGAGGAGCTATGCATAAA **GCTATGCTTATGGCTCAAGCAATAACAGGAGTTGTTTTAGGAGGACAAGTTAGAACATTTGGAAGAAAATGTTATAATTGTGGTCAA** ATTGGTCACTTAAAAAAGAATTGCCCAGTCTTAAATAAACAGAATATAACTATTCAAGCAACTACAACAGGTAGAGAGCCACCTGAC TTATGTCCAAGATGTAAAAAAGGAAAACATTGGGCTAGTCAATGTCGTTCTAAATTTGATAAAAATGGGCAACCATTGTCGGGAAAC GAGCAAAGGGGCCAGCCTCAGGCCCCACAAAACTGGGGCATTCCCCAATTCAGGCCATTTGTTCCTCAGGGTTTTCAGGGACAACAA CCCCCACTGTCCCAAGTGTTTCAGGGAATAAGCCAGTTACCACAATACAACAATTGTCCCCCGCCACAAGCGGCAGTGCAGCAG **SEQ ID 86:**

5

10

15

20

ATGGGCAACCATTGTCGGGAAACGAGCAAAGGGGCCAGCCTCAGGCCCCACAACAACTGGGGCATTCCCAATTCAGCCATTTGTTC CTCAGGGTTTTCAGGGACAACAACCCCCCACTGTCCCCAAGTGTTTCAGGGAATAAGCCAGTTACCACAATACAACAATTGTCCCCCGC CACAAGCGGCAGTGCAGCAGTAGATTTATGTACTATACAAGCAGTCTCTCTGCTTCCAGGGGAGCCCCCACAAAAAAACCCCCCACAGG GGTATATGGACCCCTGCCTAAGGGGACTGTAGGACTAATCTTGGGACGATCAAGTCTAAATCTAAAAGGAGTTCAAATTCATACTAG TGTGGTTGATTCAGACTATAAAGGCGAAATTCAATTGGTTATTAGCTCTTCAATTCCTTGGAGTGCCAGTCCAAGAGACAGGATTGC TCAATTATTACTCCTGCCATACATTAAGGGTGGAAATAGTGAAATAAAAAGAATAGGAGGGCTTGGAAGCACTGATCCAACAGGAAA GGCTGCATATTGGGCAAGTCAGGTCTCAGAGAACAGACCTGTGTGTAAGGCCATTATTCAAGGAAAACAGTTTGAAGGGTTGGTAGA CACTGGAGCAGATGTCTCTATCATTGCTTTAAATCAGTGGCCAAAAAATTGGCCTAAACAAAAGGCTGTTACAGGACTTGTCGGCAT AGGCACAGCCTCAGAAGTGTATCAAAGTACGGAGATTTTACATTGCTTAGGGCCAGATAATCAAGAAAGTACTGTTCAGCCAATGAT TACTTCAATTCCTCTTAATCTGTGGGGTCGAGATTTATTACAACAATGGGGTGCGGAAATCACCATGCCCGCTCCATCATATAGCCC CACGAGTCAAAAAATCATGACCAAGATGGGATATATACCAGGAAAGGGGACTAGGGAAAAATGAAGATGGCATTAAAATTCCAGTTGA GGCTAAAATAAATCAAGAAAGAGAAGGAATAGGGAATCCTTGC

SEQ ID 87:

TCCTAAACCCATACCATTAACTTGGAAAAACAGAAAAACCAGTGTGGGTAAATCAGTGGCCGCTACCAAAAACAAAAACTGGAGGCTTT 25 ACATTTATTAGCAAATGAACAGTTAGAAAAGGGTCATATTGAGCCTTCGTTCTCACCTTGGAATTCCCCTGTGTTTGTAATTCAGAA GAAATCAGGCAAATGGCGTATGTTAACTGACTTAAGGGCTGTAAACGCCGTAATTCAACCCATGGGGCCTCTCCAACCCGGGTTGCC CTCTCCGGCCATGATCCCAAAAGATTGGCCTTTAATTATTAATTGATCTAAAGGATTGCTTTTTTACCATCCCTCTGGCAGAGCAGGA TTGCGAAAAATTTGCCTTTACTATACCAGCCATAAATAATAAAGAACCAGCCACCAGGTTTCAGTGGAAAGTGTTACCTCAGGGAAT GCTTAATAGTCCAACTATTTGTCAGACTTTTGTAGGTCGAGCTCTTCAACCAGTTAGAGAAAAGTTTTCAGACTGTTATATTATTCA 30 TGGACTGGCAATAGCATCTGATAAGATCCAAAACCTCTACTCCTTTTCATTATTTAGGGATGCAGATAGAAAAATAGAAAAATTAAGCC ACAAAAAATAGAAAAAGAAAAGACACATTAAAAAACACTAAATGATTTTCAAAAAATTACTAGGAGATATTAATTGGATTCGGCCAAC **TCTAGGCATTCCTACTTATGCCATGTCAAATTTGTTCTCTATCTTAAGAGGAGACTCAGACTTAAATAGTAAAAGAATGTTAACCCC** AGTTAAGACTTTTACATTGTACTTGGATCAAATAGCTACATTAATCGGTCAGACAAGATTACGAATAATAAAAATTATGTGGGAATGA 35 CCCAGACAAAATAGTTGTCCCTTTAACCAAGGAACAAGTTAGACAAGCCTTTATCAATTCTGGTGCATGGAAGATTGGTCTTGCTAA TTTTGTGGGAATTATTGATAATCATTACCCAAAAACAAAGATCTTCCAGTTCTTAAAAATTGACTACTTGGATTCTACCTAAAATTAC CAGACGTGAACCTTTAGAAAATGCTCTAACAGTATTTACTGATGGTTCCAGCAATGGAAAAGCAGCTTACACAGGACCGAAAGAACG TATTATATCAGATTCTGCATATGTAGTACAGGCTACAAGGGATGTTGAGACAGCTCTAATTAAATATAGCATGGATGATCAGTTAAA 40 AGGGCCTTTGACTAAAGCAAATGAACAAGCTGACTTACTGGTATCATCTGCACTCATAAAAGCACAAAGAACTTCATGCTTTGACTCA TGTAAATGCAGCAGGATTAAAAAAACAAATTTGATGTCACATGGAAACAGGCAAAAGATATTGTACAACATTGCACCCAGTGTCAAGT CTTACACCTGCCCACTCAAGAGGCAGGAGTTAATCCCCAGAGGTCTGTGTCCTAATGCATTATGGCAAATGGATGTCACGCATGTACC TTCATTTGGAAGATTATCATATGTTCACGTAACAGTTGATACTTATTCACATTTCATATGGGCAACTTGCCAAACAGGAGAAAGTAC TTCCCATGTTAAAAAACATTTATTGTCTTGTTTTGCTGTAATGGGAGTTCCAGAAAAAATCAAAACTGACAATGGACCAGGATATTG TAGTAAAGCTTTTCCAAAAATTCTTAAGTCAGTGGAAAATTTCACATACAACAGGAATTCCTTATAATTCCCCAAGGACAGGCCATAGT 45 **ACTTAATCTAGCACTCTATACTTTAAAATTTTTTTAAACATTTATAGAAAATCAGACTACTACTTCTGCAGAACAACATCTTACTGGTAA** AAAGAACAGCCCACATGAAGGAAAACTAATTTGGTGGAAAGATAATAAAAATAAGACATGGGAAATAGGGAAGGTGATAACGTGGGG GAGAGGTTTTGCTTGTGTTTCACCAGGAGAAAATCAGCTTCCTGTTTGGATACCCACTAGACATTTGAAGTTCTACAATGAACCCAT CAGAGATGCAAAGAAAAGCACCTCCGCGGAGACGGAGACATCGCAATCGAGCACCGTTGACTCACAAGATGAACAAAATGGTGACGT CAGAAGAACAGATGAAGTTGCCATCCACCAAGAAGGCAGAGCCGCCGACTTGGGCACAACTAAGAAGCTGACGCAGTTAGCTACAA AATATCTAGAGAACACAAAAGGTGACACAAAACCCCAGAGAGTATGCTGCTGCAGGCCTTGATGATTGTATCAATGGTGGTAAGTCTCC 50 CTATGCCTGCAGGAGCAGCTGCAGC

SEQ ID 88:

55

ATGAACCCATCAGAGATGCAAAGAAAAGCACCTCCGCGGAGACGGGAGACATCGCAATCGAGCACCGTTGACTCACAAGATGAACAAA ATGGTGACGTCAGAAGAACAGATGAAGTTGCCATCCACCAAGAAGGCAGAGCCGCCAACTTGGGCACAACTAAAGAAGCTGACGCAG TTAGCTACAAAATATCTAGAGAACACAAAAGGTGACACAAACCCCCAGAGAGTATGCTGCTGCCTTGCAGCCTTGATGATTGTATCAATGGTG GTAAGTCTCCCTATGCCTGCAGGAGCAGCTGCAGCTAACTATACCTACTGGGCCTATGTGCCCTTTCCCGCCCTTAATTCGGGCAGTC ACATGGATGGATAATCCTACAGAAGTATATGTTAATGATAGTGTATGGGTACCTGGCCCCATAGATGATCGCTGCCCAGACCT GAGGAAGAAGGGATGATGATAAATATTTCCATTGGGTATCATTATCCTCCTATTTGCCTAGGGAGAGCACCAGGATGTTTAATGCCT

5

GCAGTCCAAAATTGGTTGGTAGAAGTACCTACTGTCAGTCCCATCTGTAGATTCACTTATCACATGGTAAGCGGGATGTCACTCAGG CCACGGGTAAATTATTACAAGACTTTTCTTATCAAAGATCATTAAAATTTAGACCTAAAGGGAAACCTTGCCCCAAGGAAATTCCC AAAGAATCAAAAAATACAGAAGTTTTAGTTTGGGAAGAATGTGTGGCCAATAGTGCGGTGATATTACAAAACAATGAATTCGGAACT ATTATAGATTGGGCACCTCGAGGTCAATTCTACCACAATTGCTCAGGACAAACTCAGTCGTGTCCAAGTGCACAAGTGAGTCCAGCT GTTGATAGCGACTTAACAGAAAGTTTAGACAAACATAAGCATAAAAAATTGCAGTCTTTCTACCCTTGGGAAAGGGAGAAAAAGGA ATCTCTACCCCAAGACCAAAAAATAGTAAGTCCTGTTTCTGGTCCTGAACATCCAGAATTATGGAGGCCTTACTGTGGCCTCACACCAC ATTAGAATTTGGTCTGGAAATCAAACTTTAGAAACAAGAGACCGTAAGCCATTTTATACTATTGACCTGAATTCCAGTCTAACAGTT CCTTTACAAAGTTGCGTAAAGCCCCCCTTATATGCTAGTAGGAAATATAGTTATTAAACCAGACTCCCAGACTATAAACCTGTGAA AATTGTAGATTGCTTACTTGCATTGATTCAACTTTTAATTGGCAACACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATC CCTGTGTCCATGGACCGACCGTGGGAGGCCTCGCCATCCGTCCATATTTTGACTGAAGTATTAAAAGGTGTTTTAAATAGATCCAAA AGATTCATTTTACTTTAATTGCAGTGATTATGGGATTAATTGCAGTCACAGGCTGCTGCTGCAGGAGGAGTTGCATTGCACTCT 10 TCTGTTCAGTCAGTAAACTTTGTTAATGATTGGCAAAAAAATTCTACAAGATTGTGGAAATCCACAATCTAGTATTGATCAAAAAATTG GACTGGAATACGTCAGATTTTTGTATTACACCCCCAAATTTATAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAG GGAAGAGAAGATAATCTCACTTTAGACATTTCCAAATTAAAAGAACAAATTTTCGAAGCATCAAAAGCCCCATTTAAATTTGGTGCCA GGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTTAAGACCATTGGAAGTACTACGATT ATAAATCTCATATTAATCCTTGTGTGCCCTGTTTTGTCTGTTAGTCTGCAGGTGTACCCCAACAGCTCCGAAGAGAGACAGCGACCA **SEQ ID 89:** 15 AGTTCTACAATGAACCCATCAGAGATGCAAAGAAAAGCACCTCCGCGGAGACGGAGACATCGCAATCGAGCACCGTTGACTCACAAG ATGAACAAAAATGGTGACGTCAGAAGAACAAGATGAAGTTGCCATCCACCAAGAAGGCAGAGCCGCCAACTTGGGCACAACTAAAGAAG CTGACGCAGTTAGCTACAAAATATCTAGAGAACACAAAGGTGACACAAACCCCCAGAGAGTATGCTGCTGCAGCCTTGATGATTGTA TCAATGGTGGTAAGTCTCCCTATGCCTGCAGGA **SEQ ID 90:** 20 TCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGCCATGA SEO ID 91: GTATCTACAAAAAAATCTAATCAAGCTATTTCAAATAATAGAACAATTTTGCCCCATGGTTTCCAGAACAAGGAACTTTAGATCTAAAA GATTGGAAAAGAATTGGTAAGGAACTAAAACAAGCAGGTAGGAAGGGTAATATCATTCCACTTACAGTATGGAATGATTGGGCCATT ATTAAAGCAGCTTTAGAACCATTTCAAACAGAAGAAGATAGCGTTTCAGTTTCTGATGCCCCTGGAAGCTGTATAATAGATTGTAAT 25 GAAAACACAAAGGAAAAAATCCCAGAAAGAAACGGAAGGTTTACATTGCGAATATGTAGCAGAGCCGGTAATGGCTCAGTCAACGCAA AATGTTGACTATAATCAATTACAGGAGGTGATATATCCTGAAACGTTAAAATTAGAAGGAAAAGGTCCAGAATTAGTGGGGCCATCA GAGTCTAAACCACGAGGCACAAGTCCTCTTCCAGCAGGTCAGGTGCCTGTAACATTACAACCTCAAAAGCAGGTTAAAGAAAATAAG ACCCAACCGCCAGTAGCCTATCAATACTGGCCTCCGGCTGAACTTCAGTATCGGCCACCCCCAGAAAGTCAGTATGGATATCCAGGA ATGCCCCCAGCACCACAGGGCAGGGCGCCATACCCTCAGCCGCCCACTAGGAGACTTAATCCTACGGCACCACCTAGTAGACAGGGT 30 CCTGGAGAAGGAGCCCAAGAGGAGAGAGCCTCCCACAGTTGAGGCCAGATACAAGTCTTTTTCGATAAAAAAGCTGAAAGATATGAAA GAGGGAGTAAAACAGTATGGACCCCAACTCCCCTTATATGAGGACATTATTAGATTCCATTGCTCATGGACATAGACTCATTCCTTAT GATTGGGAGATTCTGGCAAAATCGTCTCTCTCACCCTCTCAATTTTTACAATTTTAGACTTGGTGGATTGATGGGGGTACAAGAACAG GTCCGAAGAAATAGGGCTGCCAATCCTCCAGTTAACATAGATGCAGATCAACTATTAGGAATAGGTCAAAATTGGAGTACTATTAGT CAACAAGCATTAATGCAAAAATGAGGCCATTGAGCAAGTTAGAGCTATCTGCCTTAGAGCCTGGGAAAAAATCCAAGACCCAGGAAGT ACCTGCCCCCCAAAAAAGACAAAGGTTCAAAAGAGCCCTATCCTGATTTTGTGGCAAGGCTCCAAGATGTTGCTCAAAAG 35 TCAATTGCTGATGAAAAAAGCCCGTAAGGTCATAGTGGAGTTGATGGCATATGAAAACGCCAATCCTGAGTGTCAATCAGCCATTAAG CCATTAAAAGGAAAGGTTCCTGCAGGATCAGATGTAATCTCAGAATATGTAAAAGCCTGTGATGGAATCGGAGGAGCTATGTATAAA ${\tt GCTATGCTTATGGCTCAAGCAATAACAGGAGTTGTTTTAGGAGGACAAGTTAGAACATTTGGAAGAAAATGTTATAATTGTGGTCAA}$ TTATGTCCAAGATGTAAAAAAAGGAAAACATTGGGCTAGTCAATGTCGTTCTAAATTTGATAAAAATGGGCAACCATTGTCGGGAAAC GAGCAAAGGGGCCCAGCCTCAGGCCCCACAACAAACTGGGGGCATTCCCAATTCAGCCATTTGTTCCTCAGGGTTTTCAGGGACAACAA CCCCCACTGTCCCAAGTGTTTCAGGGAATAAGCCAGTTACCACAATACAACAATTGTCCCCCGCCACAAGCGGCAGTGCAGCAGTAG 40 SEQ ID 92: MGQTKSKIKSKYASYLSFIKILLKRGGVKVSTKNLIKLFQIIEQFCPWFPEQGTLDLKDWKRIGKELKQAGRKGNIIPLTVWNDWAI IKAALEPFQTEEDSVSVSDAPGSCIIDCNENTRKKSQKETEGLHCEYVAEPVMAQSTQNVDYNQLQEVIYPETLKLEGKGPELVGPS ESKPRGTSPLPAGQVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPESQYGYPGMPPAPQGRAPYPQPPTRRLNPTAPPSRQG SKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGAQEGEPPTVEARYKSFSIKKLKDMKEGVKOYGPNSPYMRTLLDSIAHGHRLIPY DWEILAKSSLSPSQFLQFKTWWIDGVQEQVRRNRAANPPVNIDADQLLGIGQNWSTISQQALMQNEAIEQVRAICLRAWEKIQDPGS 45 TCPSFNTVRQGSKEPYPDFVARLQDVAQKSIADEKARKVIVELMAYENANPECQSAIKPLKGKVPAGSDVISEYVKACDGIGGAMYK AMLMAQAITGVVLGGQVRTFGRKCYNCGQIGHLKKNCPVLNKQNITIQATTTGREPPDLCPRCKKGKHWASQCRSKFDKNGQPLSGN EQRGQPQAPQQTGAFPIQPFVPQGFQGQQPPLSQVFQGISQLPQYNNCPPPQAAVQQ **SEQ ID 93:** ATGTTAACTGACTTAAGGGCTGTAAACGCCGTAATTCAACCCATGGGGCCTCTCCAACCCGGGTTGCCCTCTCCGGCCATGATCCCA AAAGATTGGCCTTTAATTATAATTGATCTAAAGGATTGCTTTTTTACCATCCCTCTGGCAGAGCAGGATTGCGAAAAATTTGCCTTT 50 ACTATACCAGCCATAAATAATAAAGAACCAGCCACCAGGTTTCAGTGGAAAGTGTTACCTCAGGGAATGCTTAATAGTCCAACTATT **TGTCAGACTTTTGTAGGTCGAGCTCTTCAACCAGTTAGAGAAAAGTTTTCAGACTGTTATATTATTCATTGTATTGATGATATTTTA** TGTGCTGCAGAAACGAAAGATAAATTAATTGACTGTTATACATTTCTGCAAGCAGAGGTTGCCAATGCTGGACTGGCAATAGCATCT GATAAGATCCAAAACCTCTACTCCTTTTCATTATTAGGGATGCAGATAGAAAATAGAAAAATTAAGCCACAAAAAAATAGAAATAAGA AAAGACACATTAAAAACACTAAATGATTTTCAAAAATTACTAGGAGAATATTAATTGGATTCGGCCAACTCTAGGCATTCCTACTTAT

GCCATGTCAAATTTGTTCTCTATCTTAAGAGGAGACTCAGACTTAAATAGTAAAAGAATGTTAACCCCCAGAGGCAACAAAAGAAATT 55

TACTTGGATCAAATAGCTACATTAATCGGTCAGACAAGATTACGAATAATAAAATTATGTGGGAATGACCCAGACAAAATAGTTGTC AATCATTACCCAAAAACAAAGATCTTCCAGTTCTTAAAATTGACTACTTGGATTCTACCTAAAATTACCAGACGTGAACCTTTAGAA AATGCTCTAACAGTATTTACTGATGGTTCCAGCAATGGAAAAGCAGCTTACACAGGACCGAAAGAACGAGTAATCAAAACTCCATAT TATGTAGTACAGGCTACAAGGGATGTTGAGACAGCTCTAATTAAATATAGCATGGATGATCAGTTAAACCAGCTATTCAATTTATTA CAACAAACTGTAAGAAAAGAAATTTCCCATTTTATATTACACACATATTCGAGCACACACTAATTTACCAGGGCCTTTGACTAAAGCA AATGAACAAGCTGACTTACTGGTATCATCTGCACTCATAAAAGCACAAGAACTTCATGCTTTGACTCATGTAAATGCAGCAGGATTA AAAAACAAATTTGATGTCACATGGAAACAGGCAAAAGATATTGTACAACATTGCACCCAGTGTCAAGTCTTACACCTGCCCACTCAA GAGGCAGGAGTTAATCCCAGAGGTCTGTGTCCTAATGCATTATGGCAAATGGATGTCACGCATGTACCTTCATTTGGAAGATTATCA TATGTTCACGTAACAGTTGATACTTATTCACATTTCATATGGGCAACTTGCCAAACAGGAGAAAGTACTTCCCCATGTTAAAAAAACAT TTATTGTCTTGTTTTGCTGTAATGGGAGTTCCAGAAAAAATCAAAACTGACAATGGACCAGGATATTGTAGTAAAGCTTTCCAAAAA TTCTTAAGTCAGTGGAAAATTTCACATACAACAGGAATTCCTTATAATTCCCAAGGACAGGCCATAGTTGAAAGAACTAATAGAACA CTCAAAACTCAATTAGTTAAACAAAAAAAAAGAGGGGGAGACAGTAAGGAGTGTACCACTCCTCAGATGCAACTTAATCTAGCACTCTAT ACTTTAAATTTTTTAAACATTTATAGAAATCAGACTACTACTTCTGCAGAACAACATCTTACTGGTAAAAAGAACAGCCCACATGAA **GGAAAAACTAATTTGGTGGAAAGATAGTAAAAATAAGACATGGGAAATAGGGAAGGTGATAACGTGGGGGGAGAGGTTTTGCTTGTGTT** TCACCAGGAGAAAATCAGCTTCCTGTTTGGATACCCACTAGACATTTGAAGTTCTACAATGAACCCATCAGAGATGCAAAGAAAAGC ACCTCCGCGGAGACGGAGACATCGCAATCGAGCACCGTTGACTCACAAGATGAACAAAATGGTGACGTCAGAAGAACAGATGAAGTT GCCATCCACCAAGAAGGCAGAGCCGCCAACTTGGGCACAACTAAGAAGCTGACGCAGTTAGCTACAAAAATATCTAGAGAACACAAA GGTGACACAAAACCCCAGAGAGTATGCTGCTTGCAGCCTTGATGATTGTATCAATGGTGGTAAGTCTCCCCTATGCCTGCAGGAGCAGC TGCAGCTAA

SEQ ID 94:

5

10

15

MLTDLRAVNAVIQPMGPLQPGLPSPAMIPKDwPLIIIDLKDCFFTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTI 20 CQTFVGRALQPVREKFSDCYIIHCIDDILCAAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIR KDTLKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLTPEATKEIKLVEEKIQSAQINRIDPLAPLQLLIFATA HSPTGIIIQNTDLVEWSFLPHSTVKTFTLYLDQIATLIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFVGIID NHYPKTKIFQFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAVITVLQDFDQPINIISDSA YVVQATRDVETALIKYSMDDQLNQLFNLLQQTVRKRNFPFYITHIRAHTNLPGPLTKANEQADLLVSSALIKAQELHALTHVNAAGL KNKFDVTWKQAKDIVQHCTQCQVLHLPTQEAGVNPRGLCPNALWQMDVTHVPSFGRLSYVHVTVDTYSHFIWATCQTGESTSHVKKH 25 LLSCFAVMGVPEKIKTDNGPGYCSKAFQKFLSQWKISHTTGIPYNSQGQAIVERTNRTLKTQLVKQKEGGDSKECTTPQMQLNLALY TLNFLNIYRNQTTTSAEQHLTGKKNSPHEGKLIWWKDSKNKTWEIGKVITWGRGFACVSPGENQLPVWIPTRHLKFYNEPIRDAKKS TSAETETSQSSTVDSQDEQNGDVRRTDEVAIHQEGRAANLGTTKEADAVSYKISREHKGDTNPREYAACSLDDCINGGKSPYACRSS

cs SEQ ID 95:

ATGCAAAGAAAAGCACCTCCGCGGAGACGGGAGACATCGCAATCGAGCACCGTTGACTCACAAGATGAACAAAATGGTGACGTCAGAA 30 GAACAGATGAAGTTGCCATCCACCAAGAAGGCAGAGCCGCCAACTTGGGCACAACTAAAGAAGCTGACGCAGTTAGCTACAAAATAT CCTACAGAAGTATATGTTAATGATAGTGTATGGGTACCTGGCCCCATAGATGATCGCTGCCCTGCCAAACCTGAGGAAGAAGGGATG ATGATAAATATTTCCATTGGGTATCATTATCCTCCTATTTGCCTAGGGAGGAGCACCAGGATGTTTAATGCCTGCAGTCCAAAATTGG TTGGTAGAAGTACCTACTGTCAGTCCCATCTGTAGATTCACTTATCACATGGTAAGCGSGATGTCACTCAGGCCACGGGTAAATTAT TTACAAGACTTTTTCTTATCAAAGATCATTAAAAATTTAGACCTAAAGGGAAACCTTGCCCCAAGGAAATTCCCAAAGAATCAAAAAAAT 35 ACAGAAGTTTTAGTTTGGGAAGAATGTGTGGCCAATAGTGCGGTGATATTACAAAACAATGAATTCGGAACTATTATAGATTGGGCA CCTCGAGGTCAATTCTACCACAATTGCTCAGGACAAACTCAGTCGTGTCAAAGTGCACAAGTGAGTCCAGCTGTTGATAGCGACTTA ACAGAAAGTTTAGACAAACATAAGCATAAAAAATTGCAGTCTTTCTACCCTTGGGAAATGGGGAGAAAAGGAATCTCTACCCCAAGA CCAAAAATAGTAAGTCCTGTTTCTGGTCCTGAACATCCAGAATTATGGAGGCTTACTGTGGCCTCACACCACATTAGAATTTGGTCT ${\tt GGAAATCAAACTTTAGAAACAAGAGATCGTAAGCCATTTTATACTATTGACCTGAATTCCAGTCTAACAGTTCCTTTACAAAGTTGC$ GTAAAGCCCCCCTTATATGCTAGTTGTAGGAAATATAGTTATTAAACCAGACTCCCAGACTATAACCTGTGAAAATTGTAGATTGCTT 40 CGACCGTGGGAGGCCTCGCCATCCGTCCATATTTTGACTGAAGTATTAAAAGGTGTTTTAAATAGATCCAAAAGATTCATTTTTACT AACTTTGTTAATGATTGGCAAAAAAATTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAAATTGGCAAATCAAATTAAT GATCTTAGACAAACTGTCATTTGGATGGGAGACAGACTCATGAGCTTAGAACATCGTTTCCAGTTACAATGTGACTGGAATACGTCA 45 CTCACTTTAGACATTTCCAAATTAAAAGAACAAATTTTCGAAGCATCAAAAGCCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATT GCAGGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTTAAGACCATTGGAAGTACTACGATTATAAATCTCATATTA ATCCTTGTGTGCCTGTTTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGCCATG

- SEQ ID 96:
- MQRKAPPRRRRHRNRAPLTHKMNKMVTSEEQMKLPSTKKAEPPTWAQLKKLTQLATKYLENTKVTQTPESMLLAALMIVSMVVSLPM PAGAAAANYTYWAYVPFPPLIRAVTWMDNPTEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMPAVQNW LVEVPTVSPICRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEFGTIIDWA PRGQFYHNCSGQTQSCQSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIVSPVSGPEHPELWRLTVASHHIRIWS GNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKPDSQTITCENCRLLTCIDSTFNWQHRILLVRAREGVWIPVSMD RPWEASPSVHILTEVLKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTRLWNSQSSIDQKLANQIN DLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAI AGVADGLANLNPVTWVKTIGSTTIINLILILVCLFCLLLVCRCTQQLRRDSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV SEQ ID 97:

GTATCTACAAAAAATCTAATCAAGCTATTTCAAAATAATAGAACAATTTTGCCCATGGTTTCCAGAACAAGGAACTTTAGATCTAAAA GATTGGAAAAGAATTGGCGAGGAACTAAAACAAGCAGGTAGAAAGGGTAATATCATTCCACTTACAGTATGGAATGATTGGGCCATT attaaagcagctttagaaccatttcaaacaaaagatagcgtttcagtttctgatgcccctggaagctgtgtaatagattgtaat GAAAAGACAGGGGGGAGAAAATCCCCAGAAAGAAACAGGAAAGTTTACATTGCGAATATGTAACAGAGCCAGTAATGGCTCAGTCAACGCAA AATGTTGACTATAAATCAATTACAGGGGGGTGATATATCCTGAAACGTTAAAAATTAGAAGGAAAAGGTCCAGAATTAGTGGGGCCATCA GAGTCTAAACCACGAGGGCCAAGTCCTCTTCCAGCAGGTCAGGTGCCCGTAACATTACAACCTCAAACGCAGGTTAAAGAAAATAAG ACCCAACCGCCAGTAGCTTATCAATACTGGCCGCCGGCTGAACTTCAGTATCTGCCACCCCCAGAAAGTCAGTATGGATATCCAGGA ATGCCCCCAGCACTACAGGGCAGGGCGCCCATATCCTCAGCCGCCCACTGTGAGACTTAATCCTACAGCATCACGTAGTGGACAAGGT

- GGTACACTGCACGCAGTCATTGATGAAGCCAGAAAACAGGGAGATCTTGAGGCATGGCGGTTCCTGGTAATTTTACAACTGGTACAG GCCGGGGAAGAGACTCAAGTAGGAGCGCCTGCCCGAGCTGAGACTAGATGTGAACCTTTCACCATGAAAATGTTAAAAGATATAAAG 10 GAAGGAGTTAAACAATATGGATCCAACTCCCCCTTATATAAGAACATTATTAGATTCCATTGCTCATGGAAATAGACTTACTCCTTAT GACTGGGAAAGTTTGGCCAAATCTTCCCCTTTCATCCTCTCAGTATCTACAGTTTAAAAACCTGGTGGATTGATGGAGTACAAGAACAG **GTACGAAAAAATCAGGCTACTAAGCCCACTGTTAATATAGACGCAGACCAATTGTTAGGAACAGGTCCAAATTGGAGCACCATTAAC** CAACAATCAGTGATGCAGAATGAGGCTATTGAACAAGTAAGGGCTATTTGCCTCAGGGCCTGGGGAAAAATTCAGGACCCAGGAACA GCTTTCCCTATTAATTCAATTAGACAAGGCTCTAAAGAGCCATATCCTGACTTTGTGGCAAGATTACAAGATGCTGCTCAAAAGTCT ATTACAGATGACAATGCCCGAAAAGTTATTGTAGAATTAATGGCCTATGAAAATGCAAATCCAGAATGTCAGTCGGCCATAAAGCCA
- TTAAAAGGAAAAGTTCCAGCAGGAGTTGATGTAATTACAGAATATGTGAAGGCTTGTGATGGGGATTGGAGGAGCTATGCATAAGGCA 15 ATGCTAATGGCTCAAGCAATGAGGGGGGCTCACTCTAGGAGGACAAGTTAGAACATTTGGGAAAAAATGTTATAATTGTGGTCAAATC TAG

20 **SEQ ID 98:**

5

25

30

MGQTKSKTKSKYASYLSFIKILLKRGGVRVSTKNLIKLFQIIEQFCPWFPEQGTLDLKDWKRIGEELKQAGRKGNIIPLTVWNDWAI IKAALEPFQTKEDSVSVSDAPGSCVIDCNEKTGRKSQKETESLHCEYVTEPVMAQSTQNVDYNQLQGVIYPETLKLEGKGPELVGPS GTLHAVIDEARKOGDLEAWRFLVILQLVQAGEETQVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTLLDSIAHGNRLTPY $\label{eq:construction} DWESLAKSSLSSSQYLQFKTWWIDGVQEQVRKNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEAIEQVRAICLRAWGKIQDPGT$ $\label{eq:approx} AFPINSIRQGSKEPYPDFVARLQDAAQKSITDDNARKVIVELMAYENANPECQSAIKPLKGKVPAGVDVITEYVKACDGIGGAMHKA$

MLMAOAMRGLTLGGOVRTFGKKCYNCGOIGHLKRSCPVLNKONIINOAITAKNKKPSGLCPKCGKGKHWANOCHSKFDKDGOPLSGN RKRGQPQAPQQTGAFPVQLFVPQGFQGQQPLQKIPPLQGVSQLQQSNSCPAPQQAAPQ **SEQ ID 99:**

 ${\tt ctcagggttttcaaggacaacaaccacctacaggaaataccaccacctcagggagtcagccaattacaaccaatccaacagctgtcccg}$ ${\tt AGGGGTATATGGCCCGCTGCCAGAAGGGAGGGTAGGCCTTATTTTAGGGAGATCAAGTCTAAATTTGAAGGGAGTCCAAATTCATAC$ TGGGGTAATTTATTCAGATTATAAAGGGGGAATTCAGTTAGTGATCAGCTCCACTGTTCCCTGGAGTGCCAATCCAGGTGATAGAAT TGCTCAATTACTGCTTTTGCCTTATGTTAAAATTGGGGAAAACAAAACGGAAAGAACAGGAGGGTTTGGAAGTACCAACCCTGCAGG AAAAGCCACTTATTGGGCTAATCAGGTCTCAGAGGATAGACCCGTGTGTACAGTCACTATTCAGGGAAAGAGTTTGAAGGATTAGTG GATACCCAGGCTGATGTTTCTATCATCGGCATAGGCACCGCCTCAGAAGTGTATCAAAGTGCCATGATTTTACATTGTCTAGGATCT 35 GAGATTACTATCCCAGCCTCCCTATACAGCCCCAGGAATCAAAAAATCATGACTAAAAATGGGATAGCTCCCTAAAAAGGGACTAGGA

AAGAATGAAGATGGCATTAAAGTCCCAACTGAGGCTGAAAAAAATCAAAAAAGAAAAGGAATAGGGCATCCTTTTTAGAAGCGGTC ACTGTAGAGCCTCCAAAACCCATTCCATTAATTTGGGGGGGAAAAAAA

SEQ ID 100:

ATGGCATTAAAGTCCCAACTGAGGCTGAAAAAAATCAAAAAAGGAAAAGGAATAGGGCATCCTTTTTAGAAGCGGTCACTGTAGAGC 40 TACACTTATTAGCAAAGAAACAGTTAGAAAAAGGACATATTGAGCCTTCATTTTCGCCTTGGAATTCTCCTGTTTGTAATTCAGAAA AAATCCGGCAGATGGCGTATGCTAACTGACTTAAGAGCCATTAATGCCATAATTCAACCCATGGGGGGCTCTCCCATCCCGGTTGCCC **TCTCCAGCCATGGTCCCCTTTAATTATAATTGATCTGAAGGATTGCTTTTTTACCATTCCTCTGGCAAAAGAGGATTTTGAAAAATT** TGCTTTTACTATACCAGCCTAAATAAAAAAAAGAACCAGCCACCAGGTTTCAGTGGAAAGTATTGCCTCAGGGAATGCTTAATAATTCA ACTATTTGTCAGACTTTCATAGCTCAAGCTCTGCAACCAGTTAGAGACAAGTTTTCAGACTGTTATATCGTTCATTATGTTGATATT 45 CTGATAAGATTCAAACCTCTCCTCCTTTCCATTACTTGGGAATGCAGGTAGAGGAAAGGAAAATTAAACCACAAAAAATAGAAATAA GAAAAGACACATTAAAAACATTAAATGAGTTTCAAAAGTTGGTAGGAGATACTAATTGGATTCGGAGATATTAATTGGATTTGGCCA ACTCTAGGCATTCCTACTTATGCCATGTCAATTTTGTTCTTCTTTAAGAGGGGGACTTGGAAATAGTGAAAGAATGTTACCT CCAGAGGCAACTAAAGAAATTAAATTAAATTGAAGAAAAAAATTCGGTCAGCACAAGTAAATAGGATCACTTGGCCCCACTCCAAATT ATTAAGACTTTTACATTGTACTTGGATCAAATGGCTACATTAATTGGTCAGGGAAGATTACGAATAATAACATTGTGTGGAAAATGAC 50 ${\tt CCAGATAAAATCACTGTTCCTTTCAACAAGCAACAAGCTAGAACAAGCCTTTATCAGTTCTGGTGCATGGCAGATTGGTCTTGCTAAT}$ TTTCTGGGAATTATTGATAATCATTACCCAAAAACCAAAAATCTTCCAGTTCTTAAAATTGACTACTTGGATTCTACCTAAAATTACC AGACGTGAACCTTTAGAAAATGCTCTAACAGTATTTACTGATGGTTCCAGCAATGGAAAAGCGGCTTACACAGGGCCCGAAAGAACGA ATTATATCAGATTCTGCATATGTAGTACAGGCTACAAGGGATGTTGAGACAGCTCTAATTAAATATAGCACGGACGATCATTTAAAC 55 GGGCCTTTGACTAAAGCAAATGAACAAGCTGACTTACTGGTATCATCTGCATTCATAAAAGCACAAGAACTTCTTGCTTTGACTCAT GTAAATGCAGCAGGATTAAAAAAACAAATTTGATGTCACATGGAAACAGGCAAAAGATATTGTACAACATTGCACCCAGTGTCAAGTC

GAAAGAACTAATAGAACACTCAAAACTCAATTAGTTAAACAAAAGAAGGGGGAGACAGTAAGGAGTGTACCACTCCTCAGATGCAA
 CTTAATCTAGCACTCTATACTTTAAATTTTTTTAAACATTTATAGAAATCAGACTACTACTTCTGCAAAACAACATCTTACTGGTAAA
 AAGCACAGCCCACATGAAGGAAAACTAATTTGGTGGAAAGATAATAAAAATAAGACATGGGAAATAGGGAAGGTGATAACGTGGGGG
 AGAGGTTTTGCTTGTGTTTCACCAGGAGAAAATCAGCTTCCTGTTTGGATACCCACTAGACATTTGAAGTTCTACAATGAACCCCACT
 GGAGATGCAAAGAAAAGGGCCTCCACAGAGATAGTAACCCCAGTCACATGGATAATC
 SEO ID 101.

SEQ ID 101:

- - GAAGAGAAGATAATCTCACTTTAGACATTTCCAAATTAAAATAACAAATTTTCGAAGCATCAAAAGCCCATTTAAATTTGATGCCAG GAACTGAGGCAATTGCAGGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTTAAGACCATCGGAAGTACTATGATTA TAAATCTCATATTAATCCTTGTGTGCCTGTTTTGTCTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGCGACCATC GAGAACGGGCCA

SEQ ID 102:

25

30

50

55

35 SEQ ID 103:

MGQTKSKIKSKYASYLSFIKILLKRGGVKVSTKNLIKLFQIIEQFCPWFPEQGTSDLKDWKRIGKELKQAGRKGNIIPLTVWNDWAI IKAALEPFQTEEDSISVSDAPGSCLIDCNENTRKKSQKETESLHCEYVAEPVMAQSTQNVDYNQLQEVIYPETLKLEGKGPELMGPS ESKPRGTSPLPAGQVLVRLQPQKQVKENKTQPQVAYQYCRWLNFSIGHPQKVSMDIQECPQHHRAGRHTISRPLGDLILWHHLVDRV VNYMKLLINQERKEILRHGNSQ

40 SEQ ID 104:

MPPAPQGRAPYHQPPTRRLNPMAPPSRQGSELHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGAQEGEPPTVEARYKSFSIKMLKDMK EGVKQYGPNSPYMRTLLDSIAYGHRLIPYDWEILAKSSLSPSQFLQFKTWWIDGVQEQVRRNRAANPPVNIDADQLLGIGQNWSTIS QQALMQNEAIEQVRAICLRAWEKIQDPGSTCPSFNTVRQGSKEPYPDFVARLQDVAQKSIADEKAGKVIVELMAYENANPECQSAIK PLKGKVPAGSDVISEYVKACDGIGGAMHKAMLMAQAITGVVLGGQVRTFGGKCYNCGQIGHLKKNCPVLNKQNITIQATTTGREPPD LCPRCKKGKHWASQCRSKFDKNGQPLSGNEQRGQPQAPQQTGAFPIQPFVPQGFQGQQPPLSQVFQGISQLPQYNNCPSPQAAVQQ SEQ ID 105:

45 SEQ ID 105:

SEQ ID 106:

MEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQQWGAEITMPAPLYSPTSQKIMTKMGYIPGKGLGKNEDGIKVPVEAKINQEREG IGYPF

SEQ ID 107:

TTTCTGCAAGCAGAGGTTGCCAATGCTGGACTGGCAATAGCATCCGATAAGATCCAAACCTCTACTCCTTTTCATTATTTAGGGATG CAGATAGAAAATAGAAAAATTAAGCCACAAAAAATAGAAATAAGAAAAGACACATTAAAAACACTAAATGATTTTCAAAAATTACTA

	GGAGATATTAATTGGATTCGGCCAACTCTAGGCATTCCTACTTATGCCATGTCAAATTTGTTCTCTATCTTAAGAGGAGACTCAGAC
	TTAAATAGTCAAAGAATATTAACCCCAGAGGCAACAAAAGAAATTAAATTAGTGGAAGAAAAAATTCAGTCAG
	ATAGATCCCTTAGCCCCACTCCAACTTTTGATTTTTGCCACTGCACATTCTCCAACAGGCATCATTATTCAAAATACTGATCTTGTG
	GAGTGGTCATTCCTTCCTCACAGTACAGTTAAGACTTTTACATTGTACTTGGATCAAATAGCTACATTAATCGGTCAGACAAGATTA
	CGAATAACAAAATTATGTGGAAATGACCCAGACAAAATAGTTGTCCCTTTAACCAAGGAACAAGTTAGACAAGCCTTTATCAATTCT
5	GGTGCATGGCAGATTGGTCTTGCTAATTTTGTGGGACTTATTGATAATCATTACCCAAAAACAAAGATCTTCCAGTTCTTAAAATTG
	ACTACTTGGATTCTACCTAAAATTACCAGACGTGAACCTTTAGAAAATGCTCTAACAGTATTTACTGATGGTTCCAGCAATGGAAAA
	GCAGCTTACACAGGGCCGAAAGAACGAGTAATCAAAACTCCATATCAATCGGCTCAAAGAGACGAGTTGGTTG
	TTACAAGATTTTGACCAACCTATCAATATTATATCAGATTCTGCATATGTAGTACAGGCTACAAGGGATGTTGAGACAGCTCTAATT
	AAATATAGCATGGATGATCAGTTAAACCAGCTATTCAATTTATTACAACAAACTGTAAGAAAAAGAAATTTCCCATTTTATATTACT
10	TATATTCGAGCACACACTAATTTACCAGGGCCTTTGACTAAAGCAAATGAACAAGCTGACTTACTGGTATCATCTGCACTCATAAAA
10	GCACAAGAACTTCATGCTTTGACTCATGTAAATGCAGCAGGATTAAAAAACAAATTTGATGTCACATGGAAACAGGCAAAAGATATT
	GTACAACATTGCACCCAGTGTCAAGTCTTACACCTGCCCACTCAAGAGGCAGGAGTTAATCCCAGAGGTCTGTGTCCTAATGCATTA
	TGGCAAATGGATGTCACGCATGTACCTTCATTTGGAAGATTATCATATGTTCATGTAACAGTTGATACTTATTCACATTTCATATGG
	GCAACTTGCCAAACAGGAGAAAGTACTTCCCATGTTAAAAAAACATTTATTGTCTTGTTTTGCTGTAATGGGAGTTCCAGAAAAAATC
	AAAACTGACAATGGACCAGGATATTGTAGTAAAGCTTTCCAAAAATTCTTAAGTCAGTGGAAAATTTCACATACAACAGGAATTCCT
	TATAATTCCCAAGGACAGGCCATAGTTGAAAGAACTAATAGAACACTCAAAACTCAATTAGTTAAACAAAAGAAGGGGGGAGACAGT
15	AAGGAGTGTACCACTCCTCAGATGCAACTTAATCTAGCACTCTATACTTTAAAATTTTTTAAACATTTATAGAAATCAGACTACTACT
	TCTGCAGAACAACATCTTACTGGTAAAAAGAACAGCCCCACATGAAGGAAAACTAATTTGGTGGAAAGATAATAAAAAATAAGACATGG
	GAAATAGGGAAGGTGATAACGTGGGGGGAGAGGTTTTGCTTGGTTTCACCAGGAGAAAATCAGCTTCCTGTTTGGTTACCCACTAGA
	CATTTGAAGTTCTACAATGAACCCATCGGAGATGCAAAGAAAAGGGCCTCCACGGAGATGGTAACACCAGTCACATGGATGG
	CCTATAGAAGTATATGTTAATGATAGTATATGGGTACCTGGCCCCATAGATGATCGCTGCCCTGCCCAAACCTGAGGAAGAAGGGATG
	ATGATAAATATTTCCATTGGGTATCGTTATCCTCCTATTTGCCTAGGGAGAGCACCAGGATGTTTAATGCCTGCAAAAATTGG
20	TTGGTAGAAGTACCTACTGTCAGTCCCATCAGTAGATTCACTTATCACATGGTAAGCGCGGATGTCACTCAGGCCACGGGTAAATTAT
	TTACAAGACTTTTCTTATCAAAGATCATTAAAATTTAGACCTAAAGGGAAACCTTGCCCCAAGGAAATTCCCAAAGAATCAAAAAAT
	ACAGAAGTTTTAGTTTGGGAAGAATGTGTGGGCCAATAGTGCGGTGATATTATAAAACAATGAATTTGGAACTATTATAGATTGGGCA
	CCTCGAGGTCAATTCTACCACAATTGCTCAGGACAAACTCAGTCGTGTCCAAGTGCACAAGTGAGTCCAGCTGTTGATAGCGACTTA
	ACAGAAAGTTTAGACAAACATAAGCATAAAAAATTGCAGTCTTTCTACCCTTGGGAATGGGGAGAAAAAGGAATCTCTACCCCAAGA
05	CCAAAAATAGTAAGTCCTGTTTCTGGTCCTGAACATCCAGAATTATGGAGGCTTACTGTGGCCTCACACCACATTAGAATTTGGTCT
25	GGAAATCAAACTTTAGAAACAAGAGATTGTAAGCCATTTTATACTGTCGACCTAAATTCCAGTCTAACAGTTCCTTTACAAAGTTGC
	GTAAAGCCCCCTTATATGCTAGTTGTAGGAAATATAGTTATTAAACCAGACTCCCAGACTATAACCTGTGAAAATTGTAGATTGCTT
	ACTTGCATTGATTCAACTTTTAATTGGCAACACCGTATTCTGCTGGTGAGAGCAAGAGGGGGGGG
	CGACCGTGGGAGGCCTCACCATCCGTCCATATTTTGACTGAAGTATTAAAAGGTGTTTTAAATAGATCCAAAAGATTCATTTTTACT
	TTAATTGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGCTGCTGCTAGCAGGAGTTGCATTGCACTCTGTTCAGTCAG
	AACTTTGTTAATGATTGGCAAAAGAATTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAAATTGGCAAATCAAATTAAT
30	GATCTTAGACAAACTGTCATTTGGATGGGAGACAGACTCATGAGCTTAGAACATCGTTTCCAGTTACAATGTGACTGGAATACGTCA
	${\tt GATTTTTG} {\tt TATTACACCCCAAATTTATAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAGAGAAGATAAT$
	${\tt CTCACTTTAGACATTTCCAAATTAAAAGAACAAATTTTCGAAGCATCAAAAGCCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATT$
	GCAGGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTTAAGACCATTGGAAGTACATCGATTATAAATCTCATATTA
	ATCCTTGTGTGCCTGTTTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGAGAG
	ATGACGATGGCGGTTTTGTCGAAAAGAAAGGGGGGAAATGTGGGGGAAAAGCAAGGAGAGATCAAATTGTTACTGTGTCGTGTGTAG
35	
55	SEQ ID 108:
	MGPLQPGLPSPAMIPKDWPLIIIDLKDCFFTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRE
	KFSDCYIIHYIDDILCAAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDTLKTLNDFQKLL
	GDINWIRPTLGIPTYAMSNLFSILRGDSDLNSQRILTPEATKEIKLVEEKIQSAQINRIDPLAPLQLLIFATAHSPTGIIIQNTDLV
	${\tt EWSFLPHSTVKTFTLYLDQIATLIGQTRLRITKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFVGLIDNHYPKTKIFQFLKL$
	${\tt TTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRDELVAVITVLQDFDQPINIISDSAYVVQATRDVETALI$
40	KYSMDDOLNOLFNLLOOTUPKDNEDEYTTYT DAHTNLDCDLTKANEOADLUSSALTKAOFLHALTHUNAACI.KNKEDYTWKOAKDI

- 40 KYSMDDQLNQLFNLLQQTVRKRNFPFYITYIRAHTNLPGPLTKANEQADLLVSSALIKAQELHALTHVNAAGLKNKFDVTWKQAKDI VQHCTQCQVLHLPTQEAGVNPRGLCPNALWQMDVTHVPSFGRLSYVHVTVDTYSHFIWATCQTGESTSHVKKHLLSCFAVMGVPEKI KTDNGPGYCSKAFQKFLSQWKISHTTGIPYNSQGQAIVERTNRTLKTQLVKQKEGGDSKECTTPQMQLNLALYTLNFLNIYRNQTTT SAEQHLTGKKNSPHEGKLIWWKDNKNKTWEIGKVITWGRGFACVSPGENQLPVWLPTRHLKFYNEPIGDAKKRASTEMVTPVTWMDN PIEVYVNDSIWVPGPIDDRCPAKPEEEGMMINISIGYRYPPICLGRAPGCLMPAVQNWLVEVPTVSPISRFTYHMVSGMSLRPRVNY LQDFSYQRSLKFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILXNNEFGTIIDWAPRGQFYHNCSGQTQSCPSAQVSPAVDSDL
- 45 TESLDKHKHKKLQSFYPWEWGEKGISTPRPKIVSPVSGPEHPELWRLTVASHHIRIWSGNQTLETRDCKPFYTVDLNSSLTVPLQSC VKPPYMLVVGNIVIKPDSQTITCENCRLLTCIDSTFNWQHRILLVRAREGVWIPVSMDRPWEASPSVHILTEVLKGVLNRSKRFIFT LIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTRLWNSQSSIDQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTS DFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWVKTIGSTSIINLIL ILVCLFCLLLVCRCTQQLRRDSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV

SEQ ID 109:

55

⁵⁰ MNPSEMQRKAPPRRRRHRNRAPLTHKMNKMVTSEEQMKLPSTKKAGPPTWAQLKKLTQLATKYLENTKVTQTPESMLLAALMIVSMV SAGVPNSSEETATIENGP

SEQ ID 110: GAAAAAAATCAAAAAAGAA SEQ ID 111: AGCCATTAATGCCATAA SEQ ID 112: TAAATAGGATCACTT

SEQ ID 113: GGTGCGGAAATCACCATGCCCGCTCCAT SEQ ID 114: ATTATATAGCCCCACGAG 5 SEQ ID 115: CAAGATGGGATATATACCAGG SEQ ID 116: AAAACAGAAAAACCGGTG SEQ ID 117: 10 AAATCAGTGGCCGCTA SEQ ID 118: AGTTAGAAAAGGGTCAC SEQ ID 119: TGAGCCTTCGTTCTCA 15 SEQ ID 120: AGGCAAATGGCATACGT SEQ ID 121: GGCCTCTCCAACCCG SEQ ID 122: 20 GAGCAGGATTGTGAAAA SEQ ID 123: TCTTCAACCAGTGAGAGAAAA SEQ ID 124: ATTATATTGATGATATTTTA 25 SEQ ID 125: AACGAAAGATAAATT SEQ ID 126: TGACTGTTATACATT SEQ ID 127: 30 TTCATTATTTAGGGAT SEQ ID 128: AGATAGAAAATAGAAAAAT SEQ ID 129: ATTATTCAAAATACT 35 SEQ ID 130: AATAACAAAATTATGT SEQ ID 131: AGACAAAATAGTTGT SEQ ID 132: 40 TCCCTTTAACCAAGGAA SEQ ID 133: AAAAGAATGAGTCAT SEQ ID 134: CAGTATCACTTGACT 45 SEQ ID 135: TTTTAATCAGTCTATTAACATTG SEQ ID 136: AAAGGATATTGAGAGA SEQ ID 137: 50 CCTAATCAAATACATT SEQ ID 138: CGCTGTTTAATTTGT SEQ ID 139: TGCATTCATGGAAGCA 55 SEQ ID 140: ACTCAGGAGGCAAGA SEQ ID 141: TTAAGAGACATTTATT

	TAAAGCAGTTCAAAAA
	SEQ ID 143:
5	AATAGGAATTCTCTA
	SEQ ID 144:
	AAAGCTCAATTGGTTA
	SEQ ID 145:
	ACGGACGATCATTTAA

SEQ ID 146:

5

MGQTKSKIKSKYASYLSFIKILLKRGGVKVSTKNLIKLFQIIEQFCPWFPEQGTLDLKDWKRIGKELKQAGRKGNIIPLTVWNDWAI
${\tt IKAALEPFQTEEDSVSVSDAPGSCIIDCNENTGKKSQKETEGLHCEYVAEPVMAQSTQNVDYNQLQEVIYPETLKLEGKGPELVGPS$
ESKPRGTSPLPAGQVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPESQYGYPGMPPAPQGRAPYPQPPTRRLNPTAPPSRQG
SKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGAQEGEPPTVEARYKSFSIKKLKDMKEGVKQYGPNSPYMRTLLDSIAHGHRLIPY
DWEIQAKSSLSPSQFLQFKTWWIDGVQEQVRRNRAANPPVNIDADQLLGIGQNWSTISQQALMQNEAIEQVRAICLRAWEKIQDPGS
TCPSFNTVRQGSKEPYPDFVARLQDVAQKSIADEKARKVIVELMAYENANPECQSAIKPLKGKVPAGSDVISEYVKACDGIGGAMHK
AMLMAQAITGVVLGGQVRTFGRKCYNCGQIGHLKKNCPVLNKQNITIQATTTGREPPDLCPRCKKGKHWASQCRSKFDKNGQPLSGN
EQRGQPQAPQQTGAFPIQPFVPQGFQGQQPPLSQVFQGISQLPQYNNCPPPQAAVQQ
SEQ ID 147:

10 WATIVGKRAKGPASGPTTNWGIPNSAICSSGFSGTTTPTVPSVSGNKPVTTIQQLSPATSGSAAVDLCTIQAVSLLPGEPPQKTPTG VYGPLPKGTVGLILGRSSLNLKGVQIHTSVVDSDYKGEIQLVISSSIPWSASPRDRIAQLLLLPYIKGGNSEIKRIGGLGSTDPTGK AAYWASQVSENRPVCKAIIQGKQFEGLVDTGADVSIIALNQWPKNWPKQKAVTGLVGIGTASEVYQSTEILHCLGPDNQESTVQPMI TSIPLNLWGRDLLQQWGAEITMPAPSYSPTSQKIMTKMGYIPGKGLGKNEDGIKIPVEAKINQEREGIGNPC

SEQ ID 148:

- NKSRKRRNRESLLGAATVEPPKPIPLTWKTEKPVWVNQWPLPKQKLEALHLLANEQLEKGHIEPSFSPWNSPVFVIQKKSGKWRMLT
 DLRAVNAVIQPMGPLQPGLPSPAMIPKDWPLIIIDLKDCFFTIPLAEQDCEKFAFTIPAINNKEPATRFQWKVLPQGMLNSPTICQT
 FVGRALQPVREKFSDCYIIHCIDDILCAAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDT
 LKTLNDFQKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLTPEATKEIKLVEEKIQSAQINRIDPLAPLQLLIFATAHSP
 TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQIATLIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFVGIIDNHY
 PKTKIFQFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTFYQSAQRAELVAVITVLQDFDQPINIISDSAYVV
 QATRDVETALIKYSMDDQLNQLFNLLQQTVRKRNFPFYITHIRAHTNLPGPLTKANEQADLLVSSALIKAQELHALTHVNAAGLKNK
 FDVTWKQAKDIVQHCTQCQVLHLPTQEAGVNFRGLCPNALWQMDVTHVPSFGRLSYVHVTVDTYSHFIWATCQTGESTSHVKKHLLS
 CFAVMGVPEKIKTDNGPGYCSKAFQKFLSQWKISHTTGIPYNSQGQAIVERTNRTLKTQLVKQKEGGDSKECTTPQMQLNLALYTLN
 FLNIYRNQTTSAEQHLTGKKNSPHEGKLIWWKDNKNKTWEIGKVITWGRGFACVSPGENQLPVWIPTRHLKFYNEPIRDAKKSTSA
 ETETSQSSTVDSQDEQNGDVRRTDEVAIHQEGRAANLGTTKEADAVSYKISREHKGDTNPREYAACSLDDCINGGKSPYACRSSCS
 - SEQ ID 149:
- ²⁵ MNPSEMQRKAPPRRRRHRNRAPLTHKMNKMVTSEEQMKLPSTKKAEPPTWAQLKKLTQLATKYLENTKVTQTPESMLLAALMIVSMV VSLPMPAGAAAANYTYWAYVPFPPLIRAVTWMDNPTEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGYHYPPICLGRAPGCLMP AVQNWLVEVPTVSPICRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEFGT IIDWAPRGQFYHNCSGQTQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIVSPVSGPEHPELWRLTVASHH IRIWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKPDSQTITCENCRLLTCIDSTFNWQHRILLVRAREGVWI PVSMDRPWEASPSVHILTEVLKGVLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTRLWNSQSSIDQKL ANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCITPQIYNESEHHWDMVRRHLQGREDNLTLDISKLKEQIFEASKAHLNLVP GTEAIAGVADGLANLNPVTWVKTIGSTTIINLILILVCLFCLLLVCRCTQQLRRDSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVT VSV

SEQ ID 150:

SEQ ID 151:

```
SEQ ID 152:
```

SEQ ID 153:

20

25

30

5

10

15

CCACCCCTACA

SEQ ID 154:

SEQ ID 155:

 ${\tt GAGATAGGGAAAAAACCGCCTTAGGGCTGGAGGTGGGACCTGCGGGCAGCAATACTGCTTTGTAAAGCACTGAGATGTTTATGTGTATGCATATCTAAAAGCACAGCACTTAATCCTTTACATTGTCTATGATGCAAAGACCTTTGTTCAC$

35 SEO ID 156:

40 TCACAAGATGAAXAAAATGGTGAXXTCAGAAGAACAGATGAAGTTGCCATCCACCAAGAAXGCXGAGCCGCCGACTTGGGCACAAXT AAAGAAGCTGACACAGTTAGCTAXAAAAXXXXXXCTXGAGAACACAAAGGTGACACAAACTCCAGAGAXTATGCTGCTTGCAGCTTT GATGATTGTATCAATGGTGGTAAGTCTCCCXATGCCTGCAGGAGCAGCTGCAGCTAAXTATACXTACTGGGCCTATGTGCCTTTCCC GCCCTTAATTCGGGCAGTCACATGGATGGATAATCCTATTGAAGTATAGTATAATAATAGTGTATGGGXTACCTGGCCCCACAGATG ATCGTTGCCCTGCCAAACCTGAGGAAGAAGGAATGATGATAAATATTTCCATTGGGTATCXTTATCCTCCTATTTGCCTAGGGAGAG 45 CACCAGGATGTTTAATXGCCTGCAXTCCAAAATTGGTTGGTAGAAGTACCTACTGTCAGTXCCAXCAGTAGATTCACTTATCACATG GTAAGXGGXATGTCACTCAGGCCACXGGTAAATXATTTACAXGACTTTTCTTATCAAAGATCATTAAAATTTAGXCCTAAAGGGAAA CCTTGCCCCAAGGAAATTCCCCAAAGXATCAAAAXAXXCAGAAGTTTTAGTTTGGGAAGAATGTGTGGCXAATAGTGCXGTGATATTA 50 CTGGTCCTGAACATCCAGAATTATGGAXGCTTACTGTGGCCTCAXXACCACATTAGAATTTGGTCTGGAAATCAAXCTXTAGAAACA AGAGATCXTAAGCCATXTTATACTATCXACCTAAATTCCAGTCTXACAXTTCCTTTXCAAAGTTGXGTAAAGCCCCCCTTATATXGCT AGTTGTAGGAAATAXXTAGTTATTAAACCAGAXTCCCAAACTATAXXACCTGTGAAAATTGTAGATTGTTTACTTGCATTGATTCAA 55 CTTTTTAATTGGCAGCACCGTATTCTGCTXGTGAGAGCAAGAGAXGGXGTGTGGATCCCTGTGTCCATGGACCGACCGTGGGAGGCXT CXCCATCCXTCCATATTTTXACXGAAGTATTAAAAGGXXTTXTAAXTAGATCCAAAAGATTCATTTTACTTTAATTGCAGTGATTA

10 SEQ ID 158:

XXXXXXXXXX

5

20

25

SEQ ID 159:

SEQ ID 160:

40

45

SEQ ID 161: TAGGCCTTTGAGGGA SEQ ID 162: TAGGCCTTATTTTAGGG SEQ ID 163:

- GAGAAGGAGCCCAAGAG SEQ ID 164: GAGCCTCCCACAGTT
 - SEQ ID 165:
- AGGCCAGATACAAGTCT SEQ ID 166: TTTTCGATAAAAATGCTA SEQ ID 167: TTATATGAGGACATTA
 SEQ ID 168:
 - SEQ ID 168. TTATGGACATAGACTCAT SEQ ID 169: TTGGGAGATTCTGGCAAA

SEQ ID 170: AATCGTCTCTCTCACC SEQ ID 171: AATTTTTACAATTTAAGACT 5 SEQ ID 172: GTCCGAAGAAATAGG SEQ ID 173: TGCCAATCCTCCAGTT SEQ ID 174: 10 AACATAGATGCAGATCAACTAT SEQ ID 175: AGTACTATTAGTCAACAA SEQ ID 176: GTCAACAAGCATTAATGCAA 15 SEQ ID 177: CCATTGAGCAAGTTAGAG SEQ ID 178: GAGCTATCTGCCTTAGAG SEQ ID 179: 20 CTTGGGAAAAAATCCAAGAC SEQ ID 180: GAAGTACCTGCCCCTCATTTAA-TACAGTAA SEQ ID 181: CCCTACCCTGATTTT 25 SEQ ID 182: AAGGCTCCAAGATGTT SEQ ID 183: TCAATTGCCGATGAAAAAG SEQ ID 184: 30 CGGTAAGGTCATAGTGG SEQ ID 185: TGGAGTTGATGGCATAT SEQ ID 186: AAACGCCAATCCTGAGT 35 SEQ ID 187: TCAATCAGCCATTAA SEQ ID 188: AAAGGTTCCTGCAGGATCAGA SEQ ID 189: 40 AGGATCAGATGTAATCTCA SEQ ID 190: AATATGTAAAAGCCTGT SEQ ID 191: ATAAAGCTATGCTTAT 45 SEQ ID 192: AATAACAGGAGTTGTTTTAG SEQ ID 193: ACATTTGGAGGAAAAT SEQ ID 194: ATTGGTCACTTAAAAAA 50 SEQ ID 195: ATTGGTCACTTAAAAAA SEQ ID 196: GGTAGAGAGCCACCTGACTTAT 55 SEQ ID 197: AAGATGTAAAAAAGG SEQ ID 198: GCTAGTCAATGTCGTT

SEQ ID 199: GGGAAACGAGCAAAG SEQ ID 200: CCAATTCAGCCATTTG 5 SEQ ID 201: CCACTGTCCCAAGTGTTTC SEQ ID 202: AATAAGCCAGTTACCA SEQ ID 203: 10 ACAATACAACAATTG SEQ ID 204: CTCACCACAAGCGGCAGTGCAGC SEQ ID 205: TACTATACAAGCAGTCTCTCTGCTTCCAGGGGAGC 15 SEQ ID 206: AAAAAATCCCTACAGG SEQ ID 207: CACTGCCTGAGGGGACTG SEQ ID 208: 20 GACTAATCTTGGGAAGA SEQ ID 209: AAATCTAAAAGGAGTTCA SEQ ID 210: CTAGTGTGGTTGATTCAGACT 25 SEQ ID 211: CGAAATTCAATTGGTTATTA SEQ ID 212: TCTTCAATTCCTTGG SEQ ID 213: 30 AGTCCAAGAGACAGGAT SEQ ID 214: TTATTACTCCTGCCATATA SEQ ID 215: CATTAGAAAAAGGACATTG 35 SEQ ID 216: TTGGAATTCTGTTTGTA SEQ ID 217: TAACTGAGCCATTAAT SEQ ID 218: 40 AGCCATGGTCCCCTTTAATTA SEQ ID 219: TTTTACCACACCAGCCT SEQ ID 220: TTGTCAGCTCAAGCT 45 SEQ ID 221: TACATCGTTCACTAT SEQ ID 222: TTAAAAGCATTAAAT SEQ ID 223: 50 AGAAGTCCCAATTGAGG SEQ ID 224: GGTCTTGCCGATTTT SEQ ID 225: ACAATCGTTACCACA

Claims

5

25

45

55

- 1. A method for diagnosing prostate cancer, the method comprising the step of detecting the presence or absence of an expression product of a HML-2 endogenous retrovirus in a patient sample, wherein the patient sample contains prostate cells and/or wherein the patient is suspected of having prostate cancer wherein up-regulation of expression of at least 150% relative to a negative control is indicative of prostate cancer.
- 2. The method of claim 1, wherein the expression product is a RNA or a polypeptide.
- 10 **3.** The method of any preceding claim, wherein the patient sample is a prostate sample or a blood sample.
 - 4. The method of any preceding claim, wherein the expression product is a RNA having the following formula: N₁-N₂-N₃-N₄-N₅-polyA, wherein: N₁ has at least 75% sequence identity to SEQ ID 155; N₂ has at least 75% sequence identity to SEQ ID 156; N₃ has at least 75% sequence identity to SEQ ID 6; N₄ comprises any RNA sequence; N₅
- ¹⁵ has at least 75% sequence identity to SEQ ID 5; and at least one of N_1 or N_5 is present, but N_2 , N_3 , N_4 and polyA are optional.
 - 5. The method of claim 4, wherein the RNA comprises N₁.
- 20 **6.** The method of claim 5, wherein N_1 is at the 5' end of the RNA.
 - 7. The method of claim 4, wherein N_4 comprises a polypeptide-coding sequence.
 - 8. The method of claim 2, wherein the polypeptide is encoded by a mRNA having the following formula: N₁-N₂-N₃-N₄-N₅-polyA, as defined in claim 4.
 - 9. The method of claim 8, wherein the mRNA encodes one or more of the following HML-2 polypeptides: gag, prt, pol, env, cORF, tat.
- **10.** The method of claim 8 or claim 9, wherein the polypeptide is detected using an antibody.
 - 11. The method of any one of claims 1 to 6, wherein said step is preceded by a step of enriching RNA in the patient sample.
- 12. The method of any one of claims 1 to 6 or 11, wherein the expression product is detected using PCR, SDA, SSSR,
 ³⁵ LCR, TMA or NASBA.
 - **13.** The method of claim 12, wherein the PCR is RT-PCR.

40 Patentansprüche

- Verfahren zum Diagnostizieren von Prostatakrebs, wobei das Verfahren den Schritt des Nachweisens des Vorhandenseins oder Fehlens eines Expressionsprodukts eines endogenen HML-2-Retrovirus in einer Patientenprobe umfasst, wobei die Patientenprobe Prostatazellen enthält und/oder wobei der Verdacht besteht, dass der Patient Prostatakrebs aufweist, wobei eine Hochregulation der Expression von wenigstens 150 % bezogen auf eine Negativkontrolle indikativ für Prostatakrebs ist.
- 2. Verfahren gemäß Anspruch 1, wobei das Expressionsprodukt eine RNA oder ein Polypeptid ist.
- 50 **3.** Verfahren gemäß einem der vorstehenden Ansprüche, wobei die Patientenprobe eine Prostataprobe oder eine Blutprobe ist.
 - 4. Verfahren gemäß einem der vorstehenden Ansprüche, wobei das Expressionsprodukt eine RNA mit der folgenden Formel ist: N₁-N₂-N₃-N₄-N₅-polyA, wobei: N₁ wenigstens 75 % Sequenzidentität mit SEQ ID 155 aufweist; N₂ wenigstens 75 % Sequenzidentität mit SEQ ID 156 aufweist; N₃ wenigstens 75 % Sequenzidentität mit SEQ ID 6 aufweist; N₄ eine beliebige RNA-Sequenz umfasst; N₅ wenigstens 75 % Sequenzidentität mit SEQ ID 5 aufweist; und N₁ und/oder N₅ vorhanden sind, während N₂, N₃, N₄ und polyA optional sind.

- 5. Verfahren gemäß Anspruch 4, wobei die RNA N₁ umfasst.
- 6. Verfahren gemäß Anspruch 5, wobei N1 an dem 5'-Ende der RNA angeordnet ist.
- 5 7. Verfahren gemäß Anspruch 4, wobei N_4 eine Polypeptid-codierende Sequenz umfasst.
 - Verfahren gemä
 ß Anspruch 2, wobei das Polypeptid von einer mRNA, die folgende Formel gem
 ä
 ß Anspruch 4
 aufweist : N₁-N₂-N₃-N₄-N₅-polyA, codiert wird.
- Verfahren gemäß Anspruch 8, wobei die mRNA ein oder mehrere der folgenden HML-2-Polypeptide codiert: gag, prt, pol, env, cORF, tat.
 - **10.** Verfahren gemäß Anspruch 8 oder Anspruch 9, wobei das Polypeptid unter Verwendung eines Antikörpers nachgewiesen wird.
- 15

30

40

45

- **11.** Verfahren gemäß einem der Ansprüche 1 bis 6, wobei dem Schritt ein Schritt des Anreicherns von RNA in der Patientenprobe vorausgeht.
- 12. Verfahren gemäß einem der Ansprüche 1 bis 6 oder 11, wobei das Expressionsprodukt unter Verwendung von
 ²⁰ PCR, SDA, SSSR, LCR, TMA oder NASBA nachgewiesen wird.
 - 13. Verfahren gemäß Anspruch 12, wobei die PCR RT-PCR ist.

25 Revendications

- 1. Méthode de diagnostic du cancer de la prostate, la méthode comprenant l'étape consistant à détecter la présence ou l'absence d'un produit d'expression d'un rétrovirus endogène HML-2 dans un prélèvement d'un patient, le prélèvement d'un patient contenant des cellules de la prostate et/ou le patient étant suspecté d'avoir le cancer de la prostate, la régulation positive de l'expression d'au moins 150 % par rapport à un témoin négatif étant indicatrice du cancer de la prostate.
- 2. Méthode selon la revendication 1, caractérisée en ce que le produit d'expression est un ARN ou un polypeptide.
- **35 3.** Méthode selon l'une ou l'autre des revendications 1 et 2, **caractérisée en ce que** le prélèvement d'un patient est un prélèvement de la prostate ou un prélèvement de sang.
 - 4. Méthode selon l'une quelconque des revendications précédentes, caractérisée en ce que le produit d'expression est un ARN ayant la formule suivante : N₁-N₂-N₃-N₄-N₅-polyA, dans laquelle : N₁ a au moins 75 % d'identité de séquence avec SEQ ID 155 ; N₂ a au moins 75 % d'identité de séquence avec SEQ ID 155 ; N₂ a au moins 75 % d'identité de séquence avec SEQ ID 156 ; N₃ a au moins 75 % d'identité de séquence avec SEQ ID 6 ; N₄ comprend une séquence ARN quelconque ; N₅ a au moins 75 % d'identité de séquence avec SEQ ID 5 ; et au moins l'un de N₁ ou N₅ est présent, mais N₂, N₃, N₄ et polyA sont facultatifs.
 - 5. Méthode selon la revendication 4, caractérisée en ce que l'ARN comprend N₁.
 - 6. Méthode selon la revendication 5, caractérisée en ce que N₁ est à l'extrémité 5' de l'ARN.
 - 7. Méthode selon la revendication 4, caractérisée en ce que N₄ comprend une séquence codant pour un polypeptide.
- Méthode selon la revendication 2, caractérisée en ce que le polypeptide est codé par un ARNm ayant la formule suivante : N₁-N₂-N₃-N₄-N₅-polyA, telle que définie dans la revendication 4.
 - 9. Méthode selon la revendication 8, caractérisée en ce que l'ARNm code pour un ou plusieurs des polypeptides de HML-2 suivants : gag, prt, pol, env, cORF, tat.

55

10. Méthode selon la revendication 8 ou la revendication 9, **caractérisée en ce que** le polypeptide est détecté à l'aide d'un anticorps.

- **11.** Méthode selon l'une quelconque des revendications 1 à 6, **caractérisée en ce que** ladite étape est précédée d'une étape d'enrichissement de l'ARN dans le prélèvement d'un patient.
- **12.** Méthode selon l'une quelconque des revendications 1 à 6 ou 11, **caractérisée en ce que** le produit d'expression est détecté par PCR, SDA, SSSR, LCR, TMA ou NASBA.
 - **13.** Méthode selon la revendication 12, **caractérisée en ce que** la PCR est la RT-PCR.



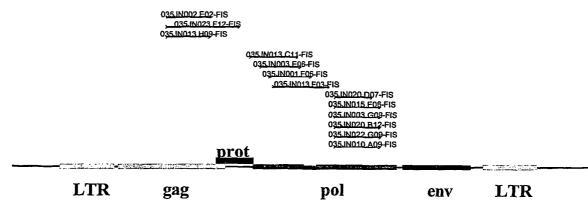
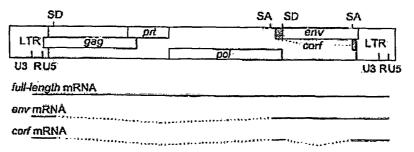


FIGURE 2 5'-LTR -LTR 3' 622 E LA cum pol 盟業 2513 pri gag env P.T RNase H 2 EN Ŗ 1033 30.0 112 鹊 DC28 5 2 3 5 5 6 4 8 ģ 10 kb i





.

.

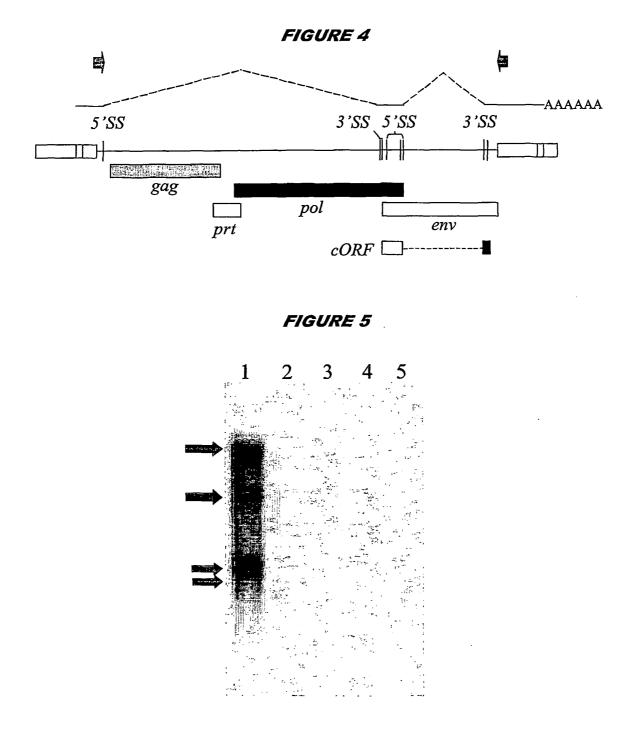


FIGURE 6

GENOMIC HERV MDA	(1)	1 BO ACATTTADAGTTCTACA
ENOMIC HERV-K TAN.	(1)	
GENOMIC AC025420	(1)	
GENOMIC AP000776		ACATTCAAGTTCTACA
NV GENOMIC HERV-K8		
V GENOMIC HERV-KI		ACATITGAAGTTCIACA
NV HERV-K AF023261	(1)	GGGGAGAGGTTTTGCTTGTGTTTCACCAGGAGAAAATCAGCTTCCTGTTTGGATACCCACTAGACATTTGAAGTTCTACA
ENV GEN AL035086	(1)	
J GENOMIC AL035587	(1)	
/ GENOMIC AC012068 / GENOMIC AF277315	(1) (1)	
GENOMIC AP277515		
GENOMIC AC078899	(1)	- ACATTTGA - ITCIACA - ACATTGA - ITCIACA - ACATTGAAGTCTACA - ATACCCACTAGACATTGAAGTCTACA - TAGACATTGAAGTTCTACA - ACATTTGAAGTTCTACA
GENOMIC HERV-KII	(1)	ACATTICAACTICIACA
GENOMIC AC008813	(1)	ATACCCACTAGACATTCAAAGTTCTACA
/ GENOMIC AC012309	(1)	TAGACATTGAAGITCIACA
GENOMIC AL121932	(1)	
GENOMIC AD000090	(1)	ACAPPICA ACAPPICATION AND A ACAPPICATION AND ACAPPICATION AND A ACAPPICATION AND ACAPPICATION AND ACAPPICATION AND ACAPPICATION AND A ACAPPICATION AND A ACAPPICATION AND A ACAPPICATION AND A ACAPPICATION AND ACAPPICA
ENV GEN AL160008	(1)	
V GENOMIC HEU32496		GOGTAATCATTGAGGACAAGTCGACGAGAGATCCCGAGGACGTCTACAGTCAGCCTTACAGTTTGAAGTTCTACA
V GENOMIC AC011467	(1)	GGTTTTGCTTGTGTTTCAOGAGGAGA-AAATCAGCTTCCTGTTTGGATGCCCACTAGACATTTGAAGGTCTACA TTTTCCTGTGTTTCACCAGGAGA-AAATCAGCTTCCTGTTTGGATACCCACTAGACATTTGAAGGTTCTACA
V GENOMIC AF235103 V GENOMIC AC026786		TTTTCTTGTGTTTTAACCAGAAAATAAATCAGCTTCCAGTTTGGATACCTACTAGAACTAGAAATTAAATCAGCTTCCAGTTTGGATACTTACAACTT
GENOMIC AC028788	- iii	CACCAGGAGA-AAATCAGCTTCCTGTTTGGGTACCCACTAGACATTTGAAGTTCTACA
V GENOMIC AC018809	(1)	-ATTTGAAGTTCTACA
HERV-K102 AF164610	(1)	TTGCTTGTGTTTCACCAGGAGA-AAATCAGCTTCCTGTTTGGATACCCACTAGACATTTGAAGTTCTACA
MIC FRAG. AF260253	(1)	
CONSENSUS	(1)	ACATTTGAAGTTCTACA
		160
CENONTO HEBU MON	(70)	81 160 ATGAACICACTGGIAGATGCAAAGAAA
V GENOMIC HERV MDA ENOMIC HERV-K TAN.	(18) (18)	
V GENOMIC AC025420	(18)	ATTICATIVATION AND A
V GENOMIC AP000776	(18)	ATGAACCCATCA EAGATGCAAAGAAXAGCACCTCGCGGAGACGGAGACACCGCAATCGAGCACC
NV GENOMIC HERV-KB	(1)	CISCARTOSAGCACC
NV GENOMIC HERV-KI	(18)	ATGAACCATCGGAGALTGCAAGAAA ATGAACCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAA ATGAACCCATCGGAGALTGCAAGAAAA ATGAACCCATCGGAGALTGCAAGAAAACCAACCACATCGAGACAACCAACCAACCGACGACACAACCAAC
NV HERV-K AF023261	(81)	atgaaccoatcogragaticcaaagaaa'agcabcticcocgagagacacaccaccesaatcoagcacce
ENV GEN AL035086	(1)	
V GENOMIC AL035587	(18)	ATGAACCCCTTGGAGATGGGAAGAAABGCGCTOTGCGGAGACAAAAAAACCTTCAATTGASCATC
V GENOMIC AC012068	(6)	ATGAACCCATCAGAGATCCAAATGAAGTGCCTCCATAGAGACAGAAAACCTGCAATCGAGCATC
V GENOMIC AF277315	(15)	ATGAACCCATCGGRAATGCAAAGAAAAGCGCCTCCGCGGGAGACAGAAAACCCCCCAATCGAGCATC
V GENOMIC AF027650	(79)	ATGAACCCATCGGAGATGCAAAGGAAAGCACCCCCACGGAGACGGAAACACCGCAATTGAACACCGC
V GENOMIC AC078899 V GENOMIC HERV-KII	(16) (18)	
V GENOMIC AC008813	(29)	
V GENOMIC AC012309	(21)	ATCIAACTIGTIGTGGGATGGAAAGAAAAGCACCTCCACGGGGGAGAGAAACACTGCAATCAAGCACC
V GENOMIC AL121932	(18)	ATGAACCTOTOGGAGATGCAAAGAAAAGCACCTCCACGGAGACAGAAACACTGCAATCAAGCACC ATGAACCCATCGGAGATGCAAAGAAAAGCGCCTCCACGGAGATGGAAACACAGCAATGGAGCACG
V GENOMIC AD000090	(18)	ATGAACCCATCAGAGATGCAAAGAAAAGCACCTCCACGGAGACGGAGACGGCAATCGAGCACCGTCAT
ENV GEN AL160008	(1)	
V GENOMIC HEU32496	(78)	ATCCACCCATCGCAGATGCAAAGAAAAGCACCTCCGCGGGAGACGGACACCCCCCAATCGAGCACCC
V GENOMIC AC011467	(74)	ATGAACCCATCGGAGATGAAAAAAAAAAAAGCGCCTCCATGGAGATGGAAACACCCAC
V GENOMIC AF235103	(72)	ATAAACCCATCCQTCAGAAATBCAAA-TAAAAGTGCCTCCACAGAGAGAAAAAACCTGCAGTGAGCATC
V GENOMIC AC026786	(58)	
V GENOMIC AC034203	(58)	
V GENOMIC AC018809	(16) (70)	
HERV-K102 AF164610 MIC FRAG. AF260253	(10)	4779477747:102490172-040642450977775-17254755077755190747591
CONSENSUS		ATGAACCCATC GAGATGCAAAGAAA AGC CCTCC CGGAGACGGAAACACCGCAATCGAGCA C
		161 240 240
V GENOMIC HERV MDA		CTCGCCAGGTAAACAAAATGGTGHTATCAGAAGAACAGAAAAAGTTGCCTTCCATCAAGGAAGCAG
ENOMIC HERV-K TAN.		OTTGACTCACAAGATGAACAAAATGGTGACGTCAGAAGAACAGAAGAAGTTGCCATCCACCAAGAAGGCAG
V GENOMIC AC025420	(83)	CITUACTCACAAGATGAA CAAAATGGTGACATCAGAA CAGATGAAGTGCCATCCACCAAGAAGGCAG
V GENOMIC AP000776 NV GENOMIC HERV-KB		GITURCTCACRAGATGAACRAAAATGGTGACGTCAGAAGAACAGATGAAGTTGCCATCCACCAAGAAGGCAG GITURCTCACAAGATGAACRAAATGGTGACGTCAGAAGAACAGATGAAGTTGCCATCCACCAAGAAA
		GIIIGACTCACAAGATGAACAAAATGGTGACGTCAGAAGAACAGAAGAGAGGTGCCATCCACCAAGAAGCGCAG
NV GENOMIC HERV-KI NV HERV-K AF023261		GITIACTCACAAGATGAACAAAATGITGACGICAGAAGAACAGATGAAGTTGCCATCCACCAAGAAGGCAG
ENV GEN AL035086		
V GENOMIC AL035587	(83)	ATCIACTCCCCAGETAAATAGTAGTATATGGTGATATCAGAAGAACAGATGAACTTGCCATCCACCAAGGAAGG
V GENOMIC AC012068		ATCACTTCCCCACGTGAAGAAAATGGTGATATCAAAAGAACAGATGAAGTTGCCATCCACCAAGAAAGCGGG
V GENOMIC AF277315		
V GENOMIC AF027650		ATTGACTCATAAGATGAA GAAAAAGGTCATGTCAGAAGAACAGATGAAGTTGCCTTCCACCAAGAAGACAG
V GENOMIC AC078899	(81)	AITGACTCETAAGATGAAGAAGAAGAAGATGACGTCATGACGACGATGACGTTGCCCACCAAGAAGAAGAAGACGC
V GENOMIC HERV-KII		
V GENOMIC AC008813	(104)	
V GENOMIC AC012309	(86)	ATCAACTCACAAGATAAAATGGTGATGTCAGAAGAATAGATGAAGATGCCATTCACCAAGAAGGCAG
V GENOMIC AL121932		ATCGACTCACAAGATGAACAGAATGGTGATGTCAGAAGAATAGATGAAGTTGCCATCCACCAAGAAGACAA
V GENOMIC AD000090	(88)	CGAGAACCATCGACTCAGAAGATGAACAGAATGGTGATGTCAGAAGAACAGATGAAGTTGCCATCCACCAAGAAGGCAG
ENV GEN AL160008	(1)	
V GENOMIC HEU32496	(143)	GITUALTUAUAGEATEAR GAAAALUGTUACATCAGAACAGATGAAGTTUCCATCCALCAAGATGACAG
V GENOMIC AC011467	(128)	

ENV GENOMIC HERV MDA
ENV GENOMIC HERV-K TAN.
ENV GENOMIC AC025420
ENV GENOMIC AP000776
ENV GENOMIC HERV-K8
ENV GENOMIC HERV-KI
ENV HERV-K AF023261
ENV GEN AL035086
ENV GENOMIC AL035587
ENV GENOMIC AC012068
ENV GENOMIC AF277315
ENV GENOMIC AF027650
ENV GENOMIC AC078899
ENV GENOMIC HERV-KII
ENV GENOMIC AC008813
ENV GENOMIC AC012309
ENV GENOMIC AL121932
ENV GENOMIC AD000090
ENV GEN AL160008
ENV GENOMIC HEU32496
ENV GENOMIC AC011467
ENV GENOMIC AF235103
ENV GENOMIC AC026786
ENV GENOMIC AC034203
ENV GENOMIC AC018809
ENV GENOMIC HERV-K102 AF164610
ENV GENOMIC FRAG. AF260253
CONSENSUS

ENV GENOMIC HERV MDA
ENV GENOMIC HERV-K TAN.
ENV GENOMIC AC025420
ENV GENOMIC AP000776
ENV GENOMIC HERV-K8
ENV GENOMIC HERV-KI
ENV HERV-K AF023261
ENV GEN AL035086
ENV GENOMIC AL035587
ENV GENOMIC AC012068
ENV GENOMIC AF277315
ENV GENOMIC AF027650
ENV GENOMIC AC078899
ENV GENOMIC HERV-KII
ENV GENOMIC AC008813
ENV GENOMIC AC012309
ENV GENOMIC AL121932
ENV GENOMIC AD000090
ENV GEN AL160008
ENV GENOMIC HEU32496
ENV GENOMIC AC011467
ENV GENOMIC AF235103
ENV GENOMIC AC026786
ENV GENOMIC AC034203
ENV GENOMIC AC018809
ENV GENOMIC HERV-K102 AF164610
ENV GENOMIC FRAG. AF260253
CONSENSUS

	3	61 240
ENV GENOMIC HERV MDA	(72)	
ENV GENOMIC HERV-K TAN.	(83)	
ENV GENOMIC AC025420	(83)	GITGACTCACAAGATGAACAAAATGGTGACATCAGAACAGATGAAGTTGCCATCCACCAAGAAGGCAGA
ENV GENOMIC AP000776	(83)	GITIJACTCACAAGATGAACAAAATGGTGACGTCAGAAGAACAGATGAAGTTGCCATCCACCAAGAAGGCAGA
ENV GENOMIC HERV-KB	(15)	GITISACTCACAAGATGAACAAGATGGTGACGTCAGAAGAACAGATGAAGTTGCCATCCACCAAGAAAGCAGA
ENV GENOMIC HERV-KI	(83)	
ENV HERV-K AF023261	(146)	
ENV GEN AL035086	(1)	
ENV GENOMIC AL035587	(83)	AUCIACTCCCCAGETAAAAAATGETGATATCAGAAGAACAGATGAAGTTGCCATCCACCAAGGAAGG
ENV GENOMIC AC012068	(71)	AICACTICCCCACGIGAAGAAATUGTGATATCAAAAGAACAGATGAAGITGCCATCACCAAGAAAGGGGA
ENV GENOMIC AF277315	(80)	
ENV GENOMIC AF027650	(144)	ATTENCTCATAAGATGAAGAAAAAAGGICATGTCAGAAGAAGAAGATGAAGTTGCCTTCCACCAAGAAGAAGAAGAA
ENV GENOMIC AC078899	(81)	Altgactchtaagatgaagaagaggggggggggggggggggg
ENV GENOMIC HERV-KII	(72)	
ENV GENOMIC AC008813	(104)	ATTEACTCACAAGATABACAAAATGGTGATGTCAAAAAAAAAAA
ENV GENOMIC AC012309	(86)	ATCAACTCACAAGA1AAAATGGTGATGTCAGAAGAATAGATGAAGTTGCCATTCACCAAGAAGGCAGA
ENV GENOMIC AL121932	(83)	ATCGACTCACAAGATGAACAAGAATGGTGHTGTCAGAAGAATAGATGAAGTTGCCATCCACCAAGAAGAAGA
ENV GENOMIC AD000090	(88)	CGAGAACCATCAACACCACCAGAAGATGAACCACGAGAACAGAACAGATGAAGTTGCCATCCACCAAGAAGGCACA
ENV GEN AL160008	(1)	-ACAGATAAAGTTGTCATCCACCAAGAAAGCGCA
ENV GENOMIC HEU32496	(143)	GITGACTCACAGGATGAA'QAAAATGGTGACATCAGA'ACAGATGAAGTTGCCATCCACCAAGATGGCAGA
ENV GENOMIC AC011467	(128)	
ENV GENOMIC AF235103	(142)	ATOGACOCACOAGGTGAACAAAATGGTGBTATCAGAAGAACAGATGAAGTTGCCATCCACCAAGAAAGCGGR
ENV GENOMIC AC026786	(58)	-ACAAAGAAGTTTCCATCTACCAAGAAAGCGGA
ENV GENOMIC AC034203	(123)	
ENV GENOMIC AC018809	(70)	
ENV GENOMIC HERV-K102 AP164610	(124)	
ENV GENOMIC FRAG. AF260253	(1)	
CONSENSUS	(161)	T GACTCACAAGATGAA AAAATGGTGA TCAGAAGAACAGATGAAGTTGCCATCCACCAAGAA GC GA

FIGURE 6 CONTD....

Αλλας ΤΑ ΓΑσΑΛΤ Α CAAAGG ΤΑ CAA TA ΤΤΑ ΓΑ ΓΑΛΑ Α CAAAGG ΤΑ CAA AG Ο ΤΑ Α CAAAGA CAAAGG ΤΑ CAA AG Ο ΤΑ Α CAAACA CAAAGG ΤΑ CAA AG Ο ΤΑ Α CAAACA CAAAGG ΤΑ CAA AG Ο ΤΑ ΓΑ CAAACA CAAAGG TA CAA AG Ο ΤΑ CAAACA CAAAGG TA CAA AG Ο ΤΑ CAAACA CAAAGG TA CAA AG Ο ΤΑ CAAAGA CAAAGG TA CAAAGA CAAAGG TA CAA AG Ο ΤΑ CAAACA CAAAGG TA CAAAGA CAAAGG TA CAA AG Ο ΤΑ CAAACA CAAAGG TA CAAAGG TA CAAAGA CAAAGG TA CAAAGG TA CAAAGA CAAAGA CAAAGG TA CAAAGA CAAAGA CAAAGG TA CAAAGA CAAAGA CAAAGG TA CAAAGA CAAAGA CAAAGA CAAAGA CAAAGG TA CAAAGA CAAAGG TA CAAAGA CAAAGG TA CAAAGA CAAAGG TA CAAAGA CAAGA CAAAGA CAA
TA TCTARGAGACACAAAGGTGACAC TA TCTARGACACAAAGGTGACAC TA TCTARGACACAAAGGTGACAC TA TCTARGACACAAAGGTGACAC TA TCTARGACACAAAGGTGACAC TA TCTARGACACAAAGGTGACAC TA TCTARGACACAAAGGTGACAC AG CTTARGACACAAAGGTACAC AG CTTARGACACAAAGGTACAC AG CTTARGACACAAAGGTGACAC AG CTTARGACACAAAGGTGACAC AA - GCTTCGAGAAACACAAAGGTGACAC AA - GCTTRGAGAACACAAAGGTGACAC AG CTTRARGAACACAAAGGTGACAC TA - TCTRARGAACACAAAGGTGACAC AG CTTRARGAACACAAAGGTGACAC AG CTTRARGAACACAAAGGTACACAC AG CTTRARGAACACAAAGGTACACAC AG CTTRARGAACACAAAGGTACACACAA AG CTTRAAGAACACAAAGGTACACACAA
ΤΑ ΤΣΤΑ Ι Α Ο ΑΛΑΛΟΚΟΥΑΛΟΥΤΑΛΟΥ ΤΑ ΤΤΤΑ Ι Α Ο ΑΛΑΛΟΚΟΥΑΛΟΥΤΑΛΟΥΤΑΛΟΥ ΤΑ ΤΤΤΑ Ι Α Ο ΑΛΑΛΟΥΛΑΛΟΥΤΑΛΟΥΤΑΛΟΥ ΤΑ ΤΓΤΑ Ι Α Ο ΑΛΑΛΟΥΓΑΛΟΥΤΑΛΟΥΤΑΛΟΥΤΑΛΟΥ ΤΑ ΤΓΤΑ Ι Α Ο ΑΛΑΛΟΥΓΑΛΟΥΤΑΛΟΥΤΑΛΟΥΤΑΛΟΥ ΤΑ ΤΓΤΑ Ι Α Ο ΑΛΑΛΟΥΓΑΛΟΥΤΑΛΟΥΤΑΛΟΥ Α Ο Ο ΤΤΑ Α ΑΛΑΛΟΥΛΑΛΟΥΤΑΛΟΥΤΑΛΟΥ Α Ο Ο ΤΤΑ Α Ο ΑΛΑΛΟΥΓΑΛΟΥΤΑΛΟΥΤΑΛΟΥ Α Ο Ο ΤΤΑ Ι Α Ο ΑΛΑΛΟΥΤΑΛΟΥΤΑΛΟΥ Α Ο ΤΤΑ Ι Α Ο ΑΛΑΛΟΥΛΟΥΤΑΛΟΥΛΟΥΤΑΛΟΥ Α Ο ΤΤΑ Ι Α Ο ΑΛΑΛΟΥΛΟΥΤΑΛΟΥΛΟΥΤΑΛΟΥ Α Ο ΤΤΑ Ι Α Ο ΑΛΑΛΟΛΟΥΛΟΥΤΑΛΟΥΛΟΥΤΑΛΟΥ Α Ο ΤΤΑ Ι Α Ο ΑΛΑΛΟΛΟΥΛΟΥΤΑΛΟΥΛΟΥΤΑΛΟΥ Α Ο ΤΤΑ Ι Α Ο ΑΛΑΛΟΛΟΥΛΟΥΤΑΛΟΥΛΟΥΥΛΟΥ Α Ο ΤΟ ΤΟ ΑΛΑΛΑΛΟΥΛΟΥΛΟΥΤΑΛΟΥΛΟΥΛΟΥΛΟΥ Α Ο ΤΟ ΤΟ ΑΛΑΛΑΛΟΥΛΟΥΛΟΥΤΑΛΟΥΛΟΥΥΛΟΥΥΛΟΥΥΛΟΥΥΛΟΥΥΛΟΥΥΛΟΥΥΛΟΥΥΛΟΥΥ
ТА ПСТАРАБАКСА СЛАЛОЗТКА СК. ТА ПСТАРАБАКА САААСТОРАСКА ТА ПСТАРАБАКА САААСТОРАСКА ТА ПСТАРАБАКА САААСТОРАСКА ТА ПСТАРАБАСА САААСТОРАСКА АС ССТАРАСААСА СААСТОРАСКА АС ССТАРАСКА САААСТОРАСКА АС ССТАРАСКА САААСТОРАСКА АА ССТСАРАВАСА САААСТОРАСКА ССТССАРАСАС САААСТОРАСКА ССТССАРАСКА САААСТОРАСКА ССТССАРАСКА САААСТОРАСКА ССТССАРАСКА САААСТОРАСКА ССТССАРАСКА САААСТОРАСКА ССТССАРАСКА САААСТОРАСКА СОСТССАРАСКА САААСТОРАСКА СОСТСАРАЛАСА САААСТОРАСКА СОСТСАРАЛАСАСААСТОРАСКА СОСТСАРАЛАСТОРА ССТАРАСТОРАСКА СОСТСАРАЛАСТОРА ССТАРАСТОРАСКА СОСТСАРАЛАСТОРА ССТАРАСТОРАСКА СОСТСАРАЛАСТОРА ССТАРАСТОРАСКА СОСТСАРАЛАСТОРА ССТАРАСТОРАСКА СОСТСАРАЛАСТОРА ССТАРАСТОРАСКА СОСТСАРАЛАСТОРА ССТАРАСТОРАСКА СОСТСАРАЛАСТОРА ССТАРАСТОРАСКА СОСТСАРАЛАСТОРА СОСТАРАСТОРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАСТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАТИРАСКАТАРАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАСТОРА СОСТАРАТИРАСКА СОСТСАРАЛАСТОРА СОСТАРАСТОРА СОСТАРАСТОРА СОСТСАРАЛАСТОРА СОСТАРАСТОРА СОСТСАРАТИРА СОСТСАРАСТОРА СОСТАРАСТОРА СОСТСАРАТИРАСТОРА СОСТАРАСТОРА СОСТСАРАСТОРА СОСТСАРАЛАСТОРА СОСТСАРАСТОРА СОСТСАРАЛАСТОРА СОСТСАРАСТОРА СОСТСАРАЛАСТОРА СОСТСАРАСТОРА СОСТСАРАЛАСТОРА СОСТСАРАСТОРА СОСТСАРАЛАСТОРА СОСТСАРАЛАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАЛАСТОРА СОСТСАРАСТОРА СОСТСАРАЛАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАРАСТОРА СОСТСАЛАСТОРА
ΤΑ ΤΕΤΑΙΑΛΟΧάζΑ ΚΑΛΑΔΟΤΟΛΟΥ ΤΑ ΤΕ ΤΑΙΑΛΟΧά ΑΛΑΛΟΥΓΙΑΛΟ ΤΑ ΤΙ ΤΑΙΔΟΛΑΛΟΛΑΛΟΥΓΙΑΛΟ ΤΑ ΤΙ ΤΑΙΔΟΛΑΛΟΥΓΙΑΛΟΥΓΙΑΛΟ ΤΑ Ο ΤΙΑΛΟΛΑΛΑΛΟΛΟΥΓΙΑΛΟΥ ΑG Ο ΤΙΑΛΟΛΑΛΑΛΟΛΑΟΟΥΓΙΑΛΟΥ ΑG Ο ΤΙΑΛΟΛΑΛΑΛΟΛΑΟΟΥΓΙΑΛΟΥ ΑΛ Ο ΤΙΑΛΟΛΑΛΑΛΟΛΑΛΟΥΓΙΑΛΟΥ ΑΛΟ ΤΙΑΛΟΛΑΛΟΛΟΛΑΟΥΓΙΑΛΟΥ ΑΛΟ ΤΙΑΛΟΛΑΛΟΛΟΛΑΟΥΓΙΑΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΥ ΑΓ Ο ΤΙΑΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟΛΟ
ТА ТСТА ГАГАЛА С СЛАДОТІЗА СІС ТА ТСТА ГАГАЛА СА СЛАДОТІЗА СА СТА ТСТА ГАГАЛА СА СЛАДОТІЗА СА ГАСАЛАВОТІЗА СА АС СТТА ГАГАЛА СА СДАДОТІЛА СА ГАС СТТА ГАГАЛА СА СЛАДОТІЛА СА ГАС СТТА ГАГАЛА СА СЛАДОТІЗА СА ГАС СТТА ГАГАЛА САЛАОТІЗА СА ГАС СТТА ГАГАЛА САЛАОТІЗА СА СС СТТА ГАГАЛА САЛАОТІЗА СА СС ПСТА ГАГАЛА САЛАОТІЗА СА ГАС ТТА ГАГАЛА САЛАОТІЗА СА ГАС ТТА ГАГАЛА САЛАОТІЗА СА ГА ПСТА ГАЛАЛА САЛАОТІЗА СА ГА ПСТА ГАЛАЛА САЛАОТІЗА СА ГА ПСТА ГАЛАЛА САЛАОТІЗА СА ГА ПСТО ГАЛАЛА САЛАОТІЗА СА ГА ПСТО ГАЛАЛА САЛАОТІЗА СА ГА ПСТО ГАЛАЛА САЛАОТІЗА СА ГА ПСТО ГАЛАЛА САЛАОТІЗА СА
ΤΑ ΤΟ ΤΑ ΒΑ GARCA CA ΑλΟσΤΟΆ COL AG Ο ΤΤΑ Ά GARCA CA ΑλΟσΤΟΆ COL AG Ο ΤΤΑ Ά GARCA CA ΑλΟσΤΑΛΑ ΤΟ AG Ο Ο ΤΟ ΑλΟΑΛΑ CA CA ΑλΟσΤΑΛΑ ΤΟ AG Ο Ο ΤΟ ΑλΟΑΛΑ CA CA ΑλΟσΤΟΆ CA AA GO TO GARCA CA ANO TO ACA AA GO TO GARCA CA ANO TO ACA AG Ο Ο ΤΑ ΒΑ CAACA CA ANO TO ACA GO Ο Ο ΤΑ ΒΑ CAACA CA ANO TO ACA GO Ο ΤΤΑ ΒΑ CAACA CA ANO TO ACA TA Ο ΤΑ ΔΑ CAACA CA ANO TO ACA TA Ο ΤΑ ΔΑ CAACA CAANGO TO ACA AG ΙΟ ΤΟ ΑΛΑΛΑ CACAN CO TO ACAC
ΑG ΟΥΤΑ ΑΆ CAACA CAA COSTANTY AG ΟΥΤG ΆΑ ΆΑΑ CAT CAA GOST ΛΆ ΤΥΤ ΑG ΟΥΤG ΆΑ GAA ΆΑ CAA GOST ΛΆ CAA ΑΑ GOYT G ΓΑ GAA CAA A GOST CAACA ΑΑ GOYT G ΓΑ GAA CAA A GOST CAACA GO ΟΥΤΑ ΓΑ GAA CAA A GOST CAACA GG ΟΥΤΑ ΓΑ GAA CAA A GOST CAACA TT ΤΥΤΑ ΓΑ GAA CAA A GOST ΛΑ TC TT ΤΥΤΑ ΓΑ GAA CAA A GOST ΛΑ TC TA ΤΥΤΑ ΓΑ GAA CAA A GOST ΛΑ TC
ΑG ΟΥΤΑ ΑΆ CAACA CAA COSTANTY AG ΟΥΤG ΆΑ ΆΑΑ CAT CAA GOST ΛΆ ΤΥΤ ΑG ΟΥΤG ΆΑ GAA ΆΑ CAA GOST ΛΆ CAA ΑΑ GOYT G ΓΑ GAA CAA A GOST CAACA ΑΑ GOYT G ΓΑ GAA CAA A GOST CAACA GO ΟΥΤΑ ΓΑ GAA CAA A GOST CAACA GG ΟΥΤΑ ΓΑ GAA CAA A GOST CAACA TT ΤΥΤΑ ΓΑ GAA CAA A GOST ΛΑ TC TT ΤΥΤΑ ΓΑ GAA CAA A GOST ΛΑ TC TA ΤΥΤΑ ΓΑ GAA CAA A GOST ΛΑ TC
ΑG Ο ΤΙ Ο ΑΝΆΑ ΑΛΤΙ ΤΑ ΤΟ ΤΟ ΤΑ Ο ΤΑ Ο ΤΑ Ο ΤΑ Ο ΤΑ Ο ΤΑ
ΑG ΟΥΤΟΑΝΘΑΝΆΑCΑΑ Ο ΟΥΤΑΊΖΟΑ ΑΑ GOTTGEAGAAA CACAAAGOTTACAC ΑΔ GOTTGEAGAACACAAAGOTTACAC ΑG OTTABACAACACAAAGOTTACAC Ο Ο ΟΥΤΤΕΛΘΑΔΑΑCΑCAAAGOTTACAC ΥΓΟ ΠΥΤΑΘΑΘΑΑCΑCAAAGOTTACAC ΑΔ ΠΥΤΑΘΑΘΑΑCΑCAAAGOTTACAC ΑΔ ΟΥΤΔΑΑGAACACAAAGOTTACAC ΑΔ ΤΟΥΓΔΑΑΑΑΑCACAAAGOTTACACA ΑΔ ΓΟΥΤGEAAAACACAAAGOTTACACA
ΙΑ GCTTGGAUAACACAAAGOTGACA AA GCTTGGAQAACACAAAOGTGACAC GG CTTAGAGACACAAAOGTGACAC GG CTTAGAGACACAAAOGTGACAC TG TTTAGAGAACACAAAOGTGACAC TA TTTAGAGAACACAAAOGTAATG IA TTTAGAGAACACAAAOGTAATG IA TTTAGAGAACACAAAOGTAATG IA TTTAGAGAACACAAAOGTAATG IA TTTAGAGAACACAAAOGTAATG IA TTTAGAGAACACAAAOGTAATG
λΑ GCTTGGAGACACAAAGTGACA AG CTABACAACACAAAGTGACA GG CTTGAGAACACAAAGGTGACA GG CTTGAGAACACAAAGGTGACA TTA CTABAGAACACAAAGGTGACA AG CTABAGAACACAAAGGTAACG AG CTABAGAACACAAAGGTAACG AG CTGAAAAACACAAAGTTAACAC AG CTGAAAAACACAA
AG Ο ΤΑΒΑΟΑΑCΑ CAAAGOTGA CAA GG Ο ΤΤΕΛΘΑΟΑ CAAAGOTGA CAA TG ΠΤΑΘΑΟΑ CAAAGOTGA CAA TA ΠΤΑΘΑΘΑ CAAAGOTGA CAA GG Ο ΤΑΑΑΔΑΑ CACAAGOTA TO TA ΤΟ ΤΑΘΑΘΑ CAAAGOTA CAA AG Ο ΤΟ ΤΟ ΔΑΑΑΑ CAAAGOTA CAA AG ΤΟ ΤΟ ΔΑΑΑΑ CAAAGOTA CAA AG Ο ΤΟ ΤΟ ΔΑΑΑΑ CAAAGOTA CAA
AG ΤΟΤΑΡΑΓΑΝΤΑ ΟΥΛΑΘΟΤΟΛΟΥ GG ΟΤΤΑΡΑΓΑΝΟΛΟΟΑΛΑΟΟΤΟΛΟΥ TG ΠΟΤΑΡΑΓΑΝΟΛΟΛΑΛΟΟΤΟΛΟΥ NG ΠΟΤΑΡΑΓΑΝΟΛΟΛΑΙΟΟΤΟΛΟΥ NG ΟΤΤΑΡΑΓΑΝΟΛΟΛΟΛΟΙΟΤΑΛΟΥ TA ΠΟΤΑΡΑΓΑΝΟΛΟΛΟΥ ΑG ΤΟΤΟΛΛΑΛΟΛΟΛΟΥ AG ΤΟΤΟΛΛΑΛΟΛΟΥ AG ΤΟΤΟΛΛΑΛΟΛΟΥ ΑG ΤΟΤΟΛΛΑΛΟΛΟΥ ΑG ΤΟΤΟΛΛΑΛΟΛΟΥ ΤΑ ΠΟΤΟΛΛΑΛΟΛΟΥ ΤΑ ΠΟΤΟΛΛΑΛΟΛΟΥ ΤΑ ΠΟΤΟΛΛΑΛΟΛΟΥ ΤΑ ΠΟΤΟΛΛΑΛΟΛΟΥ ΤΑ ΠΟΤΟΛΛΑΛΟΛΟΥ ΤΑ ΠΟΤΟΛΛΑΛΟΛΟΥ ΤΑ ΠΟΤΟΓΙΑΛΟΛΟΥ ΤΑ ΠΟΤΟΓΙΑΛΟΛΟΥ ΤΑ ΠΟΤΟΓΙΑΛΟΛΟΥ ΤΑ ΠΟΤΟΓΙΑΛΟΛΟΥ ΤΑ ΠΟΤΟΓΙΑΛΟΛΟΥ ΤΑ ΠΟΤΟΓΙΑΛΟΛΟΥ ΤΑ
GG ΟΤΤΓΑΛΟΑΛΟΑΟΆΑΟΑΟΑΑΟΑΟΤΑΑΟΑ ΤG ΟΤΤΑΓΑΛΟΑΛΟΑΟΑΑΑΟΑΟΑΑΟΑ Α.Ο. ΤΑ ΠΤΑΓΑΛΟΑΛΟΑΟΑΑΟΑΟΑΟΑ Α.Ο. ΤΑ ΟΤΤΑΛΑΟΑΛΟΑΟΑΑΟΑΟΑΟΑΟΑ Α.Ο. ΤΑ ΟΤΤΟΑΛΑΑΛΟΑΟΑΑΟΑΟΤΑΑΟΑ Α.Δ GOCTGGAAAACACAA
тететраларалараларалараларалараларалараларал
та ПТРАВАСАСАЛАССТАЛОСТОАСОС NG СТАЛОЛАСАСАЛОСТАТСС ПТ ПТРАВАСАЛАССТАЛОСТАЛСС NG ПСТОДАЛАЛСССАЛОСТАЛССА NG ПСТОДАЛАЛСАСАЛОСТАЛССА
AGCITAAAGAACACAAAGGGTAATCK ITATITAGAGAACACAAAGGTGACAC AGTICTGAAAAACACAAAGGTAACACA IAGTICTGAAAAACACAAAGAGTAACACA
AGCITAAAGAACACAAAGGGTAATCK ITATITAGAGAACACAAAGGTGACAC AGTICTGAAAAACACAAAGGTAACACA IAGTICTGAAAAACACAAAGAGTAACACA
ТА Т <u>СТАРАВА</u> АСАСАААВСТВАСАС АGТСТСАААААСАСААВСТААСАС ААССТСБААААСАСААВТААСАС
AGTCTGANANACACAA (BGTANCAC AAGCCTGGAANACACAA (A) STANCAC
AGTCTGAAAAACACAAGEGTAACAC AAGCCTGGAAAACACAAGAGTAACAC
AAGOCTGGAAAACACAAGAGIDAACAC
ALL LUCAMMALAL MALLER LAAT AT
CT GAGAACACAAAGGTGACAC
AGTOTTCCCAAGTOTGCAGGAGCAGC
AGTCTCCCTATGCCTGCAGGAGCAGC
AAGTCTCOCTATGCCTGCAGGAGCAGC
AAGTCTCCCTATGCCTGCAGGAGCAGCI
AAGTCTCCCTATGCCTGCAGGAGCAGC
AAGTCTCCCTATGCCTGCAGGAGCAGC
AAGTCTCCCTATGCCTGCAGGAGCAGCI
AAGTCTTCCCATGTCTGCAGGAGCAGCI
AAGTCTCCCCATGTCTGCAGGAGCAGC
AAGTCICCCTATATCIGCAGGAGCCGC
AAGIGTCCCCATGTCTGCAGGAGCAGC
AGTCTCCCCATTCCTGCAGGAGCAGC
AGTERCECATTICERCAGGAGCAGC
TAGTCTCCCCACCCTGCAGGAGCAGC
AAGTCTCCCCATGCCTGTAGGAGCAGCT
GTCPCCCCATGCCTGCAGGAGTAGCT
AAGTCTCCCTATGCCTGCAGGAGCAGCT
AAGTCFCCCCAGGCCTGCAGGAGCAACT
AAGICTCCCTATGCCTGCAGGAGCAGC
AAGTCTCCCTATGTCTGCAGGAGCCGC
AAGTCTCCCTATGTCTGCAGGAGCCGC
AAGTCTCCCTACATCTGTGGGAGCCGC
TAAGTCTCCC ATGCCTGCAGGAGCAGC
CAGITACATAGATGGATAATCCTATTG/
CAGTCACATGGATGGATAATCCTACAE
CAGTCACATGGATGGATAATCCTATAG
CASTCACATGGATGGATAATCCTATAG
CAGTCACATGGATGGATAATCCTATAG
CAGTCACATGGATGGATAATCCTATAG
CAGITACATGGGTGGATAATTCTATTG
CAGTCACATGGGGGGGGATAATICTBITG
SCACICACTICGATGATGATCCI GI TG
CAGTCACATGGATAGATAATCCTATTA
CAGTCACGTGGATGGATAATCCTATTG/
CAGTCACCIGGATGGATAATCCTATIG
TCACATGGATGGATAATCCTATAG
FCACATGGATGGATAATCCTATAG CAGTCTCATGGATGGATAATCCTATTG
CAGTC TCATGGATGGATAATCCTATTG CAGTC GCATGGATGGATAATCCTATTG
cagte illatggatggataatectattg cagte geatggatggataatectattg <u>-g</u> agte acatggatggataatectattgg
CAGTC ILATGGATGGATAATCCTATTG CAGTC SCATGGATGGATAATCCTATTG GAGTCACATGGATGGATAATCCTATTG ICAGTCACATGGATGGATAATCCTAT <u>G</u> G
CAGTCTLATGGATGGATAATCCTATTG ICAGTCGCATGGATGGATAATCCTATTG GAGTCACATGGATGGATAATCCTATTG ICAGTCACATGGATGGATAATCCTATTGG ICAGTCACATGGATGGATAATCCCTATGGI ICTGTCACATGGTTGGAAACCCCGGTTG
CASTC TLATGATGATAGATAATCCTATTG ICASTC SCATGATGATGATAATCCTATG GASTCACATGATGATAATCCTATG ICASTCACATGATGATAATCCTATG ILTISTCACATGGATGGATAATCCT ICASTCACATGGATGGATAATCCT
CAGTC TLATGATGATGATAATCCTATTG CAGTCGCATGGATGGATAATCCTATTG GAGTCACATGGATGGATAATCCTATTG LCAGTCACATGGATGGATAATCCTATGG LCAGTCACATGGATGGATAATCCTATGG CAGTCACATGGATGGATAATCCT
CAGTC TLATGATGATGATAATCCTATTG CAGTCGCATGGATGGATAATCCTATTG GAGTCACATGGATGGATAATCCTATTG LCAGTCACATGGATGGATAATCCTATGG LCAGTCACATGGATGGATAATCCTATGG CAGTCACATGGATGGATAATCCT
CAGTC TLATGGATGGATAATCCTATTCI GAGTCGCATGGATGGATAATCCTATTGI GAGTCACATGGATGGATAATCCTATTGI ICASTCACATGGATGGATAATCCTATGGI ICTGTCACATGGATGGATAAACCCGGTG ICTGTCACATGGATGGATAATCCTATTGI ICACATGGATGGATGATAATCCTATTGI ICACATGGATGGATAATCCTATTGI ICASTCACATGGATGATAATCCTATTGI
CAGTC TLATGATGATAGATAATCCTATTGI ICAGTC GCATGGATGGATAATCCTATTGI GAGTCACATGGATGGATAATCCTATTGI SCATCACATGGATGGATAATCCTATGI TIGTCACATGGATGGATAACCCOGGTG TIGTCACATGGATGGATAATCCT CACATGGATGGATAATCCTATTGI SCAGTCACATGGATAGATAATCCTATTGI SCAGTCACTTGGATAGATAATCCTATTGI
CAGTC TLATGATGATAGATAATCCTATTCI CAGTC GCATGGATGGATAATCCTATTCI CAGTCACATGGATGGATAATCCTATTCI CAGTCACATGGATGGATAATCCTATCI CAGTCACATGGATGGATAATCCTATCI CAGTCACATGGATGATAATCCTATTCI CCATCACTGGATAGATAATCCTATTCI CCATCACTGGATAGATAATCCTATTCI CAGTCACTTGGATAGATAATCCTATTCI CAGTCACTTGGATAGATAATCCTATTCI CAGTCACTTGGATAGATAATCCTATTCI
CAGTC TLATGATGATAGATAATCCTATTGI CAGTCGCATGGATGATAATCCTATTGI GAGTCACATGGATGGATAATCCTATTGI CLATCACATGGATGGATAATCCTATTGI TITGTCACATGGATGGATAATCCTATTGI CAGTCACATGGATGGATAATCCTATTGI CAGTCACATGGATGGATAATCCTATTGI CAGTTGGATGATGATAGATAATCCTATTGI CAGTTGGATGATGATAGATAATCCTATTGI CAGTTGGATGATGATAATCCTATTGI CAGTTGGATGATGATAATCCTATTGI CAGTTGGATGATGATAATCCTATTGI CAGTTGGATGATGATAATCCTATTGI CAGTTGGATGATGATAATCCTATTGI CAGTTGGATGATGATAATCCTATTGI CAGTTGGATGATGATAATCCTATTGI CAGTTGGATGATAATCATAATCCTATTGI CAGTTGGATGATAATCCTATTGI CAGTTGATGATGATAATCCTATTGI CAGTTGATGATAATCCTATTGI CAGTTGATGATAATCCTATTGI CAGTTGATGATAATCCTATTGI CAGTTGATGATAATCCTATTGI CAGTTGATGATAATCCTATTGI CAGTTGATGATGATAATCCTATTGI CAGTTGATGATAATCCTATTGI CAGTTGATGATAATCCTATTGI CAGTTGATGATGATAATCCTATTGI CAGTTGATGATAATCCAT
CAGTC TLATGATGATAGATAATCCTATTCI CAGTC GCATGGATGGATAATCCTATTCI CAGTCACATGGATGGATAATCCTATTCI CAGTCACATGGATGGATAATCCTATCI CAGTCACATGGATGGATAATCCTATCI CAGTCACATGGATGATAATCCTATTCI CCATCACTGGATAGATAATCCTATTCI CCATCACTGGATAGATAATCCTATTCI CAGTCACTTGGATAGATAATCCTATTCI CAGTCACTTGGATAGATAATCCTATTCI CAGTCACTTGGATAGATAATCCTATTCI

	(481AGATGTTAATAATAGTGCATGGG-TGCCTGGCCCCACAGATGACTGTTGCCCTGCCC
ENV GENOMIC HERV MDA ENV GENOMIC HERV-K TAN.	(379) (392)	AGATGTTAATBATAGTGCATGGG-TGCCTGGCCCCATAGATGATGATGGCTGCCCTGCC
ENV GENOMIC AC025420	(389)	ATATOTTAAT GATAGTOTATGGO-TACATGGCCCCATAGATGATCG CIGCCCTGCCAAACCIGAGGAAGAAGGGATGAIG
ENV GENOMIC AP000776	(392)	ATATETTAATGATAGTGTATGGE-TACLIGGCCCCA TAGATGATCGCIGCCCTGCCBAACLIGAGGAAGAAGGGATGATG
ENV GENOMIC HERV-KB	(291)	ATATOTTAATGATAGTGTATGOQ-TACCTGGCCCCACAGATGATCG CTGCCCTGCCAAACCTGAGGAAGAAAGCATGATG
ENV GENOMIC HERV-KI ENV HERV-K AF023261	(392) (455)	ATATOTIAATGATAGTETATGGG-TACCIGGCCCCACAGATGATCACTGCCCTGCCAAACCLVAGGAAGAAAGGAATGATG
ENV GEN AL035086	(178)	ATATGTTAATAATAGTSCATGGG-TACCTGGCCOCACGGATGATCGTTGCCCTGCCCAACCAAGGAAGGAATGATG
ENV GENOMIC AL035587	(392)	atatottaataatadsicatggg-taccasgcccccacagacgatigttgcccctgcccaacctmaagaagaatgatg
- ENV GENOMIC AC012068	(380)	ACATOTTAACAATAGTACATGGG-TACCAGAACCCCACAGATGACCGTGCCCTGCCC
ENV GENOMIC AF277315	(389)	ATOTOTTAATAATAGTGCATGAG-IACCAGGCCCCCACAGATGATGTTGCCCCGCCCAACCTGAAAAAGAAGGAATGATG
ENV GENOMIC AF027650 ENV GENOMIC AC078899	(454) (391)	ATATGTTAATAATGBTGTATGGG-TACATGCCCTACAGATGATCGTTGCCCTGCCAAACCCAAGGAAGAAGAAGAATGATG ATATGTTAATAATGTTGTATATGGG-TACATGCCCTACAGATGATCGTTGCCCTGCCAAACCCAAGGAAGAAGAAGAATGATG
ENV GENOMIC HERV-KII	(100)	ATATOTTAATGATAGTGTATGGG-TACCTGGCCCCACAGATGATCGCTGCCCTGCC
ENV GENOMIC AC008813	(412)	ATACOTTANTANTAGTOTATOGO-TACCTGCCCCCACAGATGATCOTTGCCCGGCCAAACCTGAGGAAGAAGAAGAATGATG
ENV GENOMIC AC012309	(391)	A TATGTTAATAA CAGTISTATGGG-TACCTGGCCCCACAGATGATCA CTGCCCTGCCAAACCTGAGGAAGAAGGAATGACG
ENV GENOMIC AL121932	(389)	ATATGTTAATAATAGCIJGIGGC-TACCTGGCCCCACAGATGATTATTGCCCTGCCAAACCTGAGGAAGAAGAATGATG
ENV GENOMIC AD000090 ENV GEN AL160008	(405) (271)	ATATETTAATGATAGTETATUGG-IACCTGUTOCAACAGATGATCGTTGCCCTGCCAAACCTGAGGAAGAAGGATGATG TTACGTTCATSCHAGNETTICGA-BGCCNGAGLCTACAGATTCUCUTGGGCCCTCTCACCCAUAGGAGGAAGGAATGTTA
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(156)	ATATGTTAATGATAGTGTATGGG-TACATGGCCCCCACAGATGATCGCTGCCCTGCC
ENV GENOMIC AF235103	(451)	ATATOTTAATAATAGTGCATGGC-TACGAGGCCCCACAGATGACCGTGCCCTGCCGAACCTGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA
ENV GENOMIC AC026786	(329)	ATACOTTAATAACAOTCCATGGG-TACOACCCCCCATGGATGCCGGGGCCTGCCGGAACCTGAAGAACGAATGATG
ENV GENOMIC AC034203	(432)	ATATGTTAATAACAGTGTATGGG-TACCAGGCCCCACAGATAACCGTGGCCCTGCCCAACCTGAAGAAGAAGGAATGATG
ENV GENOMIC AC018809	(98)	ATATGTTAATGATAGTGTATGGGGTACCTGGCCCCACAGATGATCGCCGGCCCAGAACCTGAGGAAGAAGGATTATG
ENV GENOMIC HERV-K102 AF164610 ENV GENOMIC FRAG. AF260253	(152) (1)	ATATUTTAATGATAGTUTATGGG-IACCUGGAXXCATAGAUGATCGCUCCCUGGAAACCUGAGGAAGGGAUGAUGA
CONSENSUS		ATATGTTAATAATAGTGTATGGG TACCTGGCCCCACAGATGATCGTTGCCCTGCCAAACCTGAGGAAGAAGGAATGATG
	(101)	
ENV GENOMIC HERV MDA	(455)	561 640 ACGARTATTTCCATTGGGTATCCITATCCCCCGTTTGCCTAGGGAAGGCACCAGGATCCITAAT-GCCTACAACCCAAA
ENV GENOMIC HERV-K TAN.	(455) (471)	ATAAATATTTCCATTGGGTATCATTATCCTCCTATTGCCTAGGAAAGGCACCAGGATGTTTAAT-GCCTGCAGTCCAAA
ENV GENOMIC AC025420	(468)	ATAAATATCICCAT/IGGTATCATTATCCTCCTATTTCCCTAGGAGAGCACCAGGATGTTTAAT-GCCTGGAGTCCAAA
ENV GENOMIC AP000776	(471)	
ENV GENOMIC HERV-K8	(291)	
ENV GENOMIC HERV-KI	(471)	ATAAATAT FICCATTGGGTATCGTTATCCTCCTATTTGCCTAGGGAGGAGCACCAGGATGTTTAAT-GCCTGCGCCCAAA
ENV HERV-K AF023261	(534)	ATAAATATTTCCATTGGGTATCGITATCCTCCTATTTGCTTAGGGAGAGGCACCAGGATGTTTAAT-GCCTGCAGAA
ENV GEN AL035086 ENV GENOMIC AL035587	(257) (471)	atgattattettecactoggetattgetatcetectattroccoggegegeaccaggatgetata-bootacaaccaaa ataaatgettecactoggetatcattatcetectattrocctaggeaaggeaaccaggatgettaat-bootacaaccaaa
ENV GENOMIC AD035507	(459)	ATAAACATTTCCATTGGGTATCATTATCCTCCTATTTCTCTGGGAAAGCACCAGGATCCTTAAL-BCCTACAAACCAAA
ENV GENOMIC AF277315	(468)	ATAAATATTTCCATTGGGTATCATTATCCTCCTATTTGCCTAGGGAAGGCACGGGATATTTAAT-TCCTACAACCCAAG
ENV GENOMIC AF027650	(533)	ATAAATATCTCCACTGGATATCCCTACCTACTTCCCCAGGAGAGCACTAGGATGTTTAAI-BCCTGCAAICCAAA
ENV GENOMIC AC078899	(470)	ATAAATATCTCCACTGGATATCGITATCCTCCTATTTGCCTAGGGAGGGGCACTAGGATGTTTAAI-BCCTGCAATCCAAA
ENV GENOMIC HERV-KII	(179)	ATAAATATITCCATGIIJIATCOILATCOILATCCICCTATITCCCCTAGGGAGAGCACCAGGATGTTTAAI-GCCTGCAGTCCAAA
ENV GENOMIC AC008813 ENV GENOMIC AC012309	(491) (470)	ataaatatttccachggetatcattatcctatttgccicggaagagcaccaggatgtttaaf-bcctgcagtcaaaa ataaàtatttctactggetatcetatcctatttgtctaggaagaaccaggatgtttaaf-bcctacaatccaaa
ENV GENOMIC AC012303 ENV GENOMIC AL121932	(468)	ATAAATATTTCCACTGGGTATTGTTATCCTCCTATTTGCCTAGGGAGAACACCAGGATGTTTAAI-GCCTGCAATCCAAA
ENV GENOMIC AD000090	(484)	ATAAATATITTTATTGGGFATCATTATCCTCCTATTFGCCTGGGGGGGGGCACCAGGATGTTTAAT-BCCTGCAGTCCAAA
ENV GEN AL160008	(350)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(235)	ACAAATATITCCACITOGCTATCATTATCCTCCTATTIGCCTACOGAGAGCACTAGGATGCITAAT-CCCTGCAGTCCAAA
ENV GENOMIC AF235103 ENV GENOMIC AC026786	(530) (405)	ATAAACATTTCCATIGGGTATCATTATCCTCCTATTTGCCIGGGARAAGCACCGGGATGCTTAAT-DCCIACAAACCAAA ATAAACATTTCCATIGGGTATCATTATCCTCCTATTTGCCIGGARAAGCACCAGGATGCTTAAT-GCCIACAATCCAAA
ENV GENOMIC AC028788 ENV GENOMIC AC034203	(511)	ATAAACATTTCCATTGGGTATCATTATCCTCCTATTTGCTIGGJAAAGCACCAGGATGCTTAAT-GCCTACAAGCCGAA
ENV GENOMIC AC018809	(178)	ATAAATATTTCCACTGGGTATCGTTATCCTCCTATTTGCCTAGGGAGAGCACCAGGATGCTTAAT-GCCTGCAGTCCAAA
ENV GENOMIC HERV-K102 AF164610	(231)	
ENV GENOMIC FRAG. AF260253	(1)	
CONSENSUS	(561)	ATAAATATTTCCATTGGGTATC TTATCCTCCTATTTGCCTAGGGAGAGCACCAGGATGTTTAAT GCCTGCA TCCAAA
		<u>641</u> 720
ENV GENOMIC HERV MDA	(534)	
ENV GENOMIC HERV-K TAN.	(550)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(547) (550)	
ENV GENOMIC HERV-KB	(291)	
ENV GENOMIC HERV-KI		ATTEGTTAGTAGAAGTACCTACTGTCAGTCCCATCAGTAGATTCACTTATCACATGGTAAGCGGGATGTCACTCAGGCCA
ENV HERV-K AF023261	(613)	ATTGGTTGGTAGRAGTACCTACTGTCAGTCCCHTCAGTAGATTCACTTATCACATGGTAACCGCGATGTCACTCAGGCCA
ENV GEN AL035086		ATTOGTTGGTAGAAGTACCTACTGTCAGIGCCACCAGTGGATITACTTATCACAGGGTAAGTGGAATGTCACTCAGGCCA
ENV GENOMIC AL035587		ACTGGTTGGTAGAAGTACCTACCTFCAGTGCCACCAGTAAATTTACTTATCACATGGTAAGTGJAATGTCACTCGGGCCA
ENV GENOMIC AC012068 ENV GENOMIC AP277315		ACTOGITOGTAGAAGTACCTACTGTCAGTGCCACCAGTAAATTCACTTATCACATGGTAAATGTCAAATGTCACTTGCGTCA
ENV GENOMIC AP277315 ENV GENOMIC AP027650		ATTGGTTGGTAGAAGTACCTACTGTCATTGCCAGCAATAGATTTACTTATCACATGGTAAGTGAAATGTCACTCGGGCCA ATTGGTTGGTACAAGTACCTACTGTCAGTACCATCAGTAGATTCACTTATCACATGGTACAAGGGGAATGTCACTCAGGCCA
ENV GENOMIC AP027850 ENV GENOMIC AC078899		ATTIGTTGTTGTACAAGTACCTACTGTCAGTACCATCAGTAGATTCACTTATCACATGGTAAGGGAAATGTCACTCAGGCCA
ENV GENOMIC HERV-KII	(258)	ATTGOTTGOTAGAAGTACCTACTGTCAGTCACTCACTTACACTTATCACATGGTAAGCGGCATGTCACTCAGGCCA
ENV GENOMIC AC008813	(570)	ATTGGTTGGTAGAAGTACCTACTOTCGGTACCACCAGTAGATTCACTAATCACGCGGTAAGTGGAATGTCACTCAGGCCA
ENV GENOMIC AC012309		ATTGGTTGGTAGAAGTACTTACTTACTGTACCACCAGTAGATTCACTTATCACATGGTAAGCOGAATGTCACTCAGGCCA
ENV GENOMIC AL121932		APTGGTTGGTAAAGGTACCTACTTTCAGTACCACCAGTAGATTTACTTATCACATGGTAAGTGGAATGTCACTCAGGCCA ATTGGTTGGTAGAAGTACCTACTGTCAGICCCATCAGAAGATTCACTTATCACATGGTAAGCGGGATGTCGCTCAGGCCA
ENV GENOMIC AD000090 ENV GEN AL160008		ATTUGTIGGTAGATICCAOTCATAATCAABACCAGTATCC
ENV GENOMIC HEU32496		SINGLAGIARMANALAAGILAINICROSSICHSINIC
ENV GENOMIC AC011467	(314)	ATTGGTTGGTAGAAGTACCTACTGTCAGTCCCATCAGTAGATTCACTTATCACATGGTAAGTOGGATGTCACTCAGGCCA
ENV GENOMIC AF235103	(609)	ATTGGTTGGTAGAAGTACCTACTGTCAGTGCTACCAGTAAATTCACTTATCACA-GGTAAGTGAAATGTCACTCGGGTCA
ENV GENOMIC AC026786		ATTGGTTGGTATAAGTACCTACTGTCAATGCCACCAGTAAATTTACTTATCACATGGTAAGTGGAATGTCAATCGGGTCA
ENV GENOMIC AC034203 ENV GENOMIC AC018809		ATTGGTTGGTAGAAGTACCTACTGTCAGTGGCACCAGTAAATTTACTTATCATGTGATAAGTGGAATGTCACTCGGGTCA ATTGGTTGTTAGAAGTACCTACTGTCAGTCCCATCAGTAGATTCACTTATCACACGGGTAGCAGGATGTCACTCAGGCCA
ENV GENOMIC AC018809 ENV GENOMIC HERV-K102 AF164610	(257)	ATTGETTGTTAGAAGTACCTACTGTCAGTCCCATCAGTAGATTCACTTATCACACGGGTAGCCAGGATGTCACTCAGGCCA ATNGTTGGTAGAAGTACCTACTGTCAGTCCCTCAGTAGATTCACTTATCACACGGCATGGTAGCCAGGATGTCACTCAGGCCA
ENV GENOMIC FRAG. AF260253	(1)	
CONSENSUS		ATTGGTTGGTAGAAGTACCTACTGTCAGT CCA CAGTAGATTCACTTATCACATGGTAAG GG ATGTCACTCAGGCCA

FIGURE 6 CONTD....

		721 800
ENV GENOMIC HERV MDA	(609)	TAAATAATTTACAGGACCCTTCTTATCAAAGATCATTACAATGTAGGCCTAAGGGGAAGGCTTGCCCCAAGGAAAT
ENV GENOMIC HERV-K TAN.	(630)	GGGTARATTATTACAAGACITTTCTTATCAAAGATCATTABAATTTAGACCTAAAGGGAAACCTTGCCCCAAGGAAAT
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(627) (630)	CGGETAAATTATTTACAAGACTTTTCTTATCAAAGATCATTAAAATTTTAGACCTAAAGGGAAACCTTGCCCCAAGGAAAT CGGETAAATTATTTACAAGACITTTCTTATCAAAGATCATTAAAATTTAGACCTAAAGGGAAACCTTGCCCCAAGGAAAT
ENV GENOMIC APOUNTS ENV GENOMIC HERV-KB	(291)	
ENV GENOMIC HERV-KI	(630)	CASOTRAAJTATTTACRAGACTTTICTERTCAAAGATCATTAAAATTTAGACCTTAAAASGAAACCTTGCCTCAAGGAAAT
ENV HERV-K AF023261	(693)	CGGGTAAA
ENV GEN ALO35086	(416)	CAGGTAAATTATTTGCAGGACCCITCTTATCAAAGA1GA1TAAAATCPAGGCCTAAGGGGAAGCCTTGCCCCAAAGAAAT
ENV GENOMIC AL035587 ENV GENOMIC AC012068	(630) (618)	Calatraataattiingaggaogeteettateaaagateattaaaateegggaaggaaggettagggaaggettagggaaggettgeeeeaaggaaat Caaatgaataattiingaggaeeeetteetteotaaggaagateattaaggittaggeetaagggaaagtettgeeeeeaaggaaat
ENV GENOMIC AF277315	(627)	CAGATAAATAATTTACAGGATCCITCTTATCAAAGATCATTAAAATTTAGGCCTAAGAAGCCINGCCCCAAGGAAAT
ENV GENOMIC AF027650	(692)	CGGGTAAA
ENV GENOMIC AC078899	(629)	TGGGTAAATTATTACAGGAGITTTCTTATCAAAGGTCATTAAAATTTAGGCCTAAAGGGAAACATTGCCCCCAAGGAAAT
ENV GENOMIC HERV-KII ENV GENOMIC AC008813	(338) (650)	CGEGTARANTRATTRACAAGACTTITCTTATCAAAGATCATTAAAANTTRAGACCTAAAGGGAAAACCITGCCCCAAGGAAAT CAEGTAAATTATTTACAAGACTTITCTTATCAAAGATCATTAAAANTTIAGGCCTABAGGAAAGCCITACCCCAA <u>AAA</u>
ENV GENOMIC AC012309	(629)	CABGTAAATTATTACAGGACITITCTTATCAAAGATCATTAAAATGTAGGCCTAAAGGGAAACCTTGCCCCACGGAAAT
ENV GENOMIC AL121932	(627)	CAGGTAAATTATTTACAGGACTTTTTCTTATCAAAAGATCATTAAAAATTTAGGCCTAAAGGGAAACCTIGCCCCAAGGAAAT
ENV GENOMIC AD000090	(643)	CAGGTAAATTATTTACAAGACTTTTCTTATCAAAGATCATTAAAATTTAGACCTAAAGGGAAACCTTGCCCCGAGGAAAT
ENV GEN AL160008 ENV GENOMIC HEU32496	(482) (441)	TITTIA JAA -AGITTOGTICAATTGGAACAGITIAAGCCCAGAAAAGAGAGGTGICAACAACC
ENV GENOMIC AC011467	(394)	GGGTAAATAATTTACAAGACTTTOTTATCAAAGATCATTAAAATTTAGACCTAABGGCAACCTTUCCCCGAGGAAAT
ENV GENOMIC AF235103	(688)	CAAA TAAATAATTTACAGGATTCTTCCTATAAAAGATCATTAAAATTTAGGCTTAAGGCAAAATCTTGCCCCCAAGGAAAT
ENV GENOMIC AC026786	(564)	CAAANGAATAATTTACAGUATTCITCCTATCAAAGATCATTAAAATTTAGGCCTAAGGGAAAAACCAJGCCCCGAAGGAAAT
ENV GENOMIC AC034203	(670)	CALLTAR TAR TAR TAR CALSTIC TTC CLATCRARGE TCATTRACATTER GCCTRARGE ARCANGCANCCONCERCER AND
ENV GENOMIC AC018809 ENV GENOMIC HERV-K102 AF164610	(337) (390)	CGGGTAAATAATTTACAAGACTTTTCTTATCAAAGATCATTAAAATTTTAGACCTAAAGGGAAACCTTGCCCCAAGAAAAT
ENV GENOMIC RERV-RIDZ AF164610 ENV GENOMIC FRAG. AF260253	(1)	CGGGTAAATTATTTACAAGACTTTTCTTATCAAAGATCATTAAAATTTAGAUTTAAAGGGAAACCTTGCCCCAAGGAAAT
CONSENSUS	(721)	C GETAAAT ATTTACA GACTITICITATCAAAGATCATTAAAATTTAG CCTAAAGGAAACCIIGCCCCAAGGAAAT
		801 880
ENV GENOMIC HERV MDA	(685)	TCCCAAAGAATCAAAAAGCCCAGAAGICTTAGICTGCIGAGAATGTGTGGCTGATACTGCACTG-TAGTACAAAACAATG
ENV GENOMIC HERV-K TAN. ENV GENOMIC AC025420	(710) (707)	rccranagaancaana aa tacagaagtittagiitgggaagaactotgroggcaatagtgcggtgatattacaaaachatg rcccanagaancaana aa tacagaagtittagiitgggaagaatgtgroggcaatagtgtgstgatattacaaaachatg
ENV GENOMIC AP000776	(710)	TCCCAMBERT CARRENT TACAGE ASTTTTACT TEGES ANGAST GTOTOGCCA A REST GTOTOTOGC ATATTACA A A A CARACTA A CARAC
ENV GENOMIC HERV-K8	(291)	
ENV GENOMIC HERV-KI	(710)	TCOCAAAGAATCAAAAAATACAGAAGTITTAGTITAGGAAGAATGTGTGGCCAATAGTGCGFGATATTACAAAACAATG
ENV HERV-K AF023261	(701)	
ENV GEN AL035086 ENV GENOMIC AL035587	(496) (710)	ICCURANGENTCARARAGECCNEARAGTITTAGTTTGACAAGAATGTGTGGUTGATAGEACGGTGATATTACAAAACRATG TCQURANGAATCARARAGECCREARAGCCTTAGTTTGGURAGAATGTGTGGCTRATACTGCGTGGTGTATTACAAAACRATG
ENV GENOMIC AC012068	(698)	TCCAAAAGAATCAAAAGACCCAGAAGICTTAGTTTIGGAAGAATGTGTGGCTGATACTGCAGIGGTAGTACAAAACAATG
ENV GENOMIC AF277315	(707)	ICCCAAAGAATCAAAAGCCCAAGAACACCAAGTTTGGGAAAAATGTGGTGATACGCGTGTGATACGCAAAAACAATG
ENV GENOMIC AF027650	(700)	
ENV GENOMIC AC078899 ENV GENOMIC HERV-KII	(709) (418)	PCCCAAAGGATCAAAAGACATAGAAGTTTTAGTTTGGGAAGAATGTGTGGGCCAAGAGTGCAGTGATATTACAAAATGATG
ENV GENOMIC AC008813	(110)	ncolaassaancaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
ENV GENOMIC AC012309	(709)	TCCCAAGGTATCAAAAGACACAGAAGTTTTAGTTTAGGTAGG
ENV GENOMIC AL121932	(707)	TCCCAAAGGATCAAAAGACACAGAAGTTTTAGTTTTGGAAGAATGTGTGGCCAATAGAGGAGTGATATTACAAAACGATG
ENV GENOMIC AD000090	(723)	TICCAAAGGATCAAAAAATGIBEAAGCITTAGTITTGGGAAGAATGIGGCCAATAGTGCAATAGTAATACAAAAACAGTG
ENV GEN AL160008 ENV GENOMIC HEU32496	(543) (441)	<u>TCANTAATGGPCAAAGGATTTAGAAACATTGATTTGGAAGGATTGCATCALTGATCACCACAGTGTACTGCAAAATAATC</u>
ENV GENOMIC AC011467	(474)	TCOCAAAGGATCAAAAAAATACAGAAGTTTTAGTTTGGGAAGAATGTGTGGCCAATAGTTGABTAATATTACAAAACAATG
ENV GENOMIC AF235103	(768)	TCCAMAGAATAAAAAACACTCAGAAGTCTTAGTTTBGGAAGAATGTGTGGCTGATACGCGOTGGTATTACAAAACAATA
ENV GENOMIC AC026786	(644)	TCCAAAAGAATCAAAAGACCCAGAAGICTTAGTTTGGGAACAATGTGCGGCTGATACTGCAGTGGTACTACAAAACAATA
ENV GENOMIC AC034203 ENV GENOMIC AC018809	(750) (417)	TCCAAAAGAATCAAAAGACCCAAAAAGTCTTAGTTTAGGAAGAATGTGTGGGCCAATAGTGGAAGGAA
ENV GENOMIC HERV-K102 AF164610	(470)	TCLCAAAGAATCAAAAAATACAGAAGTTTTAGTTTGGGAAGAATGTGTGGGCAATAGTGCGTGATATTACAAAACAATG
ENV GENOMIC FRAG. AF260253	(1)	
CONSENSUS	(801)	TCCCAAAG ATCAAAA A CAGAAGTTTTAGTTTGGGAAGAATGTGTGGGC AATAGTGC GTGATATTACAAAACAATG
ENV GENOMIC HERV MDA	(764)	881 960 AATTITBAACTATGATAGACTGGGTCCCCTTGAGGCCAATTATATCATA
ENV GENOMIC HERV-K TAN.	(790)	AATTOIGAACTATTATAGATTGGGCAGCTCGAGGTCAATTCTACCACA
ENV GENOMIC AC025420	(787)	AATTTIGGAACTATTATAGATTGGGCACCTCGAGGTCAATTCTACCACAATTGCTCA
ENV GENOMIC AP000776 ENV GENOMIC HERV-K8	(790)	ARTITEGAACTATTATAGATTGGGGACCTCHAGGTCAATTCTECCACA
ENV GENOMIC HERV-KB	(291) (790)	AATTGGAACLATTATAGATIAGGCACCTCGAGGTCAATTCTACGACA
ENV GENOMIC HERV-KI ENV HERV-K AF023261	(701)	LEAD AND A CHARLEN AND A AND A CHARLEN AND A CHAR
ENV GEN AL035086	(576)	AATTOGGAACCATTATAGATTOGGCACCTCGAGGTCAATTCTACTGCA
ENV GENOMIC AL035587	(790)	AATTTGGAACTACTATATATATATATAATAGTTATATATTATATATA
ENV GENOMIC AC012068	(778)	AATTTGGAACTAITTATAGACTAGGCCCCTCGAGGCCAATTATATTAT
ENV GENOMIC AF277315 ENV GENOMIC AF027650	(787)	BATTTEGAACTATTATAGACTGGGCCCCTTCAAGGCBATTATATTAT
ENV GENOMIC AC078899	(789)	AATTCAGAACTATTATAGATTGGGCAACTCAAGGTCAATTCTAACACA
ENV GENOMIC HERV-KII	(498)	AATTOGGAACTATTATAGATTGGGCACCTCGAGGTCAATTCTACCACA
ENV GENOMIC AC008813	(809)	AATTOGGAACTATTATAGATTGGGCACCTCGAGGTCAATTCTACCACA
ENV GENOMIC AC012309	(789)	AATTIGGAACTATIATAGATIGGGCACCTCGAGGTCAATTCTACCACA
ENV GENOMIC AL121932 ENV GENOMIC AD000090	(187)	AATTTGGAACIGTTATAGATAGGGCACCTCGAGGTCAATTCTACCACA
ENV GEN AL160008	(623)	CCTATEGAATCATCATTGATTGGG
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(554)	AATTCAGAACTATTATAGATTGGACATCTGGAGGTCAATTCTACCACA
ENV GENOMIC AF235103 ENV GENOMIC AC026786	(848)	AATTTGGAACTATTATAAACTGGGCCCCCTTGAGGCTAATTATATTATTATG
ENV GENOMIC AC026786 ENV GENOMIC AC034203	(724) (830)	AATTTGGAACTATTATAGACTGGGCCCCTCGAGGCCAATTATATTATG ACTTTGGAACTATTATAGACTGGGCCCCTCGAGGCCAGTTATATTATG
ENV GENOMIC ACO18809	(497)	AATTIGGAACTATTATATATATATATATATATATATATATATATA
ENV GENOMIC HERV-KL02 AF164610	(550)	AATTTGGAACTATTATAGATTGGGCACCTCGAGGTCAATTCTACCACA
ENV GENOMIC FRAG. AF260253 CONSENSUS	(1)	
CONSENSOS	(081)	AATTTGGAACTATTATAGATTGGGCACCTCGAGGTCAATTCTA CACA ATTGC CA

		9611040
ENV GENOMIC HERV MDA	(820)	GICCAGACTCATTOATGTTCACAGGGGGCATCATCCAGCCCATTAATCCAGCCTATGACGGTGATGTAACTGAAAGGCT
ENV GENOMIC HERV-K TAN. ENV GENOMIC AC025420	(846) (843)	Балсавастслогосполосологослогослогослогослоственности состорателя с составляется с с с с с с с с с с с с с с с Балсавастскогосполосологослогослогослогослоственности с с с с с с с с с с с с с с с с с с
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(845)	EGACAAACTCAGTCGESTCCAAGTGCACAAGTGAGTCCAGCTGTTGATAGCGACAGAAAGTTE
ENV GENOMIC HERV-K8	(291)	GARAAAATTCAGCIGIGICTCAAAGIGCACAAGTGAGTCCAACTGTTGATAGCGACTTTAACAGAAAGTTT
ENV GENOMIC HERV-KI ENV HERV-K AF023261	(846) (701)	GIACMAACTTAGUTGIOTTTAAGTGCACAAGTGRGTTLARCAGIAGUAACTICAAAGUAACTICA
ENV GEN AL035086	(632)	GEALAAACTCAATCGTGTCCCCAGTGCACATGC
ENV GENOMIC AL035587	(870)	GGCCAGACTCACITCATGTTTACAGGCCCCATCCATCTGGCCCCATIAATCCGGCCTATGATAGTGATATTGAAAGGCC
ENV GENOMIC AC012068 ENV GENOMIC AF277315	(834) (843)	Gederadadeeadatatisticaaadadidearectorecegeeeadateeadateareatatisticaaaadaga Bideraactereetgangticaaagadeeteeteeggeeeaaaaaaaaaaaaaaaaaaaaa
ENV GENOMIC AF027650	(700)	
ENV GENOMIC AC078899 ENV GENOMIC HERV-KII	(845) (554)	GTACAAAACTCAATCATGTCCCAGTGCACAAGTCAGTCCAGCTGTTGATAGTGACTTAACAGAAGTT GSACAAACTCAGTCCAGTCCAAGTGCACAAGTGAGTCCAGCTGTTGATAGGAACTTAACAGAAAGTCT
ENV GENOMIC HERV-KII ENV GENOMIC AC008813	(865)	GACAAACTCAGTCAGTCCCAGTCCAGAAGTGAGTCCAGCTGTTAACAGTGACTTAACAAAAAAGTTTI
ENV GENOMIC AC012309	(845)	DEACAAACTCAGTCATGTCCCAGTGCACAAGTGAGTCCAACTGTTGATAGTGACTTAACAGAAAGTTT
ENV GENOMIC AL121932 ENV GENOMIC AD000090	(843) (859)	BEACAAACTCAATCATETCCCATECECAAGTGAGTCCAGATGTTGATAGGACATAACAGAAAGTCT GEACAAACTCAATCATETCCCAGTECACAAGTGAGTCCAACTSTTGATAGGACTTCATAGAAAGTTT
ENV GENOMIC AD500090 ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467 ENV GENOMIC AF235103	(610) (904)	GOCARACTCATCSTCACCCGGTGCACAAGTGAGTCCACCTETTCATAGCAACTTAACAGAAGT BJCCACACCCATCTATOTTCACACGTCCATCAGTCTGGCCCACTAATCCGCCCATGATAGTGATCTAACTAA
ENV GENOMIC AC026786	(780)	BCCAGGCCACTCATGTTCACAGBCTCCATCTCTCGGCCCACTAATCCGGCCTATGATGTTGATGTTAACTAAAAGGCI
ENV GENOMIC AC034203	(886)	GECCAGACCCACTCGGTTCAACAGGCTCTATCTGTCTGGCCCCCCTAATCCAGCCTATAATAGTGATTTAACTAAAAAACT
ENV GENOMIC AC018809 ENV GENOMIC HERV-K102 AF164610	(553) (606)	GGGANACTORGTONGTOCOGGGCACANGTONGTOCACTOTTGNAMGTACITANCAGAANGITU GGACANACTOGTOCONGTGCACANGTONGTOCACTOTTGATAGOACTTANCAGAANGITU GGACANACTOGTOCONGTGCACANGTONGTOCAGOTGTTGATAGOACTTANCAGAANGITU
ENV GENOMIC FRAG. AF260253	(1)	
CONSENSUS	(961)	GG CAAACTCA TC TGTCC AG GCACAAG AGTCCAGCTGTTGATAG GACTTAACAGAAAGT T
THE OPPONED HERE NON	(900)	1041 1120 GOACCAGGTETAFAGAAGGETTAGAATCACTCTGCCCAAGGAAATGGGGTGAAAAGEGAATTFCATCA
ENV GENOMIC HERV MDA ENV GENOMIC HERV-K TAN.	(914)	AGACAAACATAAGCATAAAAAATTGCAGTCTTTCTACCCTTGGGAATGGGGAGAAAAAGGAATCTCTACC
ENV GENOMIC AC025420	(911)	hgacaaacataagcataaaaaattgcagtctttctacccttgggaatgggaaaaaabgaatctctacc
ENV GENOMIC AP000776 ENV GENOMIC HERV-KB	(914) (291)	AGACAAACATAAAGCATAAAAATTGCAGTGTTIICIIACCCTTGGGAATGGGGAGAAAAAGGAACCICTACC
ENV GENOMIC HERV-KI	(914)	AGACIARECATAREGCATAREARATTOCHOTUTTTTTACCUTTREGRATIGAGGAGAREARAGGARTCTUTACC
ENV HERV-K AF023261 ENV GEN AL035086	(701) (700)	BGATAAAAAFAAGTATAAAAGGTTAGRGTCINCIAATACCCTTGGGAAAGGGGAAAAGGGAAATCINCATCA
ENV GEN AL035580 ENV GENOMIC AL035587	(950)	GSACCAGGTTTAFAGAAGGCTASAATCACCCTATCCCTCGAAATCGGGTGAAAAGGGAATTTCATCA
ENV GENOMIC AC012068	(914)	RCACCAGGTITIA TAGAAGGCITAAAATCACCITATCCATGCAAAATCGGGTGAAAAGAGGATTRCTTCA
ENV GENOMIC AF277315 ENV GENOMIC AF027650	(923) (700)	gand choost the Hondoost Hondan Carata Carocana Angeostican Ageonattic Arter
ENV GENOMIC AC078899	(913)	AGACGAACATIAAGCATIAAAAAATTACAGTETTTTCERCECTFUGGAAATGGGGAAGAAAAAGGAATCITCTAGTTTGGAT
ENV GENOMIC HERV-KII ENV GENOMIC AC008813	(622) (933)	agagaaacanaaggaataaaaattacagtiittiittii oo tiittiggaattgggaatgggaaaaaaaaaaaattgiittago aggaaacanaaggaagaaattgoggiittiittiittei oo tiitggaattgggaattgggaaaaaaaaaaaattiittii
ENV GENOMIC AC012309	(913)	
ENV GENOMIC AL121932	(911)	AGACANACANAAGCACGAAAAAATTATAGTCTTTCTACCATTGAGAATGAGGAGAAAAAAGAAAACTCTCTAGT
ENV GENOMIC AD000090 ENV GEN AL160008	(927) (647)	<u>AGACGAACATAAGCATAAAAAATTACAGTCITTICTA CICTUGGAAAAGGGGGGGAGAAAAGGAAICTCITACC</u>
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(676)	AGACCAGGTTTAENGAAGGCTAGAATCACCTATCCAFGAAATCOGGTGAAAAGGCGATTACATCA
ENV GENOMIC AF235103 ENV GENOMIC AC026786	(984)	AGACCAGGTITA
ENV GENOMIC AC034203	(966)	AGACCAAGTITATAGAAGGCTAGAATGACCCTATCCATGGAAATGGGGTGAAATGGGGATTTCATCA
ENV GENOMIC AC018809 ENV GENOMIC HERV-K102 AF164610	(621)	agacapacataagcataaaaaattacagtetttetta tecctgggaatggggagaaaaaaaagaatettta e
ENV GENOMIC HERV-RI02 AF184610 ENV GENOMIC FRAG. AF260253	(6/4)	Warlydd ylwyd ylwyd yn i goraet i i fryd fy'i ffrawr yn
CONSENSUS	(1041)	AGAC AA T A TA AA TTA A TC TCTA CC TGG AATGGGG GAAAA GGAAT TC C
		1121 1200
ENV GENOMIC HERV MDA ENV GENOMIC HERV-K TAN.	(967)	
	(984)	
ENV GENOMIC HERV-K TAN. ENV GENOMIC AC025420	(991)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(991)	
ENV GENOMIC AC025420	(991)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-K8 ENV GENOMIC HERV-K8 ENV HERV-K AF023261	(981) (984) (291) (984) (701)	
ENV GENOMIC AC025420 ENV GENOMIC APO00776 ENV GENOMIC HERV-KE ENV GENOMIC HERV-KI ENV HERV-K AF023261 ENV GEN AL035086	(981) (984) (291) (984) (701) (770)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-K8 ENV GENOMIC HERV-K8 ENV HERV-K AF023261	(981) (984) (291) (984) (701) (770) (1017) (981)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-KE ENV GENOMIC HERV-KI ENV HERV-K AF023261 ENV GEN AL035086 ENV GENOMIC AL035587 ENV GENOMIC AC012068 ENV GENOMIC AF277315	(981) (984) (291) (984) (701) (770) (1017) (981) (990)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-K8 ENV GENOMIC HERV-K1 ENV GENOMIC HERV-K AF023261 ENV GENOMIC AL035587 ENV GENOMIC AC012068	(981) (984) (291) (984) (701) (770) (1017) (981) (990) (700)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-KE ENV GENOMIC HERV-KI ENV GEN AL035086 ENV GENOMIC AL035587 ENV GENOMIC AL035587 ENV GENOMIC AC012068 ENV GENOMIC AF0276315 ENV GENOMIC AF027650 ENV GENOMIC AF027859 ENV GENOMIC AC078899 ENV GENOMIC HERV-KII	(981) (984) (291) (984) (701) (770) (1017) (981) (990) (700) (993) (692)	TTCAGCGAAGACTCCAAGATCGCCACCTCCGGATACCCTAACTCAGCATTTCCCGGTTCACCTTTCCTGTTCCCA
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AL035087 ENV GENOMIC AC012068 ENV GENOMIC AF0276315 ENV GENOMIC AC078899 ENV GENOMIC AC078899 ENV GENOMIC AC078813	(981) (984) (291) (984) (701) (770) (1017) (981) (993) (700) (700) (993) (692) (1003)	TTCAGCGAAGACTCCAAGATGGCAATCGCCACCTCGGATACCCTAACTCAGCATTTCCGGGTTCACCTTTCCTGTTCCCA
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-KE ENV GENOMIC HERV-KI ENV GEN AL035086 ENV GENOMIC AL035587 ENV GENOMIC AL035587 ENV GENOMIC AC012068 ENV GENOMIC AF0276315 ENV GENOMIC AF027650 ENV GENOMIC AF027859 ENV GENOMIC AC078899 ENV GENOMIC HERV-KII	(981) (984) (291) (984) (770) (1017) (981) (990) (700) (993) (692) (1003) (981)	TTCAGCGAAGACTCCAAGATCGCCACCTCGGATACCCTAACTCAGCATTTCCGGGTTCACCTTTCCTGTTCCCA
ENV GENOMIC AC025420 ENV GENOMIC AE000776 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AL035087 ENV GENOMIC AC012068 ENV GENOMIC AC012068 ENV GENOMIC AC027859 ENV GENOMIC AC078899 ENV GENOMIC AC008813 ENV GENOMIC AC008813 ENV GENOMIC AC012309 ENV GENOMIC AL2008812 ENV GENOMIC AL008813 ENV GENOMIC AL008813 ENV GENOMIC AL008813 ENV GENOMIC AL008813	(981) (984) (291) (984) (701) (701) (701) (981) (990) (700) (993) (692) (1003) (692) (1003) (983) (981) (981)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AC02068 ENV GENOMIC AF027630 ENV GENOMIC AF027650 ENV GENOMIC AF027650 ENV GENOMIC AC078899 ENV GENOMIC AC078899 ENV GENOMIC AC02813 ENV GENOMIC AC012309 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL121932	(981) (984) (291) (984) (701) (770) (1017) (981) (990) (700) (993) (592) (1003) (983) (981) (981) (997) (647)	TTCAGCGAAGACTCCAAGATCGCCAACCTCGGATACCCTAACTCAGCATTTCCOGGTTCACCTTTCCTGTTCCCCA
ENV GENOMIC AC025420 ENV GENOMIC AC02076 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AL035086 ENV GENOMIC AC012068 ENV GENOMIC AC012068 ENV GENOMIC AC078899 ENV GENOMIC AC078899 ENV GENOMIC AC008813 ENV GENOMIC AC021309 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC HEU32496 ENV GENOMIC HEU32496	(981) (984) (291) (984) (701) (1017) (981) (993) (692) (1003) (692) (1003) (692) (1003) (693) (981) (981) (981) (997) (647) (441) (683)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AC02068 ENV GENOMIC AF027650 ENV GENOMIC AF027650 ENV GENOMIC AF027650 ENV GENOMIC AC078899 ENV GENOMIC AC078899 ENV GENOMIC AC02801 ENV GENOMIC AC02309 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL120090 ENV GENOMIC AL160005 ENV GENOMIC AC011467 ENV GENOMIC AF235103	(981) (984) (291) (984) (770) (1017) (981) (993) (700) (993) (692) (1003) (983) (983) (981) (983) (981) (981) (647) (644) (663) (1051)	
ENV GENOMIC AC025420 ENV GENOMIC AC02076 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AL035086 ENV GENOMIC AC012068 ENV GENOMIC AC012068 ENV GENOMIC AC078899 ENV GENOMIC AC078899 ENV GENOMIC AC008813 ENV GENOMIC AC021309 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC HEU32496 ENV GENOMIC HEU32496	(981) (984) (291) (701) (701) (1017) (981) (990) (700) (993) (692) (1003) (981) (981) (981) (981) (981) (987) (441) (683) (1051) (927)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AL035087 ENV GENOMIC AC012068 ENV GENOMIC AF0277315 ENV GENOMIC AF027650 ENV GENOMIC AC078899 ENV GENOMIC AC078899 ENV GENOMIC AC012309 ENV GENOMIC AC012309 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL12032 ENV GENOMIC AL12032 ENV GENOMIC AC012467 ENV GENOMIC AC034203 ENV GENOMIC AC034203 ENV GENOMIC AC034203 ENV GENOMIC AC034203	(981) (984) (291) (984) (770) (1017) (981) (990) (700) (993) (692) (1003) (981) (993) (683) (647) (441) (663) (1051) (927) (1033) (691)	
ENV GENOMIC AC025420 ENV GENOMIC AC02076 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AL035086 ENV GENOMIC AL035087 ENV GENOMIC AC012068 ENV GENOMIC AC012068 ENV GENOMIC AC0267889 ENV GENOMIC AC008813 ENV GENOMIC AC008813 ENV GENOMIC AC012309 ENV GENOMIC AC012309 ENV GENOMIC AC012309 ENV GENOMIC AL121932 ENV GENOMIC AL012496 ENV GENOMIC AC026786 ENV GENOMIC AC026786 ENV GENOMIC AC026786 ENV GENOMIC AC026786 ENV GENOMIC AC024203 ENV GENOMIC AC034203 ENV GENOMIC AC034203 ENV GENOMIC AC01467 ENV GENOMIC AC034203	(981) (984) (291) (984) (770) (1017) (981) (990) (700) (993) (592) (1003) (981) (993) (693) (1003) (981) (997) (647) (441) (663) (1051) (927) (1033) (691) (744)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776 ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC HERV-KI ENV GENOMIC AL035086 ENV GENOMIC AL035087 ENV GENOMIC AC012068 ENV GENOMIC AF0277315 ENV GENOMIC AF027650 ENV GENOMIC AC078899 ENV GENOMIC AC078899 ENV GENOMIC AC012309 ENV GENOMIC AC012309 ENV GENOMIC AL121932 ENV GENOMIC AL121932 ENV GENOMIC AL12032 ENV GENOMIC AL12032 ENV GENOMIC AC012467 ENV GENOMIC AC034203 ENV GENOMIC AC034203 ENV GENOMIC AC034203 ENV GENOMIC AC034203	(981) (984) (291) (984) (770) (1017) (981) (990) (700) (993) (692) (1003) (981) (993) (683) (647) (441) (663) (1051) (927) (1033) (691)	

			1201 1280
	ENV GENOMIC HERV MDA	(967)	
	ENV GENOMIC HERV-K TAN.		
·	ENV GENOMIC AC025420		
	ENV GENOMIC AP000776		
	ENV GENOMIC HERV-K8		
	ENV GENOMIC HERV-KI		
	ENV HERV-K AF023261		
	ENV GEN AL035086		
	ENV GENOMIC AL035587		
	ENV GENOMIC AC012068 ENV GENOMIC AF277315		
	ENV GENOMIC AF277315		
	ENV GENOMIC AC078899		CCACCCCGACTAACGCACATGCCCACTAGGGCGTGTCACACTCAGAAGTGTGAAACTCAACCGATCCCGCCCCTACCCCG
	ENV GENOMIC HERV-KII		
	ENV GENOMIC AC008813		
	ENV GENOMIC AC012309	(983)	
	ENV GENOMIC AL121932		
	ENV GENOMIC AD000090		
	ENV GEN AL160008	(647)	
	ENV GENOMIC HEU32496	(441)	
	ENV GENOMIC AC011467	(683)	
	ENV GENOMIC AF235103	(1051)	
	ENV GENOMIC AC026786	(927)	
	ENV GENOMIC AC034203	(1033)	
	ENV GENOMIC AC018809	(691)	
ENV	GENOMIC HERV-K102 AF164610	(744)	
	ENV GENOMIC FRAG. AF260253	(1)	
	CONSENSUS	(1201)	
			1281 1360
	ENV GENOMIC HERV MDA		
	ENV GENOMIC HERV-K TAN.		
	ENV GENOMIC AC025420		
	ENV GENOMIC AP000776		
	ENV GENOMIC HERV-KB		
	ENV GENOMIC HERV-KI		
	ENV HERV-K AF023261		
	ENV GEN AL035086		
	ENV GENOMIC AL035587		
	ENV GENOMIC AC012068	(981)	
	ENV GENOMIC AF277315		
	ENV GENOMIC AF027650 ENV GENOMIC AC078899		
	ENV GENOMIC ACO78899 ENV GENOMIC HERV-KII		ACCACTCCTCACCCAGCATCCATAAAAGCGCGCTGCACCTTTCGCACAGCGTGACTTCCCCTGGCGGACCAGTGAACCTC
	ENV GENOMIC ACCO8813		
	ENV GENOMIC AC012309		
	ENV GENOMIC AL121932		
	ENV GENOMIC AD000090	(997)	
	ENV GEN AL160008		
	ENV GENOMIC HEU32496		
	ENV GENOMIC AC011467		
	ENV GENOMIC AF235103		
	ENV GENOMIC AC026786		
	ENV GENOMIC AC034203	(1033)	
	ENV GENOMIC AC018809	(691)	
ENV	GENOMIC HERV-K102 AF164610	(744)	
	ENV GENOMIC FRAG. AF260253	(1)	
	CONSENSUS	(1281)	
			1361 1440
	ENV GENOMIC HERV MDA	(967)	
	ENV GENOMIC HERV-K TAN.		
	ENV GENOMIC AC025420		
	ENV GENOMIC AP000776		
	ENV GENOMIC HERV-KB		***************************************
	ENV GENOMIC HERV-KI ENV HERV-K AF023261		
		(701)	
	ENV GEN AL035086		
	ENV GENOMIC AL035587 ENV GENOMIC AC012068		
	ENV GENOMIC ACCI2008		
	ENV GENOMIC AF277315 ENV GENOMIC AF027650		
	ENV GENOMIC AC078899		
	ENV GENOMIC HERV-KII		ACCGGAGAGCTCAATAAAGAAGATTTTTGCCCTCTTGTCTTGCCTCTTGGCCTTATTGATCCACGGTGCCTTTCCATTG
	ENV GENOMIC AC008813		
	ENV GENOMIC ACOUSSIS		
	ENV GENOMIC AL121932		
	ENV GENOMIC AD000090		
	ENV GEN AL16000B		******
	ENV GENOMIC HEU32496		
	ENV GENOMIC AC011467		
	ENV GENOMIC AF235103		
	ENV GENOMIC AC026786		
	ENV GENOMIC AC034203		
	ENV GENOMIC AC018809		
ENV	GENOMIC HERV-K102 AF164610		
	ENV GENOMIC FRAG. AF260253		
	CONSENSUS	(1361)	

FIGURE 6 CONTD ...

		1441
ENV GENOMIC HERV MDA	(967)	1441CGTTGACCAAAGTAGTCCTGTTACTGGTCCTGAACATCCAGAATTA _GGAAGCTTACTGGGCC
ENV GENOMIC HERV-K TAN.	(984)	CCAAGACCAAAAATAGTAAGTCCTGTTTCTGGTCCTGAACATCCAGAATTATGGAGBCTTACTGTGGCC
ENV GENOMIC AC025420	(981)	CAAJACCAAAATAATAATAATACCTGTTTTCTGGTCCTGAACATCCAGAATTATGCAGCCTTACTGTGGCC
ENV GENOMIC AP000776	(984)	
ENV GENOMIC HERV-K8	(291)	
ENV GENOMIC HERV-KI		CAASTCOARAAATTACTATISTISTISTISTISTISTISTACATCAGASTTAGAGECTTACTACTACTOR
ENV HERV-K AF023261	(701)	CCTCBACCRAAGTTAGETAGTCCTGTTACTGGFCCTAAGCCTCCGGAACTTTGGAAGCTTACTGTGGCC
ENV GEN AL035086 ENV GENOMIC AL035587	(770) (1017)	COTORCCAAGETAGETAGECCETATACTOGCCCTAACACCCCAAACTTAGGAAGCTTACTOGCCC
ENV GENOMIC AC012068	(981)	
ENV GENOMIC AF277315	(990)	CCCCCCACCAAACITAGTCCTCTTGTIGGTCCTCAACACCCCGGAATTATGGAADCTCACTGTCGCC
ENV GENOMIC AF027650	(700)	
ENV GENOMIC AC078899	(1313)	CCTTTCATACIC CAADACCAANAATAATAATAATTACTCOTGTTTTTTCGTCCTGAACATCCAGAAATTACAGAGCCTTACTGTACCG
ENV GENOMIC HERV-KII	(692)	CCAABACCAGAAATAATAAGTCCTGTTTCTGGTCCTGAACATCCAGAATTATGGAGGCTTTGGCC
ENV GENOMIC AC008813	(1003)	CCAAGAACAAAGATAATAAGTTCTGTTTTCTGGTCCTGAACATCCAAAAATTATAGAGTCTTACTGTGGCC
ENV GENOMIC AC012309 ENV GENOMIC AL121932	(983) (981)	CCAABACCAAAAAAAAATAAGTCCTGTTTCTGGTCCTGAACATCCAGAATTATGGAGCCTACTGTGGCC CCAABACCAAAAAAAAATAAGTCCTGTTTCTGGTCCTGAACATCCAGAATTATGGAGCCTAAAGTGGCC
ENV GENOMIC AL121932 ENV GENOMIC AD000090	(997)	CATGACCAABAATAATAAGGTCCTGTTTTCTGGTCCTGAACATCCAGAATGCTGGAGGCTTAGTTAAGC
ENV GENOMIC AD000000 ENV GEN AL160008	1647)	
ENV GENOMIC HEU32496	(441)	TTTTTCCCTCA TGAACATCCA AAATTATGGA GCCTTACTOTALCC CCCCCCCAAAGTTAGTTAGTCCTCTTTTGTTGCCCCGAACACCCAGAATTATGGAABCTCACTGTCGCC
ENV GENOMIC AC011467	(683)	TTTTTGGTCATGAACATCCAAAATTATGGACGCCTAACATGTACCC
ENV GENOMIC AF235103	(1051)	
ENV GENOMIC AC026786	(927)	CCOGACCAAAGTIAGTIAGTCCTGTTGTTGGTCCIGAACAOCCAAAATTATBGAABCTCACTGTGGTC
ENV GENOMIC AC034203		CCCCGACCAAAGTTAGPTAGTCCTOTTGTIGTTCCTGAACACCCAGAATTATGGAAGCTTACTGTBGCC
ENV GENOMIC AC018809		ccaabaccaaaaaataataastcctgtttCtggtcctgaacatccagaattatggaqgCttaCtgTagcc
ENV GENOMIC HERV-K102 AF164610 ENV GENOMIC FRAG. AF260253		
CONSENSUS	(1441)	CC GACCAAA TA T AGTCCTGTT CTGGTCCTGAACATCCAGAATTATGGA GCTTACTGTGGCC
ENV GENOMIC HERV MDA	(1031)	TCAC-ACCACATTAGAATTATTCTGGAAATCAAGCTATAGGAACAAGAGATCGTAAGTCATATTATACTATCAACCTAA
ENV GENOMIC HERV-K TAN.	(1053) (1050)	TCAC-ACCACATTAGAATTIGGICIGGAAATCAAACTITAGAAACAAGAGATGGIAAGCCATITTATACTAATGGACATGA TCAC-ACCACATTAGAATTIGGICIGGAAATCAAACTITAGAAACAAGAGATGGIAAGCCATITTATACTGTIGACCIAA
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(1050)	
ENV GENOMIC HERV-K8	(291)	
ENV GENOMIC HERV-KI	(1053)	TCAC-ACCACATTAGAATTTGGTTTGGAAATCAAAACTTTAGAAACAAGAGATCATAAGGCCATTTTATAGTAATGGACTTAA
ENV HERV-K AF023261	(701)	ICGC-BCCACATTAGAATTAGGTCTGGAAATTAAGCTATAAAAACAAGAGATGGTAAGCCATCITATACTATCGACCTAA
ENV GEN AL035086	(839)	itge-accaeattagaáttiggtetggaattaágetajaaaagaágatggtaggtaigeeatgitataetaetgaeetaa
ENV GENOMIC AL035587	(1086)	TCGC-ACTACATTAGAATTTGGTCIGGAAATCAAGCTACAGAAACAAGAAATCGTAAGCCATATTATACTATCAACCTAA
ENV GENOMIC AC012068	(1046)	TCGT-ACCACATTAGAATTTCGTCTCGGAAATCAAGTTATGGGAACAAGAATCATAAGCCATATTATACTATTAACCTAA
ENV GENOMIC AF277315 ENV GENOMIC AF027650	(1059) (700)	TCCC-ACCACATTAGAATTTCCTTTCGAAAAGAAAGAACTATAGAACAAGAGATCGTAAGCCATATTATACCTATTAACCTAA
ENV GENOMIC AF027650 ENV GENOMIC AC078899	(1393)	ICHT-ACCACATTAGAATTTGOTCTGGAAATCAAACTGTAGAAACAAGAGATGTAAGCCATTTTATACTATCGACGTAA
ENV GENOMIC HERV-KII	(757)	IGAC-ACCACATTAGAATTTOGTCTOGAAATCAAACTTTAGAAACAAGAGATCGTAAGCCATTTTATACTATCGACCTAA
ENV GENOMIC AC008813	(1072)	TCAT - ACCGCATTA AAATTTGGTCTGGAAATCAA GCTATAGAAACAAGAGATGTAAGCCATTTTATACTATCGACCTAA
ENV GENOMIC AC012309	(1052)	FCAT-ACCICATTAGAATTTGATCT3GAAATCAAGCTATAGAAACAGGAGATCATAAGCCATTTTATACTATCGACCTAA
ENV GENOMIC AL121932	(1050)	TCATTACCATATTAGAATTTGGTCTAGAAATCAAGCTATAAAAACAAGAGATCATAAGCGGTGTTACACTATCAACCTAA
ENV GENOMIC AD000090	(1066)	<u>TCAT-ÀCTACATTAGAATTIGGTCTGGAAATCÄÄÄCTITAGAAACGAGA – TCATAAGCCAUITTAFACTATCGACCTAB</u>
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496 ENV GENOMIC AC011467	(441) (727)	TCAT -ACCACATTAGGATTTGATCIOGAAATCAAACTTTAGBAACAAGAGATGATAAGCCATTITATACTATCGATCIAA
ENV GENOMIC AC011467 ENV GENOMIC AF235103	(1120)	TCAT -ACCACATTAGAATTGATCIGGAAATCAAAG TITAGAAACGAGAGATGATAGACAATAAACCATATTAGATCIATGATCIAATGATCIAA
ENV GENOMIC AC026786	(996)	TCCT-ACCACATTA ANATTTTGGTATGGAAATCAAGTTATGGAAACAAGAAATCACAAGCCATATTA CACTATTAACCTAA
ENV GENOMIC AC034203	(1102)	ICAT - ACCACACTAGAATTIGGTCTGAAAATCAAGTTATGGGAACAAAAAATCAATAAGCCATATTATACTCUTAACCTAA
ENV GENOMIC AC018809	(760)	TCAT - ACCACATTAGGATTIGOTCIGGAAATCAAACITTAGAAACAAGAGATGATCAGCCOMUTATACITATGGACCTAA
ENV GENOMIC HERV-K102 AF164610	(813)	
ENV GENOMIC FRAG. AF260253	(1)	
CONSENSUS	(1521)	TCA ACCACATTAGAATTTGGTCTGGAAATCAA CT TAGAAACAAGAGATC TAAGCCAT TTATACTATC ACCTAA
		16011680
ENV GENOMIC HERV MDA	(1110)	ATTCCAGTCTGACAATTCCTTTGCAAAATTGTGTAAAACTCCCTTATATTGCTAGTTGTAGGAAAAACATAGTTATTAAA
ENV GENOMIC HERV-K TAN.	(1132)	ATTOCAGTCTAACAGITCCTTTACAAAGTTCCGTAAAGCCCCCTTATAT-GCTAGTAGGAAATATAGTTATTAAA
ENV GENOMIC AC025420	(1129)	ATTCCAGTCTAACAGTTCCTTTACAAAGTFCCGTAAAGCCCCCCTTATAT-SCTAGTGTAGGAAATATAGFTATTAAA
ENV GENOMIC AP000776	(1132)	ATTCCAGTETAACAGITECTTTACAAAGTTECGTAAAGCCCCCTTATAT-GCTAGTTGTAGGAAATATAGTTATTAAA
ENV GENOMIC HERV-KB ENV GENOMIC HERV-KI	(291) (1132)	ATACCASACTAACACTACCITTACAAAGTTGCSTAAAGCCCCCTTATAT-OCTASITGTASSAAATBTASTTATTAAA
ENV GENOMIC HERV-KI ENV HERV-K AF023261	(1132)	
ENV HERV-R AF023201 ENV GEN AL035086	(918)	
ENV GENOMIC AL035587	(1165)	ATTOCATCT GACAATTCCTTTACAAAGTTGTGTAAAACCCCCTTATAT-GCTAGTTATAGAAAACHTAGTTATTAAA
ENV GENOMIC AC012068	(1125)	ATTALAATCTCACAAATTCCTTTCCAAAGTTAGGTAAAACCCCCTTATAC- GCTACTTOTAGGAAATATAGTTATTAAA
ENV GENOMIC AF277315	(1138)	ATTOCAATCIGACAATTOCTTTGCAAAGITGTSIAAGACCCCCITATAT-STTAGTTGTAGGAAAAATAGTTATTAAA
ENV GENOMIC AF027650	(700)	

ENV GENUMIC AC02542	(1123)	ATTCLAGICTAMCAGITCCTTTACAAAGITGCGTAAAGCCCCCTTATAT-GCTAGITGTAGGAAATATAGITATTAAA
ENV GENOMIC AP00077		ättelastetaala'sitelettia'laaasteisestaaaseleettatat- settetassaaata itastitataaa
ENV GENOMIC HERV-K	3 (291)	
ENV GENOMIC HERV-K	(1132)	ATTCCAGACTAACACITCCTTTACAAAGTTCCGTAAAGCCCCCTTATAT-GCTAGTGTAGGAAATBTAGTTATTAAA
ENV HERV-K AP02326	L (701)	
ENV GEN AL03508	5 (918)	ATTCCAGTCPCACAATTCCTTTGCAAANTTTGTGTAAAGCCCCCCTTATAI- CTAGTTOTAGGAAAAA- TAGTTATTAAC
ENV GENOMIC AL03558	/ (1165)	ATTCCAATCTGACAATTCCTTTACAAAGTTGTGTAAAACCCCCTTATAI-GCTAGTATAGAAAACATAGTTATTAAA
ENV GENOMIC AC01206	3 (1125)	ATTALAATCTGACAATTCCTTTGCAAAGTTAGGTAAAACCCCCCTTATAC-GCTAGTTGTAGGAAATATAGTTATTAAA
ENV GENOMIC AF27731	5 (1138)	ATTOCAATCTGBCAATTCCTTTGCAAAGTTGTGTAGACCCCCTTATAT-GTTAGTTGTAGGAAAAGTAGTTATTAAA
ENV GENOMIC AF02765) (700)	
ENV GENOMIC AC07889) (1472)	ATTCCASTCTAACAGCTCCTTTACAAAGTTCTGTAAAGCCCCCTTATAT-GCTAGTTGTAGGAAATATAGTTATTAAA
ENV GENOMIC HERV-KI	(836)	ATTCCAGTCTAACGGTTCCTTTACAAAGTTCCTTAAAAGCCCTCTTATAT-GCTAGTTGTAGGAAATATAGTTATTAAA
ENV GENOMIC AC00881	(1151)	ATTCCAGTCTAACAGTTCCTTTGCAAAGTTTTGTAAAGCCCCCTTATAT-GCTAGTTGTAGGAAATATAGTTATTAAA
ENV GENOMIC AC01230) (1131)	ATTCAAGTCTAACGCTTTCTTTACAAAGTTGTATAAAGCGCCCTTATAT-GCTAGTTGTAGGAAATATAGTTATTAAA
ENV GENOMIC AL12193	2 (1130)	ATTCCASTCTAAOGOTTCCTTCACAAAGTTGTBTAAAGCCCCCTTATAT-GCTAGGTAGGAAATATAGTTATTAAA
ENV GENOMIC AD00009) (1143)	ATTOCASTCTAACAGTTCCITTACAAAGCTGTGTAAAGTCCCTTTATAI-BCTAGTATAGGAAATATAGTTATTAAA
ENV GEN AL16000	9 (647)	
ENV GENOMIC HEU3249	5 (441)	
ENV GENOMIC AC01146	7 (806)	ATTCCAGTCTAAQGGITCCTTTACAAASTTTCATAAAGCCCCCCTTATAT-GCTAGTTGTAGGAAATATAGTTATTAAA
ENV GENOMIC AF23510	3 (1199)	AATCCAATCCGACAATTCCTTTCCAAAGTTATGTAAAACCCCCTCATAT SCTAGTTGTAGGAAACA-TAGTTATTAAA
ENV GENOMIC AC02678		ATTCCAACTGACAATTCCTTTCCAAAGTTGTGTGTAAAACCCCCTTATAT-GCTAGGTAGGAAACATAGTTATTAAA
ENV GENOMIC AC03420		ATTCCAATCTGACAATTCCTTTGCAAAGTTGTGTGTAAAACCCCCTTATAT-GCTAGGTAGGAAACATAGTTATTAAA
ENV GENOMIC AC01880		ATTCCAGTCTAACGCTTCCTTTACAAAGTTGTGTAAAGCCCCCTTATAT-CCTAGTTGTAGGAAATATAGTTATTAAA
ENV GENOMIC HERV-K102 AF16461		
ENV GENOMIC FRAG. AF26025		
CONSENSU		ATTCCAGTCT ACA TTCCTTT CAAAGTTG GTAAAGCCCCCTTATAT GCTAGTTGTAGGAAATA TAGTTATTAAA

EP 1 415 005 B1

ENV GENOMIC HERV MDA ENV GENOMIC HERV-K TAN.	(1190) (1209)	1081 COTEATCOCRARCCATA ATCTORGARRATTOIGEAATOTTACTTECATTEATTCACTTTAATTEGCCAECA CORACTCCCAEACCATA ACCTETEAAAATTGTAGAATTGCTACTTECATTEAATTGACTTTAATTEGCCAECA
ENV GENOMIC AC025420	(1206)	CCAGACACITCAAAACTATA ACCTGTGAAAATTGTAGATTGCTTACTTGCATTGATTCAACTTTTAATTGGCAACAC
ENV GENOMIC AP000776	(1209)	CCAGECTCCCAGACTATA ACCTGTGAAAATTISTAGATTICTACTTGCATTGATTCAACTTTTAATTRGCCAACAC
ENV GENOMIC HERV-K8	(291)	
ENV GENOMIC HERV-KI	(1209)	CCAGACTCCCBGACTATABCCTGTGAAAAATTGTAGATTGCTTACTTGCATTGATTCAACTTTTAATTGTCAACAA
ENV HERV-K AF023261 ENV GEN AL035086	(701) (995)	CCAGATTCCCAAACTATAACCTGTGAAAATTGTAGATTGTTTACTTGCATTGATTCAACTTTTAATTGGCAGCAG
ENV GENOMIC AL035587	(1242)	CCAGATICCCAAACTATAACCTGTGAAAACTGTAGATTGTTTACTTGCATTGATTTGACTITTAATTGGCAGCGC
ENV GENOMIC AC012068	(1202)	CCAGATICCCAAACTATA ACCTOTGAAAACIGTAGATTGTTTACTTGCATTGATTCAACTTTCGACTGGCAGCAT
ENV GENOMIC AF277315	(1215)	CCAGATUCCCAAACTATAACCTGRGAAAATTOFAGATTGTTTACTTGCATTGATTCGACTTCRATTGGCAGCAC
ENV GENOMIC AF027650	(700)	
ENV GENOMIC AC078899	(1549)	CCAGACGCCCAAACTATAACCISTGAAAATTGCAGATIGITIACTABCAITGATTCAACTITTAATTGGCAGCAC
ENV GENOMIC HERV-KII ENV GENOMIC AC008813	(913)	CCAGACTCCCAAACTATA ACCIGIGAAAATTGTAGATTGTTTACITGCATTGATTCAACTTTTAAPTGGCGCAC
ENV GENOMIC AC008813 ENV GENOMIC AC012309	(1228) (1208)	CCAGACTCCC CCAGACTCCCCAAACTATA TAACCTCTGAAAATTGCAGATTGCTTTACTTGCATTGATTCAACTTTCAATTGGCAGCAC
ENV GENOMIC ACUI2309 ENV GENOMIC AL121932	(1208)	CCAGACICCCAAACIAIN IARCINGAAAAITGCAGACIGTIIACIIGCAIIGAIICARUTICARIIGGCAGCAC CCAGACICCCAAACIAIN ACCIGIGAAAAITGCAGACIGTIIACIIGCAIGCAIGCAICGAIIGAIICARIICARIIGGCAGCAC
ENV GENOMIC AD000090	(1220)	CCAGACIUTICAAACTATAACCESTIGAAAACIIICAGATUGTTTACTIIGCATUGATUGATTCAACTITTTAATTGGCAACAC
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(883)	
ENV GENOMIC AF235103	(1276)	
ENV GENOMIC AC026786 ENV GENOMIC AC034203	(1152) (1258)	
ENV GENOMIC AC034203 ENV GENOMIC AC018809	(1258)	그는 이 전에 동안하는 것이 가지 못 한 것이 있는 것이 가지 않는 것이 있는 것이 없는 것이 있는 것이 있는 것이 있는 것이 없는 것이 있는 것이 없는 것이 있는 것이 없는 것이 없
V GENOMIC HERV-K102 AF164610	(910)	
ENV GENOMIC FRAG. AF260253	(1)	
CONSENSUS	(1681)	CCAGA TCCCAAACTATA ACCTGTGAAAAATTGTAGATTGTTTACTTGCATTGATTCAACTTTTAATTGGCAGCAC
		1761 1
ENV GENOMIC HERV MDA	(1268)	ATTCTACTAGGAAGAGCAAGAGAGGCTTTGTGGATCCTTGTGCCATCGACCATGGAGGCTTCGCTATCCAT
ENV GENOMIC HERV-K TAN.	(1287)	ATTCTGCTG9T0AGAGCAAGAGAGGGTGTGGATCCCTGTGTCCATGGACGACGTGGGAGGCCCCCCATCCA
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(1284) (1287)	ATTC/GCTGJTGAGAGGCAAGAGAGGGGTGTGGGATCCCTGTGTCCATGGACCGACC
ENV GENOMIC HERV-K8	(291)	RTICIGATERADA ANGRADES A TOUS TOUS TOUS AND A TOUS
ENV GENOMIC HERV-KI	(1287)	ATTY: TIG: TAGTISAGAGCAAGAGAGGGGGTGTGGATCCCTGTGTCCATGGACCGTAGGAGGCCTCGCCATGGG
ENV HERV-K AF023261	(701)	
ENV GEN AL035086	(1073)	
ENV GENOMIC AL035587	(1320)	
ENV GENOMIC AC012068	(1280)	ATTCIGTUAGTCAGGGCAAGAGAGAGAGOTATUGGGATCCCTGTGTCCATGGACTGACCGTGGGAGGGTTCTCCATCTGG
ENV GENOMIC AF277315	(1293)	
ENV GENOMIC AF027650 ENV GENOMIC AC078899	(700) (1627)	ATTCTGCTAGTGAGAGCAAGAGAAGGTGTGTGGATCCCTGTGTCCATGGACTGACT
ENV GENOMIC ACU78899 ENV GENOMIC HERV-KII	(991)	
ENV GENOMIC AC008813	(1238)	
ENV GENOMIC AC012309	(1288)	ATTCHOCHAGTGAGAGCAAGGGAAGGGGTGTGGGATCCCTGTGTCCATGGACGGAGGCCTGGGAGGCCTGGCCATCCAT
ENV GENOMIC AL121932	(1285)	ATTCTSCTASTGAGAGGAAGAGAAGGGTGTGGATCCCGTGTCCATGCACCGTGGGAGGCCTCGCCATCTGT
ENV GENOMIC AD000090	(1298)	
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496 ENV GENOMIC AC011467	(441)	
ENV GENOMIC ACOII487 ENV GENOMIC AF235103	(961) (1354)	
ENV GENOMIC AC026786	(1230)	
ENV GENOMIC AC034203		ATTCIG TIAGTA AGGCAAGAAAAGTTIGTGGATCCCTGTGTCCATGGACCGACCGTGGGAAGCTT
ENV GENOMIC AC018809	(992)	ATTCTGCTGGTGAGAGCAAGAGAGGGGGTGTGGATCCCTGTGTCCATGGACOGACGGTGGGAGGCCICGCCATCCG3
W GENOMIC HERV-K102 AF164610	(1047)	ATTERSTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
ENV GENOMIC FRAG. AF260253 CONSENSUS	(1)	
CONSENSUS	(1/81)	ATTERET GLAANGLANGNAN OF GLGEGATEEEIGIGEEEAIGGAEGAACGIGGAAGGE TE EEAITE T
ENV GENOMIC HERV MDA	(1348)	1841 TATTTTAACGGAAGTATTAAAAGGAATTCTAACTAGATCCAAAAGATTCATTTTTACTTTGATCGCAGTGATTATGG
ENV GENOMIC HERV-K TAN.	(1367)	
ENV GENOMIC AC025420	(1364)	INTITI'GACTGAAGTATTAAAAGGTGTTTTAAATGATCCAAAAGATTCATTITTACTTTAATTGCAGTGATTATGG
ENV GENOMIC AP000776	(1367)	
ENV GENOMIC HERV-K8	(291)	
		TATTTTGACTGAAGTATTAAAAAGGTGTTTTAAATAGATCCAAAAGATTCATTTTAACTTTAATTGCAGTGATTATAG
ENV HERV-K AF023261	(701)	
ENV GEN ALUSSU86	(1200)	Inteltance in the second secon
ENV GENOMIC AC012068	(1360)	TATCITIACGIALAGAAGTATTCAACTACTACTACTACTACTACTACTACTACTACTA
ENV GENOMIC AF277315	(1373)	TATITTAATGGAAGTATTAAAAGGAGTTCPAACTAGATCCAAAAGATTCATTTTTACTTTAATUGCAGTCATTANGG
ENV GENOMIC AF027650	(700)	
ENV GENOMIC AC078899	(1707)	IATTITCACTINGGTATTAAAAGACAFTTTAAATAGATCCAAAAGATTCATTTTTACTTTAATTGCAGTGATTATGC
		TATTITGACTBAAGTATTAAAAGACATITTAAATAGATCCAAAAAGATTCATTTTTACCTTAATTGCAGTGATTATGC
ENV GENOMIC AC008813	(1238)	
ENV GENOMIC AC012309	(1368)	TATTITGACTGAAGTATTAAAAGTCATTTIAAATAGATCCAAAAGATTCATTTTACTTTAATTGCAGTGATTATGG TATTITGACTGAAGTATTAAAAGGCATTTIAAAATAGATCCAAAAGATTCGTTTTTACTTTAATTGCAGTG-TTATCC
ENV GENOMIC AD121932	(1378)	TRITITGACIGAAGTATAAAAGGCATGITAAAATAGATCCAAAAGATICGTITTTACITTAATAGCAGTGATTAICG TATTTTGACTGAAGTATTAAAAGGCTITTTAATAGATCCAAAAGATTCATTTTTACTTTAATAGCAGTGATTAICG
ENV GENOMIC HEU32496	(441)	
		TATITTGACTGAAGTATTAAAAGGCATTITTAAATAGATCCAAAAGATTCATTTTTACTTTAATTGCAGTGATTATGC
ENV GENOMIC AF235103	(1434)	TATCITIANCAGAAGTATTAAAAGGAGTTCTAACTAGATCTAAAAGATTCATTTTTACTTTAATTGCAGTGATTATGG
ENV GENOMIC AC026786	(1310)	TATCTTAACAGRAGTATTAAAAGGAGTTCTAACTAAATCTAAAAGATTCATTTTACTTTGATTGCAGTGATTATGG
ENV GENOMIC AC034203	(1403)	Татттсастрая татталалссолтттралагасатссалалсяттсатттасттталттасаотссттатса
FNV GENOMIC ACOLOGOO	(1072)	TATTTGACTGAA-TATTAAAAAGCAATTTAAAAAAAAAAA
W GENOMIC WEBU CON ACCIECCY		
IV GENOMIC HERV-KL02 AF164610 ENV GENOMIC FRAG. AF260253	(1127)	TATTITCACTERAGIATTAAAAGGICTTTTAAATAGATCCAAAAGATTCATTITTACTITAATTSCAQTGATTATGG

		2000
ENV GENOMIC HERV MDA	(1428)	
ENV GENOMIC HERV-K TAN.	(1447)	TCATTGCAGTCACAGCTAGIGCTGGGGGGGGGGGGGGGGG
ENV GENOMIC AC025420	(1444)	TAATTGCAGTCACAGCTA TGBCTGCTGTAGCAGGAGTTGCATTGCACTCTITITGTTCAGTCAGTCAGTAAACTTITGTTAATGAT
ENV GENOMIC AP000776	(1447)	TAATTGCAGTCACAGCTACGGCTGCTGTAGCAGGAGTTGCACTCTCTCT
ENV GENOMIC HERV-K8	(291)	RANTIGCAGICACAGCTACOBCTGCTGTABCABASCTIGCATTGCACTCTGTGTGAGTCAGTAASCTITIGTTAATGAT
ENV GENOMIC HERV-KI ENV HERV-K AP023261	(1447) (701)	<u>EARTIGCAUTCAUSCIACOCTUCIUTARADAUSTICATIGIACICUCUUTCAUTAUAUAUAUAUAU</u>
ENV GEN AL035086	(1233)	TARGTGCAGTCACAGTTATGGCGCTGTGGCAGGAGTTGCATTGCACTCTTCAGAGGGTACACTTFGTTAACAAT
ENV GENOMIC AL035587	(1475)	TTATTGCAGTCACAGCTACTGCTGCAGCTGCTGCAATTGCTTTACACTCCTCTGTTCAAACTGCAGAATATGTAAAIGAT
ENV GENOMIC AC012068	(1440)	htattgcautcacaderactisciscisciscisciationalitectiviacacticitoriotiticaaaliticacaalahadigaa taan
ENV GENOMIC AF277315	(1453)	ITGTTGCAGFCACAGCTACTGCTGGCTGGTGGAATTGCTTTACACTCCTCTGTTCAAATGGTBABATATGTABATAAT
ENV GENOMIC AF027650 ENV GENOMIC AC078899	(700) (1786)	
ENV GENOMIC ACU78899 ENV GENOMIC HERV-KII	(1151)	TG
ENV GENOMIC AC008813	(1238)	
ENV GENOMIC AC012309	(1448)	TANT COLOR OF TRACE OF THE TANDER OF THE TANDER OF THE TARGET THE TARGET TARGET TARGET TARGET TARGET TARGET TAR
ENV GENOMIC AL121932	(1444)	TAAGTGCAGTCACAGCTACGCCGCTGTGGCAGGAGTTGCATTGCACTGTTGCTGTTCAGACAGTAAACLTTGGTTAACGAT
ENV GENOMIC AD000090	(1458)	TAATTGCAGTCACAGGTATGSCIGCTGTGGCAGGAGITGCATTGCACTCIGTTCAGTCAGTAAACTTTGTTAAIGAT
ENV GEN AL160008 ENV GENOMIC HEU32496	(647) (441)	
ENV GENOMIC AC011467	(1121)	TANTTGCASTCACAGCTACGCTGCTGTGCAFGAGTTGCACTCTCTCTCAGTCAGTCAGTCAGTCAAGTAAAGTTGTTAATGAT
ENV GENOMIC AF235103	(1514)	ITATTOCAGTCACAGCIGCT9CIAGGCCT9CTGTAATTGCTTTACCCTCCTCTTTCACACTGCAGAATATGTGAATAAT
ENV GENOMIC AC026786	(1390)	TATTSCAUTCACAGCIACTSCTUCECTUCEGAATTSCAUTACACIGCTCTUTCAAACTACAAAATAbuGGAATAAT
ENV GENOMIC AC034203	(1403)	
ENV GENOMIC AC018809	(1151)	TAATTGCAGTCACAGCTACGGCTGCTGTGGCAGGAGTTGCGCTGCACTCTTCTGTTCAGTCAG
ENV GENOMIC HERV-K102 AF164610	(1207)	iarthcagteragetracegetetetageabgacitgeattgeactetetetageteabetaacittettaatgat
ENV GENOMIC FRAG. AF260253 CONSENSUS		T ATTECAGTCACAGCTAC GETEC G GC GGA TTEC TT CACTC TETETTCA C G A A T TET AAT AT
CONSENDOS	(1)21)	
		2001 2080
ENV GENOMIC HERV MDA	(1508)	
ENV GENOMIC HERV-K TAN.	(1527)	IggCraaaagaa - TTCCTCAAAATTGTGGAATTCTCAGATCCAA - ATAGATCAAAATTGGCAAACCAAATTAATGATCATC IggCraaaaaaa - TTCTACAACATTGTGGAATTCACA - ATQTAGTATTAATCAAAAATTGGCAAAACCAAATTAATGATCTT
ENV GENOMIC AC025420	(1524)	TGGCANAAAAA-ITTCTACAAGATTGTGGAATTCACA-ATGTAGTATTGATCAAAAATTGGCAAATCAAAT
ENV GENOMIC AP000776	(1527)	
ENV GENOMIC HERV-K8	(291)	
ENV GENOMIC HERV-KI	(1527)	
ENV HERV-K AF023261	(701)	NGCCAAAAAAAATTCTACAAGATTGTGGAATTCACA -ATDTGGTATTGATCAAAAATTGGCAAATTAGATTAGATTA
ENV GEN AL035086 ENV GENOMIC AL035587	(1313) (1555)	ICCCGAAAGAA-ITCCTCAAAATTGTCGAATTCTCAGACCAA-ACAATCAAAAATTGCCAAA GAAATTAATGATCTT
ENV GENOMIC AD033367	(1520)	
ENV GENOMIC AP277315	(1533)	
ENV GENOMIC AF027650	(700)	
ENV GENOMIC AC078899	(1788)	BCNAAAGAA- ITCTACAAGATTGTGGAATTGACA-ATCTGGTATTGATCAAAATTABCAAATCAAATTAATGATTTT
ENV GENOMIC HERV-KII	(1231)	Ingglaaaa cab - It ctacaa catter caca - a titagtait cataaaaattigg caaatlaaattigatiett
ENV GENOMIC AC008813 ENV GENOMIC AC012309	(1238)	IGGCAANAGAA-FTCTACAAGATGETGGAATTCACA-ATTITGGTAITGATCAAAAATTCGCAAATCAAATTAATGATCTA
ENV GENOMIC AC012309 ENV GENOMIC AL121932	(1528) (1524)	IRSCAARIGAGA-JITUTALAAGATIGIRGAATXUALA-MILIGOIALIGATXEAAARTI IRSCAARITAARITAA ISA ISI' 1720/193 Archinego
ENV GENOMIC AD000090	(1538)	
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(1201)	TOGCAAAAGAA-TTCCACAAAATTUTUGAATTCACA-ATLTGGTATTCAAAAATTGGCAAATCAAAATTAATUATCTT
ENV GENOMIC AF235103 ENV GENOMIC AC026786	(1594)	PGGCAAAAGAA-DTCCTCAAAATTGTGGAATTCTCAGACTCAA-ATAGATCAAAAATTGGCAAATCAAAT
ENV GENOMIC AC028788 ENV GENOMIC AC034203	(1470)	TRECAMACIA - I TOTIZACO UN TREATILITARE COMPANYA I MELANTIANA MATLANA MATLA
ENV GENOMIC AC018809	(1231)	IGGCABAAAAA-IIGUTACAAGATHGTGGAATTCACA-ATCIGGTATIIGATCAAAAATTGGCAAATCAAAATTAATGATTTT
ENV GENOMIC HERV-K102 AF164610		TGCAAAAGAA-TTTTACAACATTGTGGAATTCACA-ATCTAGTACTGAAAAAATTGGCAAATCAAAT
ENV GENOMIC FRAG. AF260253	(1)	
CONSENSUS	(2001)	TGGCAAAA AA TTC CAA ATTGTGGAATTC CA A C AT GATCAAAAATTGGCAAATCAAAT
		2081 2160
ENV GENOMIC HERV MDA	(1586)	
ENV GENOMIC HERV-K TAN.	(1605)	
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(1602) (1605)	
ENV GENOMIC HERV-KS	(291)	
		AAAAAACUSTCAIITIGGATAGGAGACAGALICAIGAGCTIAGAACATOGIIICACGIIRGHAGTGACIGGAAAAACUST
ENV HERV-K AF023261		
ENV GEN AL035086	(1392)	AGACAAAATOTCATTTTGGATGGGAGATAAGCTCGTGAGCTIGGAACATCATTTTCAGTTACAGTGTGACTGGAATACBTC
ENV GENOMIC AL035587	(1633)	AGACAAACTGTCATTTGGCTGGGAAATAGGCTCATGAGCTTGGAATATCTTTTTCAGTIACAGTGTGACTGGAATATGTC
		AGACAAACTGTCATTTTGGATGGGAGATAGGCTCATAAGTITTGGAATATCTTTTTTCAGTTACAGTGTGACTGGAATATGTC
ENV GENOMIC AF277315 ENV GENOMIC AF027650	(1011)	AGACAAACTGTCATTTGGATGCGAGATAGGCTCATGAGCTIGAAATATCTTTTTCAGTTACAGUGTGACTGGAATACGTC
ENV GENOMIC AC078899	(1864)	AGACAAACTGTCATTTGGATGGGAGATAGGCTTATGAGCTTAGAACATCATTICCAGTTACAGIGTGAATACTTC
ENV GENOMIC HERV-KII	(1309)	AGACAAACTGTCATTTGGATGGGAGA'CAGALTCATGAGCTTAGAACATTGTTICCAGTTACAGTGTGACTGGAATACGTC
ENV GENOMIC AC008813	(1238)	
ENV GENOMIC AC012309	(1606)	AGACTOTCATTIGGATGGGAGATAGGCTCATGAGCTIAGAACATCATTICCAGTIACAGINTGACTGGAATACCTC
ENV GENOMIC AD000090	(1616)	AGACAAACTOTCATITGGAIGGGAGACAGGCPCATGAGCTTAGAACATGGTTICCASTTACAGPTGACTGGAAIACGTC
ENV GEN ALIGOUDS ENV GENOMIC HEU32496		
		AGACAAACTGTCATTTGGGTGGGAGACAGACTCATGAGCTTAGAACATCATTICCAGTTAGATGTGACTGGAATATCTC
ENV GENOMIC AF235103	(1672)	AGACAAACTGTCATTTGGATGGGAGATAGGCTCATAAGTTTGGAATATCTTTTCAGTTACTGTGTGACTACAATACATC
ENV GENOMIC AC026786	(1548)	AGACAAACTGTCATTTGGATGGGAGATAGGCTCATGAGCCTGGABTATCTTTTTCAGTTAGTGTGACTGGACTAGTC
ENV GENOMIC AC034203	(1403)	AGACAAACTGTCATTTGGATGGGAGACAGGCTCATGAGCTTAGAACATGGTTICCAGTTACAATGTGACTGGAATATGTC
ENV GENOMIC AC018809	(1309)	AGACAAACTGTCATTTGGATGGGAGACAGGCTCATGAGCTTAGAACATGGTTTCCAGTTACAATGTGACTGGAATATGTC
ENV GENOMIC HERV-K102 AF164610 ENV GENOMIC FRAG. AF260253	(1365)	AGACAAACIGICATITGGAGGGGAGACAGACICATGAGCITAGAACATGGITTCCAGTIACAANATGACIGGAAIACITC
		AGACAAACTUTCATTTCGATCGGAGA AG CTCATGAGCTT GAA ATC TTT CAGTTACA TGTGACTUGGAATACGTC

		2161 224 AGATTTTTGTGTTACACCACAAGCCTATAATGAGTCTGAGCATCACTGGGACATGGTTAGATGCCATCIGCAAGGAGA
ENV GENOMIC HERV MDA ENV GENOMIC HERV-K TAN.	(1664) (1685)	AGATTTTTGTATTAGACGACAAGCCIATAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAGA AGATTTTTGTATTAGACGCCAAATTTATAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAGA
ENV GENOMIC AC025420	(1682)	AGATTITITISTATTACACCCCARATTIATAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAGA
ENV GENOMIC AP000776	(1685)	BGATTTITISTATTACACC CEARATTTATEAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAGA
ENV GENOMIC HERV-KS	(291) (1685)	AGATTIITISTATTACACCOCAAATTIATAATGAGTCTGAGCATCA-TAGGACATAGTTAGACGCCATCTACAAGGAAGA
ENV GENOMIC HERV-KI ENV HERV-K AF023261	(1665)	
ENV GEN AL035086	(1472)	AGATTTTTTGTATTACACCCCAAGCCFGFAATGAOTCTGAGCCTCACTGGGACATGGTTAGATGCCATCTACAAGGAAGA
ENV GENOMIC AL035587	(1713)	AGATTTTTTTTTTTACACCOCAAGCOTATAATGAGTCTGAGCATCACTGGGACATGGTTAGATGCCATCTACAGGGAAGA
ENV GENOMIC AC012068	(1678)	AGATITTTTGTATTACACITCAAGCCTATAATGAAICCTAAACATCACTGGGACATGGTTAGATGCCATCIACAAGGAAGA
ENV GENOMIC AF277315	(1691) (700)	AGATITTTIGTATTACACCCCGAGCCCTACAATGAGTCTGAGCATCACCGGGCCATGGTTAGATGCCATCTACAAGGAAGA
ENV GENOMIC AF027650 ENV GENOMIC AC078899	(1944)	AGATTTTTGTATTACACCCCSAGTTTATAATGAGTCTCAGCATCACTGGGACATCGTTAGACCCCCATCTACAAGGAAAA
ENV GENOMIC HERV-KII	(1389)	AGATTTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAGA
ENV GENOMIC AC008813	(1238)	
ENV GENOMIC AC012309	(1682)	AGATITTINITATTACACCTGGAGTITATAATGAGTCTGAGCATCACTGGGACATGGTTAGATGCCGICTACAGGAAGA
ENV GENOMIC AL121932 ENV GENOMIC AD000090	(1538) (1696)	AGAITTITISTATTACACCCCAAGTITATAATGAGTCTISAGCATGACTGGGACATGGTTAGACGCCATCTACAGGGAAGA
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(1359)	agattittistattacaccocaaattitataatgagtctgagcatcactggggacatggttagatgccatctacagggagga
ENV GENOMIC AF235103	(1752)	agattitingtattacaccitcaacccitataatgaatctgaacatcactgggacatggttagatgccatctacaaggaaga Agattitingtattacaccitcggccitataatgaatctgaacatcactgggacatggttaaacgcccittacaaggaaga
ENV GENOMIC AC026786 ENV GENOMIC AC034203	(1626) (1403)	
ENV GENOMIC AC018809	(1389)	AGATTTTTTGTATTACACCCCAAGT-TGTAATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACACGCAAGA
ENV GENOMIC HERV-K102 AF164610	(1445)	AGATTTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCACTGGGACATGGTTAGAGGCCATCTACACGGAAGA
ENV GENOMIC FRAG. AF260253		CIGGTAAGAGCTAGTCTACAGGAGAGA
CONSENSUS	(2161)	AGATTTTTTGTATTACACC CAA TATAATGAGTCTGAGCATCACTGGGACATGGTTAGA GCCATCTACA GGAAGA
ENV GENOMIC HERV MDA	(1744)	2241 233 Алдатаатст-тастятадасаттасаалааттааладаа, тосссадссаадалалатеттто ассатсалалдосса.
ENV GENOMIC HERV-K TAN,	(1765)	AAGATAATCTCACTTTAGACATTTCCAAATTAAAAGAACAAATTTTCGAAGCATCAAAAGCCCAT
ENV GENOMIC AC025420	(1762)	алдаталтстісастттадасатттісалатталалдараБалаттітісаласатсалалдасатсалалдосат алдаталтстісастттадасатттісалатталалдалаСалаттттісаласатсалалдоссат алдаталтстісастттісалаттасалатталалдалаСалатттісаласатсалалдоссат
ENV GENOMIC AP000776	(1765)	BAGAIAATUICACITIIGACATIICCBBAITAAAAGAACAAAIITIICCAAGCAICAABAAGCCCAI
ENV GENOMIC HERV-KS ENV GENOMIC HERV-KI	(291) (1765)	BAGATABACECECTTTAGACATTTCCAAATTABAAGAA
ENV GENOMIC HERV-KI	(701)	
ENV GEN AL035086	(1552)	AAGATAATCTITACTICAGACATTICAAAATTAAAAGAA
ENV GENOMIC AL035587	(1793)	AAGATAATCTTACTTTAGATATTTCAAAATTAAAATAACAAAATTTTGAGGCATCGAAAGCCCAT
ENV GENOMIC AC012068	(1758)	AGATAATCTTACTTTACATATTTCAAAATIGAAAAAA
ENV GENOMIC AF277315 ENV GENOMIC AF027650	(1771) (700)	ABGATABATCHICACTITTAGACATTTICCAABATTABABAGAA AAGATABATCHICACTITTAGACATTTICCAABATTABABAGAA AAGATABATCHICACTITTAGACATTTICCAABATTABABATABA AAGATABATCHICACTITTAGATATTTICAABATTABABATABA AAGATABATCHICACTITTAGATATTICAABATTABABAGA AAGATABATCHICACTITTAGACATTTICCAABATTABABAGAA AAGATABATCHICACTITTAGACATTTICCAABATTABABAGAA ABGATABATCHICACTITTAGACATTTICCAABATTABABAGAA ABGATABATCHICACTITTAGACATTTICCABATTABABAGAA ABGATABATCHICACCTITTAGACATTTICCABATTABABAGAA ABGATABATCHICACCTITTAGACATTTICCABATTABABAGAA ABGATABATCHICACCTITTAGACATTTICCABATTABABAGAA ABGATABATCHICACCTITTAGACATTTICCABATTABABAGAA ABGATABATCHICACCTITTAGACATTTICCABATTABABAGAA ABGATABATCHICACCTITTAGACATTTICCABATTABABAGAA ABGATABATCHICACCTITTAGACATTTICCABATTABABAGAA
ENV GENOMIC AC078899	(2024)	ARGATAATCTCACTTTAGACATTTCCAAATTAAAAGAACAAATTTTTGAAGCATCAAAAGCCCA
ENV GENOMIC HERV-KII	(1469)	BAGATAATET CACTITAGACATTYCCAAATTAAAATAACAAATTTCGBAGCATCAAAAAGCCCA
ENV GENOMIC AC008813	(1238)	
ENV GENOMIC AC012309	(1762)	BAGATAATCTCACTTTAGACATTTCCAAATTAAAGAA
ENV GENOMIC AL121932 ENV GENOMIC AD000090	(1538) (1776)	
ENV GENOMIC AD000090 ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(1439)	AAGATAATCTCACTTTAGACATTTCCAAATTAAAAGAACAAATTTTCAAAACATCAAAAAGCCCAT
ENV GENOMIC AF235103	(1832)	
ENV GENOMIC AC026786 ENV GENOMIC AC034203	(1706) (1403)	
ENV GENOMIC AC018809	(1468)	AAGATAATCTCACTTTAGACATTTCCAAATTAAAAGAACAAATTTTCAAAACATCAAAAGCCCCA
ENV GENOMIC HERV-K102 AF164610	(1525)	алдаталтот саститада салтисалалталалда
ENV GENOMIC FRAG. AF260253	(29)	ATATCICACITI-AGACATICAATAAGTACAATATCICAACCATC-AAAGCCCA
CONSENSUS	(2241)	AAGATAATCT ACTTTAGACATTIC AAATTAAAAGAA CAAAATTIT AA CATCAAAAAGCCCAM
ENV GENOMIC HERV MDA	(1824)	2321 TAAATTTIGGTGCCACGAACGGAGAGAATCGTGAAAGCTGCTGATAGCCTCACAAATCTTAAGCCAGTCACTTGGGTTA
ENV GENOMIC HERV FLA	(1831)	
ENV GENOMIC AC025420	(1828)	
ENV GENOMIC AP000776	(1831)	ÎBAATUNGINGCAGGĂACNGAGGAĂŃĪĬĊAGGĂGĪNGENGATĢGCCNCĠĊĂĂĂŦĊĪŬĂĂĊĊĊŊŢĊĂĊŢŊŔĠŢŢŢĂ
ENV GENOMIC HERV-K8 ENV GENOMIC HERV-KI	(291)	FAAATTIAGTGCCAGGAACTGAGGCAATTECAGGAGTFGCTGATGGCCTCGCAAATCTTAACCCTGTCACTT
ENV GENOMIC HERV-KI ENV HERV-K AF023261	(701)	
ENV GEN AL035086	(1618)	TAAATTTUGTECCAGGAACTGAAGCAAETACGGAAGCTACTGATEGCCTCGCAAATCTTAACCCAGTCACTTSGGTTA
ENV GENOMIC AL035587	(1859)	TAAATTTGOGAACTGAGGCAATCGTGAAAGCTACTGATGGCCTCACAAATCTTAACACTGTCACTGGTTAA
		TAAATCIGGTGCCAGAAACTGAGGCAACGGTGAAAGCTGTTGATAGCCTCACAAATCITAACCCTGTCACTTGGGTTA
		TAAATCINEGTSICAGAAACTGAGGCAATCITGAAAGCINGCTGATGGCCTCACAAATCTTAATGCCGTCACITGGGTTA
ENV GENOMIC AF027650 ENV GENOMIC AC078899		TAAATTTOGTACCOBGAACTGAGGCAATCATGGGAGTTGCTGACTGCCTCTCAAATCTTAACCCIGTCACTTGGGTTA
		TANATITICATCCCAGGAACTCAGGCAATTECAGGAGTTCCTGATCGCCTCCCAAATCTTAACCCTCTCACTTGGGTTA
ENV GENOMIC AC008813	(1238)	
ENV GENOMIC AC012309	(1826)	TAAATTTISTACCAGGAACTGAGGCAATCATCG-AGTINGCTGATGGTCTCACAAATCITAACCCTGTCACTTCGGTTA
ENV GENOMIC AL121932	(1538)	
ENV GENOMIC AD000090 ENV GEN AL160008	(1842) (EAT)	TAAATTTUGTGECAGGAACTGAG <u>CCAAT</u> TGCAGGAGUIGCTGATGGCCTCGTAAATCTTAACCCTGTCACTUGGGTTA
ENV GEN ALIGO008 ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(1505)	TAAATTTGGTGCOGGGAACTGAGGCAATCGCAGGAGTTOCTGCTGCCCCTCTCAAATCTTAACCCTGTCACTTGGGTTA
ENV GENOMIC AP235103	(1898)	TARATTIGGIGCCAGAAACTGGCAAGAAGCAAAAGCIGIGGAAAAAGCIGAACGAAACGAAGACGAAGACGAAGACGAAGACGAA
ENV GENOMIC AC026786	(1772)	TAAATCTGGTGCCAGAAACTGAGACAATGGTAAAAGCTGTTGATAGCCTCACAAATCTTAACCCTGTCACITGGGTTA
ENV GENOMIC AC034203 ENV GENOMIC AC018809		TARATTTGGTGCCAGGAACTGAGGCAATCGCAGGAGTTGCTGATGGCCTCACAAATCTTRACCCTGTCACTTGGGTTA
ENV GENOMIC AC018809 ENV GENOMIC HERV-K102 AF164610		TAAATTTGGTGCCAGGAACTGAGGCAATCSCAGGAGTIGCTGATGGCCTCGCAAATCTTAACCCTGTCACITGGGTTA TAAATTTGGTGCCAGGAACTGAGGCAATTSCAGGAGTIGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTTA
ENV GENOMIC FRAG. AF260253	(79)	AATTGGTGCABAACTGACGCAATGCAGAGIGCTGATGGCCTGGCAAATCTTAACCCTGTCACTTGGGTTA
CONSENSUS	(2321)	TAAATTTGGTGCCAGGAACTGAGGCAAT G AG TGCTGATGGCCTC CAAATCTTAACCCTGTCACTTGGGTTA

		24012480
ENV GENOMIC HERV MDA	(1904)	AGCATCAGAAGTITCACTATTGTAAATTTCATATTAATCCTTGTATGCCTGTCTGT
ENV GENOMIC HERV-K TAN.	(1911)	ACCATTGBAAGTAGTAGATTATAAAACCTCATATIAATCCTTGTGTGCCCTGTTTTGTCTGTTGTTAGTCTGCAGGGET
ENV GENOMIC AC025420	(1908)	ACCATTGGAAGTACTACGATTATAAAATCTCATATTAATCCTTGTGTGCCTGTTTTGTCTGTTGTFIAGTCTGCAGGTGT
ENV GENOMIC AP000776 ENV GENOMIC HERV-K8	(1911) (291)	ACCATTGEAGTACTACAATTATAAATCTCATATTAATCCTTGTGTGCCCCGTTTTGTCTGTGTGT-AGTCTCCAGGTGT
ENV GENOMIC HERV-KI	(1906)	ACCAYCGGAAGTACTACAATTARAAATCICATATTAATCCITRTGTGCCTGTTTBGTCTGTTGTI-AGTCTGCAGCRST
ENV HERV-K AF023261	(701)	
ENV GEN AL035086	(1698)	accateacateactic cattatian a anticateatateactic for creater for the construction of th
ENV GENOMIC AL035587	(1933)	ACCATCADAAGTICCACITATICTAAATTICATATTAATCCTIGTAIGCCIGTICTGTCIGTIGTIAGTCCACAGGIGI
ENV GENOMIC AC012068 ENV GENOMIC AF277315	(1904) (1917)	actacted baaring careerig caratter state at the state of
ENV GENOMIC AF277313 ENV GENOMIC AF027650	(700)	
ENV GENOMIC AC078899	(2170)	ACCAYCAGAAGTACTACTATTATAAATTTCATATTAATCCTTGTGTGCCTGTICTGTGTGTTAGTCTGCAGGTGT
ENV GENOMIC HERV-KII	(1615)	ACCATCGERAGTACTATGATTATAAATCTCATATTATCCTTGTGTGCCTGTTTGTCTGTGTGTAGTCTGCAGGTGT
ENV GENOMIC AC008813	(1238)	ACCAYCORAGETICTACTATIONAAATTICCTATTAATCCTINTIGRGECTGFICINGICUSTIGFIACICIGCACGUUT
ENV GENOMIC AC012309 ENV GENOMIC AL121932	(1905) (1538)	ACCAYCOMACTIVITACTATIGTAAAATTTCCTATTAATCCTITATCAGECLUTICUTCUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTU
ENV GENOMIC AD121932 ENV GENOMIC AD000090	(1922)	ACUATO GIAAGTAUTATAATTATAAAATTIYAATATTAATYOTTIGT GIGCONGTICIGYUHATIATI - AGIOTGCAGGGET
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(1585)	ACCATTOGAAGTACTACTATTATAAATTTOATATTAATCCTTGTCGCCTGTTTTUTCTGTTGTACTCTCCAGGTGT
ENV GENOMIC AF235103 ENV GENOMIC AC026786	(1978) (1852)	accatt Giavantic a ctatic caatitt grattatic city for concernation of the concernation of the concerned of th
ENV GENOMIC AC020788 ENV GENOMIC AC034203	(1403)	
ENV GENOMIC AC018809	(1614)	ACGATTGGAAGTACTACGATTATAAATTTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTCAAGTCTGCAGGTGT
ENV GENOMIC HERV-K102 AF164610	(1671)	ACCATTGGAAGTACTACCATTATAAAATCTCATATPAATCCPTGTGTGCCTGTTTTGTCTGTTGTTAGTCTGCAGGTGT
ENV GENOMIC FRAG. AF260253	(151)	ATCCATCHAAGTAGTAGGAT-GTAAATCTCATATTAATCATTGTGTGCCTGTTTTGTCTGTFGTTGGTCTGCAG3TGT
CONSENSUS	(2401)	ACCAT GAAGT C AC ATT TAAAT TCATATTAATCCTTGT TGCCTGTT TGTCTGTIGTT AGTCT CAGGTGT
		2481 2560
ENV GENOMIC HERV MDA	(1982)	ATCCASCAGCTCCAAAAAAAAAAACAGCAACCAGAAAAATGGGCCATAGTGACGATGGTJGTTTTGTCAAAAAAAAAA
ENV GENOMIC HERV-K TAN.	(1989)	ACCCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGCCATGATGACGATGGCGGGTTTTGTCGAAAAGAAAAGAGGGGG
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(1986) (1989)	ACCCAALAGCTCCGAAGAGAGAGAGCGGCCTTGAGAACGGGCCATGATGACGATGGTGGTTTTTTTGGAAAAGAGAAAAGGGGG ACCCAACAGGTCTGAAGAGAGAGAGCAGCATGGAGAACGGGCCATGATGAGGATGGCGGTTTTTTTGAAAAAAAA
ENV GENOMIC HERV-K8	(291)	
ENV GENOMIC HERV-KI	(1984)	RCCCBALAGCTCOGAAGAGAGAGAGGGGACCATCGAGAACAGGCCATCATGACGATAGTGGTTTTGTCGAAAAGAAAAGGGGGG
ENV HERV-K AF023261	(701)	ACCCASCAGCTCCGAAGAGACAGCGACCATCAAGAACGGGGCCATGATGACGATGGCAGTTTTGTCAAAAAGCGAAGAGGAG
ENV GEN AL035086	(1776)	accccascarccaaaaaaaaaaaaaaaaaaaaaaaaaaa
ENV GENOMIC AL035587	(2011)	ACCESCIACTCCCAAGACACCACCACCACCACGACGACGACGACGACGAC
ENV GENOMIC AC012068 ENV GENOMIC AF277315	(1982) (1995)	atecageagetectgagagagagagegecageagagagagagagagagagagag
ENV GENOMIC AF277313 ENV GENOMIC AF277515	(700)	N-N-KATASTICAMAGUADHA, MUCHARAUMAGUAGUANAGUANAGUANAGUANAGUANAGUANAGUA
ENV GENOMIC AC078899	(2248)	RCCCAGCAGCTCCAAAGAGACAGGBACCATCAAGAACGGGGCCATGATAACGATGGTGGTTGTTCAAAAATAAAAGGGGG
ENV GENOMIC HERV-KII	(1693)	BCCCAACASCTCCCGAACAGCGACAGCGACCATTGAGAACGGGCCATGATGACGATGGCSGTTTTOTCGAAAAGAAAASGGGG
ENV GENOMIC AC008813	(1238)	aciccaacagerecaaagagagegggageetegagaaggggggggtatgacgatereggyttttetccaaaagaaaagggg
ENV GENOMIC AC012309	(1983)	BCCCAACAGCECCAACAACAACAGCCACCAACAACAGCCCCACGACGACGCCAACAA
ENV GENOMIC AL121932 ENV GENOMIC AD000090	(1538) (2000)	acecaacagenceqaagagagagaggaggagagagaggagagaggaggaggag
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	ACCCAACAGCICCCAADAGACAGCAACAACCACCACAA ATCCACCACCACCACACACACCACCACCACCACCACACACAC
ENV GENOMIC AC011467	(1663)	ACCCAACAGCTCCCAACAACACCACCAACAACCAACCAAC
ENV GENOMIC AF235103	(2056)	ATCCRECAGCTCCCGAGAGACAGCGGCTAGCAGAACOGACCATGATGATGATGCGGGTTTTGTCAAAAAGAAAAG
ENV GENOMIC AC026786 ENV GENOMIC AC034203	(1930) (1403)	
ENV GENOMIC AC034203 ENV GENOMIC AC018809	(1694)	accoracagetettgaagagagagggggeettaagaacgggeetatgatgatgatgatgatgatgatgatgatgagggggggg
ENV GENOMIC HERV-K102 AF164610	(1749)	ACCCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTCGAAAAGAAAAGGGGG
ENV GENOMIC FRAG. AF260253	(228)	ACCCAACAGCTCGGAAGAGAGAGCGACCATCGAGAACGGGCCATGATGACGGTCTTTGTCGAAAAGAAAAGGGGG
CONSENSUS	(2481)	A CCA CAGCTCCGAAGAGAGAGAGGGGCCAC C AGAACGGGCCATGATGACGATGG GGTTTTGTC AAAAGAAAAG
		25612640
ENV GENOMIC HERV MDA	(2062)	ggatatgtaaggaaaaagagagatcagactticactgt-gtctatgtagaaa-aggaagacataagaaactccattt
ENV GENOMIC HERV-K TAN.	(2069)	AAATIST.GGGGAAAAGCAAGAGAGAGAGATCAGATIGTITACTIST- GICT.GTGTAGAAAGAAGTAGACATAGGAGACACCAITT
ENV GENOMIC AC025420 ENV GENOMIC AP000776	(2066) (2069)	haatgtgeggaaaaccaagagatcagattgttgttactgt-gfctgtagaaagaagtagaatggagactaccattg haatgtgeggaaaaccaagagagatcagattg1Cactgt-gtctgtagaaagaagtagacatbgagagactccattt
ENV GENOMIC APOUNTS		
ENV GENOMIC HERV-KI	(2064)	RALITSTOSSGAAAAGCAAGAGAGAGATCAG8TTYTTACTOT-STCTOTSTAGAAAGAAGTAGAACATAGBAGACHCCATTT
ENV HERV-K AF023261	(701)	
ENV GEN AL035086	(1856)	ahtor gegegaaagaabagaabagaatcagattettactot-etettatgeagaa-aggaagacattagbabteccattr
ENV GENOMIC AL035587	(2086)	ATATOTAGGAAAAGAGAGATCOACTOTTACTGT-GTCTACATAGAAA-AGGAAGACATAA
ENV GENOMIC AC012068	(2052)	htatogialoganaadanaadaotoodictottactor-ototratotagani-mganagaratahdagactocatti htatocaucangadanagagacagacagacagitactottactottactagagad-nogangacastahdanactocatti
ENV GENOMIC AF277315 ENV GENOMIC AF027650	(700)	
		NAGIGTGEGEAAAAGAAAGAGAGAGAGAGACIGTTACTAT-STCTATGTAGAAAGAAGAAGAAGAAGAAGAAGACTCCATTT
ENV GENOMIC HERV-KII		RAANSTGGGGAAAAGCAAGAGAGATCAGATHITTACTGT-BTCTGTGTAGAABCAAGTAGACATBGGAGACTCCATTT
ENV GENOMIC AC008813	(1238)	AAATG EGAAAAG AGAGATCAGACTGTTACGT-GTLTATGCAGAAATAAGTAGACATAAGAACATTAAGTACACTTCGTTT
ENV GENOMIC AC012309 ENV GENOMIC AL121932	(2063)	AATGEGAAAAGAGAGATCAGACIGTTACCGT-GILLATGCAGAAATAAGTAGACATAAGAGACILLGITT
ENV GENOMIC ALIZISZ ENV GENOMIC AD000090	(2080)	AAATATGGGGAAAAGAAGAAGAAGAAGATCAGATTGTTACTOT-GTCTGTGTAGAAAGAAGAAGAAGAGAGAGAGAGAGAGAGA
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(1699)	
		ATATISTAGGGAAAAGAGAGAGAGAGACCAGACTOTTACTGT- GTCTATGTAGAA-GGGAAGACATAAGAGACTCCATTT AAATISTAGGGAAAAGAAGAAGAGAGACCAGACTOTCACTGT-GCCTATGTAGAA-GGGAAGACATAAGAGACTCCATTT
ENV GENOMIC AC026786 ENV GENOMIC AC034203	(1403)	
ENV GENOMIC AC018809	(1774)	AAATGTGGGGAAAAGAGAGATCAGATTGTTACTGTGTCGTGTAGAAATAAGTAGACATAGGAGACTCCATTT
ENV GENOMIC HERV-K102 AF164610	(1829)	Anatyficggganangcangagagatcagattgttactgf-gictgtgganaggaggagacatgggagactccattt;
ENV GENOMIC FRAG. AF260253	(308)	BAATGTOGGGAAAAGCAAGAGAGATCAGACTGTTACTGT-ETCTGTGTAGAAGAAGTAGAAGTAGACATAGGAGACTCCATT
CONSENSUS	(2561)	A ATGT GGGAÄAAG AGAGAGATCAGA TGTTACTGT GTCT TGTAGAAA A G AGACATA GAGACTCCATTT

FIGURE 6 CONTD ...

		2641 2707
ENV GENOMIC HERV MDA	(2136)	TGATCIGTACTAA
ENV GENOMIC HERV-K TAN.	(2146)	TGTTATGTACTAA
ENV GENOMIC AC025420	(2143)	TETTATGTAGTAA
ENV GENOMIC AP000776	(2146)	IGTTAIGTACTAA
ENV GENOMIC HERV-K8	(291)	
ENV GENOMIC HERV-KI	(2141)	TGTTATSTACTAA
ENV HERV-K AF023261	(701)	
ENV GEN AL035086	(1931)	TGATCRGTACTAA
ENV GENOMIC AL035587	(2146)	
ENV GENOMIC AC012068	(2138)	TGAAAAAGACCTGTACTTTGAACAATT
ENV GENOMIC AF277315	(2152)	NACCINTA
ENV GENOMIC AF027650	(700)	
ENV GENOMIC AC078899	(2405)	TGTTC/GTACTAAGAGAAAATTCTTCTGCCTTGAGATGCTGTTAA
ENV GENOMIC HERV-KII	(1850)	MITTCINTACTAA
ENV GENOMIC AC008813	(1238)	
ENV GENOMIC AC012309	(2133)	BYITOIGTACCAAG
ENV GENOMIC AL121932	(1538)	
ENV GENOMIC AD000090	(2157)	IGTTCHGTAGTAA
ENV GEN AL160008	(647)	
ENV GENOMIC HEU32496	(441)	
ENV GENOMIC AC011467	(1699)	
ENV GENOMIC AP235103	(2212)	TUAAAAAGACCTGTACTTTGAACAATTGCTTTGCTCAGATGTTGTTAATTTGTAGTTTT
ENV GENOMIC AC026786	(2086)	TJAAAAAGACCTGTACTTTAAACAATTGCTTTGCTGAGATGTTGTTAATTTGTAGCTTTCCCCAGCC
ENV GENOMIC AC034203	(1403)	
ENV GENOMIC AC018809	(1846)	TGCTCTGTACTAAG
ENV GENOMIC HERV-K102 AF164610	(1906)	TGTTATGTGTTAAGAAAAATTCTT
ENV GENOMIC FRAG. AF260253	(385)	ngttchgtactaa
CONSENSUS	(2641)	TG TGTAC

FIGURE 7

		1 60
GI_4185938_EMB_CAA76878.1_ ((1)	MGQTKSKIKSKYASYLSFIKILLKRGGVKVSTKNLIKLFQIIEQFCPWFPEQGTL
GI_4185942_EMB_CAA76881.1_	(1)	MGQTKSKIKSKYASYLSFIKILLKRGGVKVSTKNLIKLFQIIEQFCPWFPEQGTL
GI_4185946_EMB_CAA76884.1_	(1)	MGQTKSKIKSKYASYLSFIKILLKRGGVKVSTKNLIKLFQIIEQFCPWFPEQGTL
GI_5931704_EMB_CAB56602.1_	1)	MGQTKTKSKYASYLSFIKILLKRGGVRVSTKNLIKLFQTTEQFCPWFPEQGNL
GAG OF AB047240	(1)	MGQTKSKTKSKYASYLSFIKILLKRGGVRVSTKNLIKLFQIIEQFCPWFPEQGTL
TRANSLATION OF ORF99	(1)	YKKAGLGQTKSKTKSKYASYLSFIKILLKRGGVRVSTKNLIKLFQIIEQFCPWFPEQGTL
TRANSLATION OF G226TOP-LINK	(1)	
TRANSLATION OF G591TOP-LINK	(1)	
TRANSLATION OF LNCAP-GAG	(1)	MGQTKSKTKSKYASYLSFIKILLKRGGVRVSTKNLIKLFQIIEQFCPWFPEQGTL
GAG106-135	(1)	
GAG186-215	(1)	
GAG46-75	(1)	CPWPPEQGTL
PDG-G1	(1)	
PGD-G2	(1)	
PGD-G3	(1)	
CONSENSUS	(1)	CPWFPEQG L
		61 120
GI_4185938_EMB_CAA76878.1_ (5	56)	DLKDWKRIGKELKQAGRKGNIIPLTVWNDWAIIKAALEPFQTEEDSVSVSDAPGSCTIDC
GI_4185942_EMB_CAA76881.1_ (5	56)	DLKDWKRIGKELKQAGRKGNIIPLTVWNDWAIIKAALEPFQTEEDSVSVSDAPGSCIIDC
GI_4185946_EMB_CAA76884.1_ (5	56)	DLKDWKRIGKELKQAGRKGNIIPLTVWNDWAIIKAALEPFQTEEDSVSVSDAPGSCTIDC
GI_5931704_EMB_CAB56602.1_ (5	54)	DLEDWKRIGKELKQAGRKGNIIPLTVWNDWPIIKAALEPFQTEDS-VSVSDAPGSCIIDC
GAG OF AB047240 (5	56)	DLKDWKRIGEELKQAGRKGNIIPLTVWNDWAIIKAALEPFQTKEDSVSVSDAPGSCŰIDC
TRANSLATION OF ORF99 (6	51)	DLKDWKRIGEELKQAGRKGNIIPLTVWNDWAIIKAALEPFQTKEDSVSVSDAPGSCVIDC
TRANSLATION OF G226TOP-LINK	(1)	
TRANSLATION OF G591TOP-LINK	(1)	
TRANSLATION OF LNCAP-GAG (56)	DLKDWKRIGEELKQAGRKGNIIPLTVWNDWAIIKAALEPFOTKEDSVSVSDAPGSCVIDC
GAG106-135	(1)	DAPGSCIIDC
GAG186-215	(1)	
	-	DLKDWKRIGKELKQAGRKGN
GAG46-75 (1	11)	
	(1) (1)	DWKRIGKBLKQAGRKGDWKRIGKBLKQAGRKGDWKRIGKBLKQAGRKG
PDG-G1		DWKRIGKBLKQAGRKG
PDG-G1 PGD-G2	(1)	DWKRIGKBLKQAGRKG
PDG-G1 PGD-G2 PGD-G3	(1) (1)	DWKRIGKBLKQAGRKG
PDG-G1 PGD-G2 PGD-G3	(1) (1) (1)	DWKRIGKBLKQAGRKG

FIGURE 7 CONTD...

		121 180
GI_4185938_EMB_CAA76878.1_		NENTRXKSQKETEGLHCEYVAEPVMAQSTQNVDYNQLQEVIYPETLKLEGKGPELVGPSE
GI_4185942_EMB_CAA76881.1_	(116)	NENTRKKSQKETESLHCEYVAEPVMAQSTQNVDYNQLQEVIYPETLKLEGKGPELVGPSE
GI_4185946_EMB_CAA76884.1_	(116)	NENTRAKSQKETEGLHCEYVAEPVMAQSTQNVDYNQLQEVIYPETLKLEGKGPELVGPSE
GI_5931704_EMB_CAB56602.1_ GAG OF AB047240	(113) (116)	NEKTRKKSQKETETLHCEYVAEPLMAQSTQNVDYNQLQEVIYPETLKLEGKGPELVGPLE NEKTGRKSQKETESLHCEYVTEPVMAQSTQNVDYNOLQGVIYPETLKLEGKGPELVGPSE
TRANSLATION OF ORF99	(110)	NEKTGRKSQKETESLHCEYVTEPVMAQSTQNVDYNQLQGVIYPETLKLEGKGPELVGPSE
TRANSLATION OF G226TOP-LINK	(1)	
TRANSLATION OF G591TOP-LINK	(1)	
TRANSLATION OF LNCAP-GAG	(116)	NEKTGRKSQKETESLHCEYVTEPVMAQSTQNVDYNQLQGVIYPETLKLEGKGPELVGPSE
GAG106-135	(11)	NENTRKKSQKETEGLHCEYV
GAG186-215	(1)	
GAG46-75 PDG-G1	(31) (17)	
PGD-G1 PGD-G2	(1)	
PGD-G3	(1)	
CONSENSUS	(121)	NE T KKSQKETE LHCEYV
		101 040
CT 4195029 END CAA76970 3	(176)	
GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76881.1_		SKPRGTSPLPAGQVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPESQYGYPGMPP SKPRGTSRLPAGQVPVTLQPQTQVKENKTQPPVAYQYWPPAELQYRPPVESQYGYPGMPP
GI 4185946 EMB CAA76884.1	(176)	Ph. The second s
GI 5931704 EMB CAB56602.1	(173)	
GAG OF AB047240	(176)	
TRANSLATION OF ORF99	(181)	
TRANSLATION OF G226TOP-LINK	(1)	SQYGYPGMPP
TRANSLATION OF G591TOP-LINK	(1)	
TRANSLATION OF LNCAP-GAG	(176)	
GAG106-135	(31)	
GAG186-215	(1)	AGQVPVTLQPQKQVKENKTQPPVAYQYWPP
GAG46-75 PDG-G1	(31) (17)	
PDG-G1 PGD-G2	(1)	
· PGD-G3	(1)	
CONSENSUS	(181)	AGQV VTLOPO QVKENKTO PVAYOYWPP SOYGY GMPP
		241 200
GT 4195938 PMP (3376979)	(236)	
GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76881_1		${\tt APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA$
GI_4185942_EMB_CAA76881.1_	(236)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRRGSELHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA
	(236) (236)	${\tt APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA$
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_	(236) (236) (233)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRRGSELHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB56602.1 GAG OF AB047240 TRANSLATION OF ORF99	(236) (236) (233) (236) (241)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRGSELHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSQQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF ORF99 TRANSLATION OF G226TOP-LINK	(236) (236) (233) (236) (241) (11)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRGSELHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRQCYGTT
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF ORF99 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK	(236) (236) (233) (236) (241) (11) (11)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRQCYGTT- ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET APQGRAPYPQPPTRRLNPTA-
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF ORP99 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG	(236) (236) (233) (236) (241) (11) (11) (236)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRGSELHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQQRAPYPQPPTRRLNPTAPPSRQSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQQRAPYPQPPTRRQCYGTT
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF ORP99 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135	(236) (236) (233) (236) (241) (11) (11) (236) (31)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET APQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215	(236) (236) (233) (236) (241) (11) (11) (236)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRGSELHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQQRAPYPQPPTRRLNPTAPPSRQSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQQRAPYPQPPTRRQCYGTT
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF ORP99 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135	(236) (233) (233) (236) (241) (11) (11) (236) (31) (31)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRGSELHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQCSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTVRLNPTASRSQQGSTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB56602.1 GAG OF AB047240 TRANSLATION OF ORF99 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75	(236) (233) (233) (236) (241) (11) (11) (236) (31) (31) (31)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET APQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG186-215 GAG46-75 PDG-G2 PGD-G3	(236) (233) (233) (241) (11) (11) (236) (31) (31) (31) (17)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRGSELHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQQRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQQREPYPQPPTRRQCYGTT
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF 0RP99 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PDG-G1 PGD-G2	(236) (236) (233) (236) (241) (11) (236) (236) (31) (31) (31) (17) (1) (1)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG186-215 GAG46-75 PDG-G2 PGD-G3	(236) (236) (233) (236) (241) (11) (236) (236) (31) (31) (31) (17) (1) (1)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVMLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG186-215 GAG46-75 PDG-G2 PGD-G3	(236) (236) (233) (236) (241) (11) (236) (236) (31) (31) (31) (17) (1) (1)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVHLEPMPPGEGA APQQRAPYPQPPTRRQCYGTT
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG106-135 GAG186-215 GAG186-215 PDG-G2 PGD-G3 CONSENSUS	(236) (236) (233) (236) (241) (11) (236) (31) (31) (31) (31) (11) (1) (1) (241)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTR
GI 4185942 EMB CAA76881.1 GI 4185946 EMB CAA76884.1 GI 5931704 EMB CAB76884.1 GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG186-215 GAG46-75 PDG-G2 PGD-G3	(236) (236) (233) (236) (241) (11) (236) (31) (31) (31) (31) (11) (241) (241)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVHLEPMPPGEGA APQQRAPYPQPPTRRQCYGTT
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG186-215 GAG46-75 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_	(236) (236) (233) (241) (11) (236) (31) (31) (31) (17) (1) (241) (241) (296) (296)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTVRLNPTASRSQQGSTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTR
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG106-135 GAG106-215 GAG46-75 PDG-G1 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76878.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_	(236) (236) (233) (241) (11) (236) (31) (31) (31) (17) (1) (241) (241) (296) (296)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQQRAPYPQPPTRRQCYGTT
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG106-135 GAG186-215 GAG186-215 GAG46-75 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76881.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240	(236) (236) (233) (236) (241) (11) (236) (31) (31) (31) (31) (11) (241) (241) (296) (296) (296) (296) (296) (296)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PDG-G1 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185946_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF ORF99	(236) (236) (233) (236) (241) (11) (236) (31) (31) (31) (31) (17) (11) (241) (296) (297) (APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTVRLNPTASRSQQGSTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET AQGGRPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET AQGGRPYPQPPTR AQGAPAPSSIKKLKDMKEGVKQYGPNSPYMRTLLDSIAHGHRLIPYDWEILAK QEGEPPTVEARYKSFSIKKLKDMKEGVKQYGPNSPYMRTLLDSIAHGHRLIPYDWEILAK QGGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTLLDSIAHGNRLTPYDWESLAK QVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTLLDSIAHGNRLTPYDWESLAK
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PDG-G1 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76878.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK	(236) (236) (233) (241) (11) (236) (31) (31) (17) (13) (17) (241) (296) (297)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQQRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQQRAPYPQPPTRRQCYGTT
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG106-135 GAG106-135 GAG106-215 GAG46-75 PDG-G1 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76878.1_ GI_4185942_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK	(236) (236) (233) (236) (241) (11) (236) (31) (31) (31) (17) (13) (17) (12) (296) (1) (1) (206) (231) (231) (231) (231) (231) (31) (1) (232) (23	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 CONSENSUS GI_14185938_EMB_CAA76878.1_ GI_5931704_EMB_CAA76878.1_ GI_59317000000000000000000000000000000000000	(236) (236) (233) (236) (236) (241) (11) (236) (31) (31) (11) (241) (296) (21) (21) (21) (21) (21) (21) (21) (21	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG106-135 GAG186-215 GAG46-75 PDG-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76881.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135	(236) (236) (233) (236) (241) (11) (236) (31) (31) (31) (11) (241) (296) (296) (296) (296) (296) (296) (296) (301) (11) (296) (31)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET AQGGRPPTVEARYKSFSIKKLKDMKEGVKQYGPNSPYMRTLLDSIAHGHRLIPYDWEILAK QEGEPPTVEARYKSFSIKKLKDMKEGVKQYGPNSPYMRTLLDSIAHGHRLIPYDWEILAK QEGEPPTVEARYKSFSIKKLKDMKEGVKQYGPNSPYMRTLLDSIAHGHRLIPYDWEILAK QVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTLLDSIAHGNRLTPYDWESLAK QVGAPARAETRCEPFTMKMLKDIKEGVKQYGSNSPYIRTLLDSIAHGNRLTPYDWESLAK
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 CONSENSUS GI_14185938_EMB_CAA76878.1_ GI_5931704_EMB_CAA76878.1_ GI_59317000000000000000000000000000000000000	(236) (236) (233) (236) (241) (11) (236) (31) (31) (31) (17) (241) (296) (296) (296) (296) (296) (296) (301) (31) (11) (296) (31)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PGD-G2 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76878.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215	(236) (236) (236) (236) (236) (241) (11) (236) (31) (31) (17) (13) (17) (296)	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQQRAPYPQPPTRRQCYGTT
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PDG-G1 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76878.1_ GI_5931704_EMB_CAA76881.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK GAG186-215 GAG186-215 GAG46-75	(236) (236) (236) (236) (236) (241) (11) (236) (31) (31) (131) (11) (241) (296) (297	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG106-135 GAG166-215 PDG-G1 PGD-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76878.1_ GI_5931704_EMB_CAA76884.1_ GI_5931704_EMB_CAA76884.1_ GI_5931704_EMB_CAA76884.1_ GI_5931704_EMB_CAA76884.1_ GI_5931704_EMB_CAA5684.1_ GAG OF AB047240 TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG166-215 GAG46-75 PDG-G1	(236) (236) (236) (236) (236) (241) (11) (236) (31) (31) (131) (11) (241) (296) (297	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET
GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAB76884.1 GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PDG-G2 PGD-G3 CONSENSUS GI_4185938_EMB_CAA76878.1_ GI_4185946_EMB_CAA76878.1_ GI_5931704_EMB_CAB56602.1 GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PDG-G1 PGD-G2	(236) (236) (236) (236) (236) (241) (11) (236) (31) (31) (11) (241) (296) (297	APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQGRAPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTAPPSRQGSKLHEIIDKSRKEGDTEAWQFPVTLEPMPPGEGA APQDREPYPQPPTRRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET ALQGRAPYPQPPTVRLNPTASRSGQGGTLHAVIDEARKQGDLEAWRFLVILQLVQAGEET

FIGURE 7 CONTD...

GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76881.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PDG-G1 PGD-G2 PGD-G3 CONSENSUS	(356) (356) (356) (356) (361) (31) (31) (31) (31) (31) (17) (17) (11) (361)	361 420 SSLSPSQPLQFKTWWIDGVQEQVRRNRAANPPVNIDADQLLGIGQNWSTISQQALMQNEA SSLSPSQPLQFKTWWIDGVQEQVRRNRAANPPVNIDADQLLGIGQNWSTISQQALMQNEA SSLSPSQPLQFKTWWIDGVQEQVRRNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEA SSLSSSQYLQFKTWWIDGVQEQVRKNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEA SSLSSSQYLQFKTWWIDGVQEQVRRNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEA SSLSSSQYLQFKTWWIDGVQEQVRKNQATKPTVNIDADQLLGTGPNWSTINQQSVMQNEA
GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76884.1_ GI_5931704_EMB_CAB76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PGD-G2 PGD-G3 CONSENSUS	(416) (416) (254) (416) (421) (31) (131) (31) (31) (31) (17) (17) (17) (12)	421 480 IEQURAICLRAWEKIQDPGSTCPSFNTVRQGSKEPYPDFVARLQDVAQKSIADEKARKVI IEQVRAICLRAWEKIQDPGSTCPSFNTVRQGSKEPYPDFVARLQDVAQKSIADEKARKVI IEQVRAICLRAWGKIQDPGTAFP-INSIRQGSKEPYPDFVARLQDAQKSITDDNARKVI IEQVRAICLRAWGKIQDPGTAFP-INSIRQGSKEPYPDFVARLQDAQKSITDDNARKVI IEQVRAICLRAWGKIQDPGTAFP-INSIRQGSKEPYPDFVARLQDAQKSITDDNARKVI
GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG106-135 GAG186-215 GAG46-75 PDG-G1 PGD-G2 PGD-G3 CONSENSUS	(476) (476) (254) (475) (480) (31) (17) (31) (31) (31) (17) (17) (12)	VELMAYENANPECQSAIKPLKGKVPAGSDVISEYVKACDGIGGAMHKAMLMAQAITGVVL VELMAYENANPECQSAIKPLKGKVPAGVDVITEYVKACDGIGGAMHKAMLMAQAMRGLTL VELMAYENANPECQSAIKPLKGKVPAGVDVITEYVKACDGIGGAMHKAMLMAQAMRGLTL
GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76881.1_ GI_4185946_EMB_CAA76884.1_ GI_5931704_EMB_CAA566602.1_ GAG OF AB047240 TRANSLATION OF G225TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PGD-G2 PGD-G3 CONSENSUS	(536) (536) (254) (535) (540) (31) (1) (535) (31) (31) (31) (17)	GGQVRTFGKKCYNCGQIGHLKRSCPVLNKQNIINQAITAKNKKPSGLCPKCGKGKHWANQ GGQVRTFGKKCYNCGQIGHLKRSCPVLNKQNIINQAITAKNKKPSGLCPKCGKGKHWANQ

FIGURE 7 CONTD....

CONSENSUS (601) C SKFDK GQPLSGN	GI_4185938_EMB_CAA76878.1_ GI_4185942_EMB_CAA76881.1_ GI_4185946_EME_CAA76884.1_ GI_5931704_EMB_CAB56602.1_ GAG OF AB047240 TRANSLATION OF G226TOP-LINK TRANSLATION OF G226TOP-LINK TRANSLATION OF G591TOP-LINK TRANSLATION OF LNCAP-GAG GAG106-135 GAG186-215 GAG46-75 PDG-G1 PGD-G2 PGD-G2	(595) (595) (254) (595) (301) (595) (31) (31) (31) (17) (17) (1)	CRSKFDKNGQPLSGNEQRGQPQAPQQTGAFPIQPFVPHGPQGQQP-PLSQVFQGISQLPQ CRSKFDKNGQPLSGNEQRGQPQAPQQTGAFPIQPFVPQGFQGQQP-PLSQVFQGISQLPQ CRSKFDKDGQPLSGNRKRGQPQAPQQTGAFPVQLFVPQGFQGQQPLQKIPPLQGVSQLQQ CHSKFDKDGQPLSGNRKRGQPQAPQQTGAFPVQLFVPQGFQGQQPLQKIPPLQGVSQLQQ CRSKFDKNGQPLSGNRKRGQPQAPQQ- CHSKFDKDGQPLSGNRKRGQPQAPQQTGAFPVQLFVPQGFQGQQPLQKIPPLQGVSQLQQ
	PGD-G3	(1)	CRSKFDKNGQPLSGNE

GI_4185942_EMB_CAA76881.1_ (63 GI_4185946_EMB_CAA76884.1_ (64 GI_5931704_EMB_CAB56602.1_ (62 GAG OF AB047240 (64 TRANSLATION OF ORF99 (64 TRANSLATION OF G226TOP-LINK (17 TRANSLATION OF G591TOP-LINK (17 TRANSLATION OF G591TOP-LINK (17 TRANSLATION OF G591TOP-LINK (17 GAG106-135 (17 GAG186-215 (17 GAG186-215 (17 GAG46-75 (17 PGD-G1 (17) PGD-G3 (17)	661 673 54) YNNCPPPQAAVQQ 54) YNNCPPPQAAVQQ 54) YNNCPPPQAAVQQ 54) 55) SNSCPAPQQAAPQ 60) SNSCPAPQQAAPQ 31) 55] SNSCPAPQQAAPQ 31) 31) 31) 17)
	61)

FIGURE 8

		1 60
GI_4185939 EMB CAA76879.1_	(1)	MLTDLRAVNAVIQPMGPLQPGLPSPAMIPKDWPLIIIDLKDCFFTIPLAEQDCEKFA
GI_4185943_EMB_CAA76882.1_	(1)	MLTDLRAVNAVNAVIQPMGPLQPGLPSLAMIPKDWPLIIIDLKDCFFTIPLAEQDCEKFA
GI_4185947 EMB CAA76885.1_	(1)	MLTDLRAVNAVIQPMGPLQPGLPSPAMIPKDWPLIIIDLKDCFFTIPLAEQDCEKFA
GI 5931705 EMB CAB56603.1	(1)	MIPKDWPLIIIDLKDCFFTIPLAEQDCEKFA
ENV OF AB047240	(1)	
TRANSLATION OF P386TOP-LINK	(1)	
TRANSLATION OF POL349-LINK	(1)	
LNCAP-GENOMEA-POLORF	(1)	
TRANSLATION OF LNCAP-POL-GENA-GOODA	(1)	
TRANSLATION OF ORF111-10	(1)	
PGD-P1	(1)	
PGD-P2	(1)	
PGDP3	(1)	
CONSENSUS	(1)	
		61 120
GI_4185939 EMB CAA76879.1_	(58)	61 120 FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC
GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_	(58) (61)	
		FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC
GI_4185943_EMB_CAA76882.1_	(61)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC
GI_4185943 RMB CAA76882.1 GI_4185947 EMB CAA76885.1 _	(61) (58)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC
GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1_	(61) (58) (32)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC
GI_4185943 EMB_CAA76882.1_ GI_4185947 EMB_CAA76885.1_ GI_5931705 EMB_CAB56603.1_ ENV OF AB047240	(61) (58) (32) (1)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC
GI_4185943 EMB_CAA76882.1_ GI_4185947 EMB_CAA76885.1_ GI_5931705 EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK	(61) (58) (32) (1) (1)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC
GI_4185943 EMB CAA76882.1 GI_4185947 EMB CAA76885.1 GI_5931705 EMB CAB56603.1 ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF POL349-LINK	(61) (58) (32) (1) (1) (1)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC
GI_4185943 EMB_CAA76882.1_ GI_4185947 EMB_CAA76885.1_ GI_5931705 EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386T0P-LINK TRANSLATION OF P0L349-LINK LNCAP-GENOMEA-POLORF	(61) (58) (32) (1) (1) (1) (1)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYFDDILC
GI_4185943 EMB_CAA76882.1_ GI_4185947 EMB_CAA76885.1_ GI_5931705 EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF POL349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF LNCAP-GENA-GOODA	(61) (58) (32) (1) (1) (1) (1) (1)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKFSDCYIIHYFDDILC
GI_4185943 EMB CAA76882.1 GI_4185947 EMB CAA76885.1 GI_5931705 EMB CAA76885.1 ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF POL349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF INCAP-POL GENA-GOODA TRANSLATION OF ORP111-10	(61) (58) (32) (1) (1) (1) (1) (1) (1) (1)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTLCQTFVGRALQPVRKFSDCYIIHCIDDILC
GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P01349-LINK LNCAP-GENOMEA-P01AF TRANSLATION OF INCAP-P01-GENA-GODA TRANSLATION OF ORF111-10 FGD-F1	(61) (58) (32) (1) (1) (1) (1) (1) (1) (1) (1)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYFDDILC
GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386T0P-LINK TRANSLATION OF P01349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF LNCAP-OL-GENA-GOODA TRANSLATION OF ORF111-10 PGD-P1 PGD-P2	(61) (58) (32) (1) (1) (1) (1) (1) (1) (1) (1)	FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHCIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYIIHYIDDILC FTIPAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVRDKFSDCYIIHYFDDILC

FIGURE 8 CONTD ...

		121 180
GI 4185939 EMB_CAA76879.1	(118)	AAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDT
GI 4185943 EMB CAA76882.1		AAEMKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLEMQIENRKIKPPKIEIRKDT
GI_4185947_EMB_CAA76885.1_		AAETKDKLIDCYTFLQAEVANAGLAIASDKIQTSTPFHYLGMQIENRKIKPQKIEIRKDT
GI_103547_AMB_CAB56603.1_	(92)	AABTKDKLIDCYTFLQAEVANAGLAIASDKIQISTPFHYLGMQIENRKIKPQKIEIRKDT
ENV OF AB047240	· (1)	
TRANSLATION OF P386TOP-LINK	(1)	
TRANSLATION OF POL349-LINK	(1)	
LNCAP-GENOMEA-POLORF	(1)	
TRANSLATION OF LNCAP-POL-GENA-GOODA	(1)	
TRANSLATION OF ORF111-10	(1)	
PGD-P1	(1)	IENRKIKPQKIEIRKD-
PGD-P2	(1)	
PGDP3	(1)	
CONSENSUS	(121)	
		181 240
GI_4185939_EMB_CAA76879.1_	(178)	LKTLNDFOKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLTPEATKEIKLVEE
GI_4185943_EMB_CAA76882.1_	(181)	LKTLNDFOKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLTPEATKEIKLVEE
GI 4185947 EMB CAA76885.1		LKTLNDFOKLLGDINWIRPTLGIPTYAMSNLFSILRGDSDLNSKRMLTPEATKEIKLVEE
GI_5931705_EMB_CAB56603.1_		LKTLNDFOKLLGDINWIRPTIGIPTYAMSNLFSILRGDSDLNSKRMLTPRATKEIKLVEB
ENV OF AB047240	(1)	
TRANSLATION OF P386TOP-LINK	(1)	
TRANSLATION OF POL349-LINK	(1)	
LNCAP-GENOMEA-POLORF	(1)	
TRANSLATION OF LNCAP-POL-GENA-GOODA	(1)	
TRANSLATION OF ORF111-10	(1)	
PGD-P1	(17)	
PGD-P2	(1)	***************************************
PGDP3	(1)	
CONSENSUS	(181)	
		241 300
GI_4185939_EMB_CAA76879.1	(238)	KIQSAQINRIDPLAPLOLLIFATAHSPTGIILONTDLVEWSFLPHSTWKTFTLYLDOTAT
GI_4185943_EMB_CAA76882.1_		KIQSAQINRIDPLAPLOLLIFATAHSPTGIILONTDLVEWSFLPHSTVKTFTLYLDOMAT
GI 4185947 EMB CAA76885.1		KIQSAQINRIDPLAPLQLLIFATAHSPTCIIIQNTDLVEWSFLPHSTVKTFTLYLDQLAT
GI 5931705 EMB CAB56603.1		KIQSAQINRIDPLAPLQLLIFATAHSPTGIIIQNTDLVEWSFLPHSTYKTFTLYLDQLAT
ENV OF AB047240		A South and the second se
		Į AI
TRANSLATION OF P386TOP-LINK	(1)	
TRANSLATION OF POL349-LINK	(1)	
LNCAP-GENOMEA-POLORF	(1)	DHLAPLOILIFITAHSLIAIIWONTDLVIWSPLPHSTIKTFTLYLDOMAT
TRANSLATION OF LNCAP-POL-GENA-GOODA		DHLAPLOILIFGTAHSLTAIINONTDLVDWSFLPHSTHKTFTLYLDOMAT
TRANSLATION OF ORF111-10		YKKAGSDHLAPLQILIFGTAHSLTAIIVQNTDLVDWSFLPHSTIKTFTLYLDQMAT
PGD-P1	(17)	
	(477	
PGD-P2	(1)	
PGD-P2	(1)	
PGD-P2 PGDP3	(1) (1)	
PGD-P2 PGDP3	(1) (1)	
PGD-P2 PGDP3	(1) (1)	
PGD-92 PGD93 Consensus	(1) (1) (241)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT
PGD-P2 PGDP3	(1) (1) (241) (298)	D LAPLQLLIFATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_	(1) (1) (241) (298) (301)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76882.1_	(1) (1) (241) (298) (301) (298)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFVGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFVGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAA856603.1_	(1) (1) (241) (298) (301) (298) (272)	D LAPLQLLIFATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240	(1) (1) (241) (298) (301) (298) (272) (4)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFVGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFVGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1_ ENV_OF_AB047240 TRANSLATION OF P386TOP-LINK	(1) (1) (241) (241) (301) (298) (272) (4) (1)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185947_EMB_CAA76879.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF POL349-LINK	(1) (1) (241) (241) (301) (298) (272) (4) (1) (1)	D LAPLQLLIFATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFYGIIDNHYPKTKIF LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGOTRLRIIKLCGNDPDKIVVPLTKEV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P013467UP-LINK TRANSLATION OF P01349-LINK LNCAP-GENOMEA-POLORF	(1) (1) (241) (241) (301) (298) (301) (298) (272) (42) (11) (11) (51)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOTRLRIIKLCGNDPDKIVVPLTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIIKLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P386TOP-LINK LNCAP-GENOMEA-POLORP TRANSLATION OF LNCAP-GENOMEA-POLORP	(1) (1) (241) (298) (301) (298) (272) (4) (1) (1) (51) (51)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEOVROAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185947_EMB_CAA76879.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1 ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P0L349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF POL349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORF111-10	(1) (1) (298) (301) (298	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEOVROAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76822.1_ GI_4185947_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK ITRANSLATION OF P0149-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF LNCAP-POL-GENA-GODDA TRANSLATION OF ORF111-10 PGD-P1	(1) (1) (241) (298) (301) (298) (297) (298) (297) (298) (297) (298) (297) (298) (297	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIIKLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK INCAP-GENOMEA-POLORF TRANSLATION OF P0L349-LINK LNCAP-F0L-GENA-GOODA TRANSLATION OF ORF111-10 PGD-P2 PGD-P2	(1) (1) (241) (241) (298) (272) (298) (272) (298) (272) (11) (51) (51) (57) (17) (12)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEOVROAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P0L349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORJ11-10 PGD-P1 PGD-P1 PGD-P2 PGDP2	(1) (1) (241) (241) (298) (272) (272) (272) (4) (1) (51) (51) (51) (57) (17) (1) (1)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK INCAP-GENOMEA-POLORF TRANSLATION OF P0L349-LINK LNCAP-F0L-GENA-GOODA TRANSLATION OF ORF111-10 PGD-P2 PGD-P2	(1) (1) (241) (241) (298) (272) (272) (272) (4) (1) (51) (51) (51) (57) (17) (1) (1)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIIKLCGNDPDKIVVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P0L349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORJ11-10 PGD-P1 PGD-P1 PGD-P2 PGDP2	(1) (1) (241) (241) (298) (272) (272) (272) (4) (1) (51) (51) (51) (57) (17) (1) (1)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P0L349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORJ11-10 PGD-P1 PGD-P1 PGD-P2 PGDP2	(1) (1) (241) (241) (298) (272) (272) (272) (4) (1) (51) (51) (51) (57) (17) (1) (1)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P0L349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORJ11-10 PGD-P1 PGD-P1 PGD-P2 PGDP2	(1) (1) (241) (241) (298) (272) (272) (272) (4) (1) (51) (51) (51) (57) (17) (1) (1)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185943_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1 ENV OF AB047240 TRANSLATION OF PO1866707-LINK IRCAP-GENOMEA-POLORF TRANSLATION OF PO1949-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORF111-10 PGD-P1 PGD-P1 PGD-P2 PGDP3 CONSENSUS	(1) (1) (241) (241) (298) (272) (272) (272) (4) (1) (51) (51) (51) (57) (17) (1) (1)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P0L349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORJ11-10 PGD-P1 PGD-P1 PGD-P2 PGDP2	(1) (1) (241) (241) (241) (272) (272) (4) (1) (51) (51) (51) (51) (51) (51) (51)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKIVVPFTKQOVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185943_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1 ENV OF AB047240 TRANSLATION OF PO1866707-LINK IRCAP-GENOMEA-POLORF TRANSLATION OF PO1949-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORF111-10 PGD-P1 PGD-P1 PGD-P2 PGDP3 CONSENSUS	(1) (1) (298) (301) (298) (272) (4) (1) (51) (51) (57) (17) (17) (1301) (301)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPISGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCCNDPDKIVVPLTKEQVRQAPISGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVVPFNKQQVRQAPISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVVPFNKQQVRQAPISGAWQIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVVPFNKQQVRQAPISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVVPFNKQQVRQAPISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQQVRQAPISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQVRQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPDKIVPFNKQVRQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPKIVPFNKQVRQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPKIVPFNKQVRQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPKIVPFNKQVRQVRQAFISGAWIGLANFYGIIDNHYPKTKIF SGQRLRIITLCGNDPKIVPFNKQVRQXRQVRQAFISGAWIGLANFYGIIDNHYPKTKIF
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA7682.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF PO18670P-LINK TRANSLATION OF PO38670P-LINK INCAP-GENOMEA-POLORF TRANSLATION OF ONF911-10 PGD-P1 PGD-P2 PGD-P2 PGD-P3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_	(1) (1) (298) (301) (298) (272) (4) (1) (51) (51) (51) (57) (17) (17) (17) (13) (301) (358) (358) (361)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPTKEQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGAUQICANFYGIIDNHYF SGAUQICANFYGIIDNHYF SGAUQICANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQICANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNHYF SGAUQUCANFYGIIDNH
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA7682.1_ GI_4185947_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK IRCAP-GENOMEA-POLORF TRANSLATION OF P0149-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF LNCAP-POL-GENA-GODDA TRANSLATION OF ORF111-10 PGD-P1 PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA768879.1_ GI_4185943_EMB_CAA76882.1_	(1) (1) (241) (241) (241) (272) (272) (4) (1) (51) (51) (51) (51) (51) (51) (51)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTLRIIKLCGNDPDKIVVPLTKEQVRQAPISGAWQIGLANFYGIIDNHYPKTKIP LIGQGLRIITLCCNDPDKIVVPTNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGAQURAPHYTYTLGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGAQRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGAQRLRIITLCGNDPDKITVPFNKQQVRQAPISGAWQIGLANFYGIDNHYPKTKIF SGAQRLRIITLCGNDPDKITVPFNKQQVRQAPISGAWQIGLANFYGIDNHYPKTKIF SGAQRLRIITLCGNDPDKITVPFNKQQVRQAPISGAWQIGLANFYGIDNHYPKTKIF SGAQRLRIITLCGNDPDKITVFTYFTGSSNGKAAYTGPKERVIKTPQSAQRAELVAV SGALTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P386TOP-LINK INCAP-GENOMEA-POLORF TRANSLATION OF POL349-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF ORF111-10 PGD-P1 PGD-P1 PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_	(1) (1) (298) (301) (298) (272) (4) (1) (51) (51) (51) (51) (51) (51) (51)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCCNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQTLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGAU
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA7682.1_ GI_5931705_EMB_CAA7682.1_ GI_5931705_EMB_CAB56603.1_ ENV OF P386700-LINK IRANSLATION OF P0L349-LINK IRCAP-GENOMEA-POLORP TRANSLATION OF INCAP-POL-GENA-GOODA TRANSLATION OF INCAP-POL-GENA-GOODA TRANSLATION OF ORP111-10 PGD-P1 PGD-P2 PGD-P3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EME_CAA76882.1_ GI_4185947_EMB_CAA5682.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240	(1) (1) (298) (301) (298) (272) (4) (1) (51) (51) (51) (51) (51) (51) (51)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIIKLCGNDPDKIVVPTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPTKEQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGAUQFLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGAUQFLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SGAUQFLRIITLCGNDPDKITVPFNKQQVRQAPISGAWQIGLANFYGIIDNHYPKTKIF SGAUQFLRIITLCGNDPDKITVFFNKQQVRQAPISGAWQIGLANFYGIIDNHYPKTKIF SGAUQFLRITWILFKTTREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA7682.1_ GI_4185947_EMB_CAA7682.1_ GI_531705_EMB_CAA7685.1_ GI_531705_EMB_CAA7685.1_ ENV OF AB047240 TRANSLATION OF POIS46TOP-LINK IRCAP-GENOMEA-POLORF TRANSLATION OF LNCAP-GENOMEA-POLORF TRANSLATION OF LNCAP-GENOMEA-POLORF TRANSLATION OF DIA19-LINK LNCAP-GENOMEA-POLORF TRANSLATION OF POIS409-LINK GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76879.1_ GI_4185947_EMB_CAA76882.1_ GI_531705_EMB_CAB5603.1_ ENV OF PA8047240 TRANSLATION OF P386T0P-LINK	(1) (1) (298) (301) (272) (272) (272) (272) (272) (272) (351) (551) (551) (551) (551) (551) (551) (551) (51) (D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOTRLRIIKLCGNDPDKIVVPTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPTKEQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF S161 420 GFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF P38870P-LINK TRANSLATION OF P38870P-LINK LNCAP-GENOMEA-POLORP TRANSLATION OF ORF111-10 PGD-P1 PGD-P2 PGD-P3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76887.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_ ENV OF AB047240 TRANSLATION OF P386700P-LINK	(1) (1) (298) (301) (298) (298) (298) (298) (298) (298) (272) (41) (51) (51) (51) (51) (51) (51) (51) (5	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCCNDPDKIVVPLTKQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKIVVPTNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKIVVPNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKIVVPTNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKIVVPTNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF SIGQ RLRIILCGNDPDKIVVPTNKQQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF SIGQ RLRIILCGNDPDKIVVPTNKQQVRQAFISGAWQIGLANFYGIIDNHYPKTKIF SIGQ RLRIILCGNDPDKIVPTNKQQVRQAFISGAWQIGLANFYGIIDNHYPKTKIF SIGQ RLRIILCGNDPDKIVPTNKQQVRQAFISGAWQIGLANFYGIIDNHYPKTKIF SIGQ RLRIILCGNDPDKIVPTNKQQVRQAFISGAWQIGLANFYGIIDNHYPKTKIF SIGQ FLKITTWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKITWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76879.1_ GI_4185947_EMB_CAA76885.1_ GI_5931705_EMB_CAB56603.1_ ENV OF P386TOP-LINK TRANSLATION OF P386TOP-LINK INCAP-GENOMEA-POLOGP TRANSLATION OF POL349-LINK LNCAP-GENOMEA-POLOF PGD-P1 PGD-P2 PGD-P2 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76885.1_ GI_5931705_EMB_CAB76885.1_ GI_5931705_EMB_CAB76885.1_ GI_5931705_EMB_CAB76885.1_ CONSENSUS CONSENSUS	(1) (1) (298) (301) (298) (272) (4) (1) (51) (57) (17) (1) (1) (301) (358) (361) (358) (311) (358) (311) (358) (311) (358) (311) (358) (311) (358) (311) (31	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQTRLRIIKLCGNDPDKIVVPTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SIGQTLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGQTLRIITLCGNDPDKITVPFNKQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGQTLRIITLCGNDPDKITVPFNKQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGQTLRIITLCGNDPDKITVPFNKQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGQTLRIITLCGNDPDKITVPFNKQVRQPAPISSGAWQIGLANFYGIDNHYPKTKIF SIGQTLRIITLCGNDPDKITVPFNKQVRQPAPISGAWIGLANFYGIDNHYPKTKIF SIGQTLNITTLCGNDPDKITVPFNKQVRQPAPISGAWQIGLANFYGIDNHYPKTKIF SIGQTLANFYGIDNHYPKTKIF SIGQTLANFYGIDNHYPKTKIF SIGQTLANFYGIDNHYPKTKIF SIGQTLANFYGIDNHYPKTKIF SIGQTLANFYGIDNHYPKTKIF SIGQTLANFYGIDNHYPKTKIF SIGQTLANFYGIDNHYFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV SIGLATTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV SIGNGKAAYTGPKERVIKTPYGSAQRAELVAV SIGNGKAAYTGPKERVIKTPYGSAQRAELVAV SIGNGKAAYTGPKERVIKTPYGSAQRAELVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA7682.1_ GI_4185947_EMB_CAA76882.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_ ENV OF AB047240 TRANSLATION OF PO134670-LINK TRANSLATION OF PO1349-LINK INCAP-GENOMEA-POLORF TRANSLATION OF LNCAP-PO1-GENA-GOODA TRANSLATION OF PO1349-LINK DECP1 PGD-P1 PGD-P2 PGDP1 PGD-P2 CONSENSUS GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76882.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAA5885.1_ TRANSLATION OF P386TOP-LINK TRANSLATION OF P386TOP-LINK TRANSLATION OF P01349-LINK	(1) (1) (298) (301) (298) (272) (272) (272) (272) (1) (57) (57) (17) (1) (1) (57) (57) (17) (17) (11) (301) (358) (361) (358) (31) (358) (358) (358) (31) (358) (358) (358) (31) (358) (358) (31) (358) (31) (358) (31) (358) (31) (358) (31) (358) (31) (358) (31) (358) (31) (358) (31) (358) (31) (358) (31) (358) (31) (31) (31) (31) (31) (31) (31) (31	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPTKEOVRQAFINSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKIVVPTKEQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKIVVPTKEQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKIVVPTNKQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKITVPFNKQVRQAFISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQVRQAFISSGAWQIGLANFYGIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQVRQAFISSGAWQIGLANFYGIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQVRQAFISSGAWQIGLANFYGIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVFFNKQVRQAFISGGAWQIGLANFYGIDNHYPKTKIF S161 420 OFLKJTWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKLTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAA76885.1_ ENV OF AB047240 TRANSLATION OF P386T07-LINK INCAP-GENOMEA-POLORF TRANSLATION OF POL349-LINK LNCAP-GENOMEA-POLORF PGD-P1 PGD-P1 PGD-P2 PGDP3 CONSENSUS GI_4185943_EMB_CAA76879.1_ GI_4185943_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ ENV OF AB047240 TRANSLATION OF P386T09-LINK IRANSLATION OF P386T09-LINK IRANSLATION OF P01349-LINK INCAP-GENOMEA-POLORF TRANSLATION OF POL349-LINK INCAP-GENOMEA-POLORF	(1) (1) (298) (298) (298) (298) (298) (298) (298) (298) (298) (298) (298) (298) (298) (298) (298) (10) (51) (51) (51) (51) (51) (51) (51) (51	D LAPLQLLIPATAHS TGIIIQNTDLVEWSPLPHSTVKTFTLYLDQMAT 301 360 LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGQTRLRIIKLCGNDPDKIVVPLTKEQVRQAPISGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCCNDPDKIVVPLTKQVRQAPISGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKIVVPTNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGQGRLRIITLCGNDPDKIVVPFNKQQVRQAPISGAWQIGLANFYGIIDNHYPKTKIF STATANA S61 420 OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV OFLKUTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAELVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA7682.1_ GI_5931705_EMB_CAA7682.1_ GI_5931705_EMB_CAB56603.1_ ENV OF P386TOP-LINK TRANSLATION OF P386TOP-LINK INCAP-GENOMEA-POLOGP TRANSLATION OF P01349-LINK LNCAP-GENOMEA-POLOGP TRANSLATION OF LNCAP-POL-GENA-GOODA TRANSLATION OF P01349-LINK ENCAP-GENOMEA-POLOGP PGD-P1 PGD-P2 PGD-P2 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_ TRANSLATION OF P0386TOP-LINK TRANSLATION OF P0386TOP-LINK INCAP-GENOMEA-POLORF TRANSLATION OF ORP111-10 PGD-P1	(1) (1) (298) (301) (298) (272) (4) (1) (51) (57) (17) (17) (17) (13) (358) (361) (358) (361) (358) (364) (358) (358) (358) (358) (111) (111) (111) (117)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKIVVPTKEOVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SIGORLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGORLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGORLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGORLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPTTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV SINGKAAYTGPKERVIKTPYGSAQRAELVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAA76885.1_ ENV OF AB047240 TRANSLATION OF P386T07-LINK INCAP-GENOMEA-POLORF TRANSLATION OF POL349-LINK LNCAP-GENOMEA-POLORF PGD-P1 PGD-P1 PGD-P2 PGDP3 CONSENSUS GI_4185943_EMB_CAA76879.1_ GI_4185943_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ GI_5931705_EMB_CAA76882.1_ ENV OF AB047240 TRANSLATION OF P386T09-LINK IRANSLATION OF P386T09-LINK IRANSLATION OF P01349-LINK INCAP-GENOMEA-POLORF TRANSLATION OF POL349-LINK INCAP-GENOMEA-POLORF	(1) (1) (298) (298) (221) (298) (228) (228) (228) (272) (4) (1) (51) (51) (51) (51) (51) (51) (51)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKIVVPTKEQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKIVVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF S161 420 OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV OFLKJTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYQSAQRAEUVAV
PGD-P2 PGDP3 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA7682.1_ GI_5931705_EMB_CAA7682.1_ GI_5931705_EMB_CAB56603.1_ ENV OF P386TOP-LINK TRANSLATION OF P386TOP-LINK INCAP-GENOMEA-POLOGP TRANSLATION OF P01349-LINK LNCAP-GENOMEA-POLOGP TRANSLATION OF LNCAP-POL-GENA-GOODA TRANSLATION OF P01349-LINK ENCAP-GENOMEA-POLOGP PGD-P1 PGD-P2 PGD-P2 CONSENSUS GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_ GI_5931705_EMB_CAA76885.1_ TRANSLATION OF P0386TOP-LINK TRANSLATION OF P0386TOP-LINK INCAP-GENOMEA-POLORF TRANSLATION OF ORP111-10 PGD-P1	(1) (1) (298) (298) (221) (298) (228) (228) (228) (272) (4) (1) (51) (51) (51) (51) (51) (51) (51)	D LAPLQLLIPATAHS TGIIIQNTDLVEWSFLPHSTVKTFTLYLDQMAT 301 360 LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWKIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPLTKEQVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOTRLRIIKLCGNDPDKIVVPTKEOVRQAPINSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKIVVPTKEOVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIP LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF LIGOGRLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIIDNHYPKTKIF SIGORLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGORLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGORLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGORLRIITLCGNDPDKITVPFNKQQVRQAPISSGAWQIGLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPKTKIF SIGORLANFYGIDNHYPTTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV OFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYGSAQRAELVAV SINGKAAYTGPKERVIKTPYGSAQRAELVAV

FIGURE 8 CONTD....

		421 480
GI 4185939_EMB CAA76879.1_	(418)	ITVLODFDOPINIISDSAYVVOATRDVETALIKYSMDDQLNOLFNLLOOTVRKRNFPFYI
GI_4185943_EMB_CAA76882.1_	(421)	ITVLQDFDQPINIISDSAYVVQATRDVETALIKYSMDDQLNQLFNLLQQTVRKRNFPFYI
GI_4185947_EMB_CAA76885.1_	(418)	ITVLQDFDQPINIISDSAYVVQATRDVETALIKYSMDDQLNQLPNLLQQTVRKRNFPFYI
GI 5931705 EMB CAB56603.1_	(392)	ITVLQDFDQPINIISDSAYVVQATRDVETALIKYSMDDQLNQLFNLLQQTVRKRNFPFYI
ENV OF AB047240	(124)	ITVLQDFDQPINIISDSAYVVQATRDVBTALIKYSTDDHLNQLFNLLQQTVRKRNFPFYI
TRANSLATION OF P386TOP-LINK	(31)	
TRANSLATION OF POL349-LINK	(28)	***************************************
LNCAP-GENOMEA-POLORF	(171)	ITVLQDFDQPINIISDSAYVVQATRDVETALIKYSTDDHLNQLFNLLQQTVRKRNFPFYI
TRANSLATION OF LNCAP-POL-GENA-GOODA	(171)	ITVLQDFDQPINIISDSAYVVQATRDVETALIKYSTDDHLNQLFNLLQQTVRKRNFPFYI
TRANSLATION OF ORF111-10	(177)	ITVLQDFDQPINIISDSAYVVQATRDVETALIKYSTDDHLNQLFNLLQQTVRKRNFPFYI
PGD-P1	(17)	
PGD-P2	(17)	***************************************
PGDP3	(1)	
CONSENSUS	(421)	ITVLQDFDQPINIISDSAYVVQATRDVETALIKYS DD LNQLFNLLQQTVRKRNFPFYI

		481 540
GI_4185939_EMB_CAA76879.1_	(478)	THIRAHTNLPGPLTKANEOADLLVSSALIKAOBLHALTHVNAAGLKNKFDVTWKOAKDIV
GI_4185943_EMB_CAA76882.1_	(481)	THIRAHTNLPGPLTKANEQADLLVSSALIKAQELHALTHVNVAGLKNKFDVTWKQAKDIV
GI_4185947_EMB_CAA76885.1_	(478)	THIRAHTNLPGPLTKANEQADLLVSSALIKAQELHALAHVNAAGLKNKFDVTWKQAKDIV
GI_5931705_EMB_CAB56603.1_	(452)	THIRAHTNLPGPLTKANEQADLLVSSAFIKAQELHALTHVNAAGLKNKFDVTWKQAKDIV
ENV OF AB047240	(184)	THIRAHTNLPGPLTKANEQADLLVSSAFIKAQELLALTHVNAAGLKNKFDVTWKQAKDIV
TRANSLATION OF P386TOP-LINK	(31)	
TRANSLATION OF POL349-LINK	(28)	
LNCAP-GENOMEA-POLORF	(231)	THIRAHTNLPGPLTKANEQADLLVSSAFIKAQELLALTHVNAAGLKNKFDVTWKQAKDIV
TRANSLATION OF LNCAP-POL-GENA-GOODA	(231)	THIRAHTNLPGPLTKANEQADLLVSSAFIKAQELLALTHVNAAGLKNKFDVTWKQAKDIV
TRANSLATION OF ORF111-10	(237)	THIRAHTNLPGPLTKANEQADLLVSSAFIKAQELLALTHVNAAGLKNKFDVTWKQAKDIV
PGD-P1	(17)	
PGD-P2	(17)	
PGDP3	(1)	
CONSENSUS	(481)	THIRAHTNLPGPLTKANEQADLLVSSA IKAQEL ALTHVNAAGLKNKFDVTWKQAKDIV

600

	541 600
	541 600
(538)	OHCTOCOVLHLPTOEAGVNPRGLCPNALWOMDVTHVPSFGRLSYVHVTVDTYSHFIWATC
(541)	OHCTOCOVLHLPTOEAGVNPRGLCPNALWOMDVTHVSSFGRLSYLHVTVDTYSHFIWATC
(538)	QHCTQCQVLHLPTQEAGVNPRGLCPNALWQMDVTHVPSFGRLSYVHVTVDTYSHFIWATC
(512)	QHCTQCQVLDLPTQBAGVNPEVCVLMHYGKWMSHMYLHLGRLSYVHVTVDTYSHFMCATC
(244)	QHCTQCQVLHLSTQEAGVNPRGLCPNALWQMDGTHVPSFGRLSYVHVTVDTYSHFIWATC
(31)	
(28)	
(291)	OHCTOCOVLHLSTOEAGVNPRGLCPNALWOMDGTHVPSFGRLSYVHVTVDTYSHFIWATC
(291)	OHCTOCOVLHLSTOEAGVNPRGLCPNALWOMDGTHVPSFGRLSYVHVTVDTYSHFIWATC
(297)	QHCTQCQVLHLSTQEAGVNPRGLCPNALWQMDGTHVPSFGRLSYVHVTVDTYSHFIWATC
(17)	
(17)	
(1)	
(541)	QHCTQCQVLHL TQEAGVNPRGLCPNALWQMD THV SFGRLSYVHVTVDTYSHFIWATC

	601 660
(598)	QTGESTSHVKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFQKFLSQWKISHTTGIPYNSQG
(601)	OTGESTSHVKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFOKFLSOWKISHTTGIPYNSOG
(598)	OTGESTSHVKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFOKFLSOWKISHTTGIPYNSOG
(572)	QTGESTSHVKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFQKFLSQWKISHTTGIPYNSQG
(304)	QTGESTSHVKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFQKFLSQWKISHTGIPYNSQG
(31)	
(28)	
(351)	OTGESTSHVKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFOKFLSOWKISHTTGIPYNSOG
(351)	OTGESTSHVKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFOKFLSOWKISHTTGIPYNSOG
(357)	QTGESTSHAKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFQKFLSQWKISHTTGIPYNSQG
(17)	
(17)	
(1)	
(601)	QTGESTSHVKKHLLSCFAVMGVPBKIKTDNGPGYCSKAFQKFLSQWKISHTTGIPYNSQG

	661 720
(658)	QAIVERTNRTLKTQLVKQKEGGDSKECTTPOMQLNLALYTINFLNIYRNQTTTSAEQHLT
(661)	QAIVERINRTLKTQLVKQKEGGDSKECTTPQMQLNLALYTIMPLNIYRNQTTTSAEQHLT
(658)	OAIVERTNRTLKTOLVKOKEGGDSKECTTPOMOLNLALYTINPLNIYRNOTTTSAEQHLT
(632)	OAIVERTNRTLKTOLVKOKEGGDSKECTTPOMOLNLALYTLNPLNIYRNOTTTSAE-HLT
(364)	QAIVERTNRTLKTQLVKQKEGGDSKECTTPQMQLNLALYTINPLNIYRNQTTTSAKQHLT
(31)	
(28)	
(411)	QAIVERTNRTLKTQLVKQKEGGDSKECTTPQMQLNLALYTLNPLNIYRNQTTTSAKQHLT
(411)	OAIVERTNRTLKTOLVKOKEGGDSKECTTPOMOLNLALYTLNFLNIYRNOTTTSAKOHLT
(417)	QAIVERTNRTLKTQLVKQKEGGDSKECTTPQMQLNLALYTLNFLNIYRNQTTTSAKQHLT
(17)	
(17)	***************************************
(1)	HLT
(661)	QAIVERTNRTLKTQLVKQKEGGDSKECTTPQMQLNLALYTLNFLNIYRNQTTTSA QHLT

GI_4185939_EMB_CAA76879.1 GI_4185943_EMB_CAA76882.1 GI_4185947_EMB_CAA76885.1 GI_5931705_EMB_CAB56603.1 ENV OF AB047240 TRANSLATION OF P386TOP-LINK TRANSLATION OF P01349-LINK LNCAP-GENOMEA-P0LORF TRANSLATION OF LNCAP-P0L-GENA-GOODA TRANSLATION OF ORF111-10 PDD-P1

GI_4185939_EMB_CAA76879.1_

GI_4185939_EMB_CAA76879.1 GI_4185943_EMB_CAA76882.1 GI_4185947_EMB_CAA76885.1 GI_5931705_EMB_CAB7685.1 ENV OF AB647240 TRANSLATION OF P386TOF-LINK TRANSLATION OF P0L349-LINK LNCAP-GENOMEA-P0LORF TRANSLATION OF LNCAP-FOL-GENA-GOODA TRANSLATION OF ORF111-10 PEGD-P1

GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76882.1_ GI_4185947_EMB_CAA76885.1 GI_5931705_EMB_CAB56603.1_ ENV OF AB047240 TRANSLATION OF POL349-LINK TRANSLATION OF POL349-LINK LANCAP-GENOMEA-POLORF TRANSLATION OF LNCAP-POL-GENA-GOODA TRANSLATION OF ORF111-10 FGD-P1

- - - PGD-P1 PGD-P2

PGD-P1 PGD-P2 PGDP3 CONSENSUS

PGD-P1 PGD-P2 PGDP3 CONSENSUS

- PGDP3
- CONSENSUS

		721 500
GI_4185939_EMB_CAA76879.1_	(710)	721 GKKNSPHEGKLIWWKDSKNKTWEIGKVITWGRGFACVSPGENQLPVWIPTRHLKPYNEPI
GI_4185955_EMB_CAA768879.1_ GI_4185943_EMB_CAA76882.1_		GKKNSPHEGKLIWWKDSKNKIWEIGKVIIWGRGFACVSPGENQLPVWIPIRHLKFYNEPI GKKNSPHEGKLIWWKDNKNKTWEIGKVIIWGRGFACVSPGENQLPVWIPIRHLKFYNEPI
GI 4185947 EMB CAA76885.1		GKKNSPHEGKLIWWKDNKNKIWEIGKVIIWGRGFACVSPGENQLPVWIPTHLKFYNEPI
GI_5931705_EMB_CAB56603.1_	(691)	
ENV OF AB047240	(424)	
TRANSLATION OF P386TOP-LINK	(31)	
TRANSLATION OF POL349-LINK	(28)	
LNCAP-GENOMEA-POLORF		GKKHSPHEGKLIWWKDNKNKTWEIGKVITWGRGFACVSPGENQLPVWIPTRHLKFYNEPI
TRANSLATION OF LNCAP-POL-GENA-GOODA	(471)	GKKHSPHEGKLIWWKDNKNKTWEIGKVITWGRGFACVSPGENQLPVWIPTRHLKFYNEPI
TRANSLATION OF ORF111-10	(477)	GKKHSPHEGKLIWWKDNKNKTWEIGKVITWGRGFACVSPGENQLPVWIPTRHLKFYNEPI
PGD-P1	(17)	
PGD-P2	(17)	
PGDP3	(4)	GKKNSPHEGKLIC
CONSENSUS		GKK SPHEGKLIWWKD KNKTWEIGKVITWGRGFACVSPGENQLPVWIPTRHLKFYNEPI
		781 840
GI_4185939_EMB_CAA76879.1_	(778)	RDAKKSTSAETETS
GI_4185943_EMB_CAA76882.1_	(781)	GDAKKSTSAETETP
GI_4185947_EMB_CAA76885.1_	(778)	RDAKKSTSAETETS
GI_5931705_EMB_CAB56603.1_	(703)	
ENV OF AB047240	(484)	GDAKKRASTEMVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIVYRYP
TRANSLATION OF P386TOP-LINK	(31)	
TRANSLATION OF POL349-LINK	(28)	
LNCAP-GENOMEA-POLORF	(531)	GDAKKRASTEMVTFVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIVYRYP
TRANSLATION OF LNCAP-POL-GENA-GOODA	(531)	GDAKKRASTEMVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIVYRYP
TRANSLATION OF ORF111-10	(537)	GDAKKRASTEMVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIVYRYP
PGD-P1	(17)	
PGD-P2	(17)	
PGDP3	(17)	
CONSENSUS	(781)	DAKK SE T
		841 900
GI_4185939_EMB_CAA76879.1_		
GI_4185943_EMB_CAA76882.1_	1 . 1	
GI_4185947_EMB_CAA76885.1_	(792)	
GI_5931705_EMB_CAB56603.1_		
ENV OF AB047240		PICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDPSYQRSLKFR
TRANSLATION OF P386TOP-LINK	(31)	
TRANSLATION OF POL349-LINK	(28)	
LNCAP-GENOMEA-POLORF		PICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSLKFR
TRANSLATION OF LNCAP-POL-GENA-GOODA		PICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDPSYQRSLKFR
TRANSLATION OF ORF111-10		PICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDPSYQRSLKPR
PGD-P1		
PGD-P2 PGDP3	(17)	
CONSENSUS	(841)	
CONSENSOS	(041)	
		901 960
GI_4185939_EMB_CAA76879.1_	(792)	QSSTVDSQDEQNGDVRRTDEVALH
GI_4185943_EMB_CAA76882.1_	(795)	QSSTVDSQDEQNGDVRRTDEVALH
GI 4185947 EMB CAA76885.1	(792)	QSSTVDSQDEONGDVRRTDEVAIH
GI 5931705 EMB CAB56603.1	(703)	
ENV OF AB047240	(604)	PKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEFGTEIDWAPRGQFYHNCSGQTQS
TRANSLATION OF P386TOP-LINK	(31)	
TRANSLATION OF POL349-LINK	(28)	
LNCAP-GENOMEA-POLORF		PKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEPGTTIDWAPRGQFYHNCSGQTQS
TRANSLATION OF LNCAP-POL-GENA-GOODA		PKGKPCPKEIPKESKNTEVLVWEECVANSAVILONNEFGTIIDWAPRGQFYHNCSGOTOS
TRANSLATION OF ORF111-10		PKGKPCPKEIPKESKNTEVLVWEECVANSAVILONNEFGTLIDWAPRGOFYHNCSGOTOS
PGD-P1		
PGD-P2	(17)	
PGDP3	(17)	
CONSENSUS	(901)	TI
		Ac
di 4105020 BBD 03306080 1	(01 ()	961 1020 QEGRAANLGTTKEADAVSYKISREHKGDTNPREVAACSLDDCINGGKSPYACRSSCS
GI_4185939_EMB_CAA76879.1_ GI_4185943_EMB_CAA76879.1_		QEGRAANLIJTTKEADAVSYKISKEHKGDTNPREYAACSIDDCINGGKSPIACRSSCS QESRAADLIGTTKEADAVSYKISREHKGDTNPREYAACSIDDCINGGKSPYACRSSCS
GI_4185943_EMB_CAA76882.1_ GI 4185947 EMB CAA76885.1		QEGRAANLGTTKEADAVSYKISREHKGDTNPREYAACGIDDCINGGRSPHACRSSCS QEGRAANLGTTKEADAVSYKISREHKGDTNPREYAACSLDDCINGGRSPHACRSSCS
	(202)	VEGRAMMING ITAKADAVSIAISARANGUTMPREXARCSUDUCINGGASPIACRSSUS
GI_5931705_EMB_CAB56603.1_ ENV OF AB047240	(664)	CPSAQVSPAVDSpltesldkhkhkklosfypwewdekgistprpeiispvsgpehpelwr
TRANSLATION OF P386TOP-LINK	(204)	CESAQUSEAUDSUDTESSIDAMAAAADQSFIPHEMQEAGISIFAEBIIGE VAGEBAFEDWR
TRANSLATION OF PS8810F-LINK TRANSLATION OF POL349-LINK	(28)	
LNCAP-GENOMEA-POLORF	(711)	CPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEMGEKGISTPRPEIISPVSGP
TRANSLATION OF LNCAP-POL-GENA-GOODA	(713)	CPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIISPVSGPEHPELWR
TRANSLATION OF ORF111-10	(717)	CPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIISPVSGPEHPELWR
PGD-P1	(17)	
PGD-P2		
	· (17)	
PGDP3	·(17)	
	·(17)	

FIGURE 8 CONTD...

		1021	1035
GI_4185939_EMB_CAA76879.1_	(873)		
GI_4185943_EMB_CAA76882.1_	(876)		
GI_4185947_EMB_CAA76885.1_	(873)		
GI_5931705_EMB_CAB56603.1_	(703)		
ENV OF AB047240	(724)	LWPDTTLER	GLEIKL
TRANSLATION OF P386TOP-LINK	(31)		
TRANSLATION OF POL349-LINK	(28)		
LNCAP-GENOMEA-POLORF	(764)		
TRANSLATION OF LNCAP-POL-GENA-GOODA	(771)	LWPDTTLEI	GLEIKL
TRANSLATION OF ORF111-10	(777)	LWPDTTLEI	GLEIKL
PGD-P1	(17)		
PGD-P2	(17)		
PGDP3	· (17)		
CONSENSUS	(1021)		

FIGURE 9

		1 60
GI 4185940 EMB CAA76880.1	(1)	
GI 4185944 EMB CAA76883.1	(1)	
GI_4185948 EMB_CAA76886.1	(1)	
GI 5931706 EMB CAB56604.1	(1)	
ENV OF AB047240	(1)	MATLIGOGRLRIITLCGNDPDKITVPFNKOOVROAFISSGAWOIGLANFLGIIDNHYPKT
TRANSLATION OF E207TOP-LINK	(1)	MAIDIOQGADATITICONDEDATIVETMAQQVAQAFIBSGAMQIGDAAFIBGIIDAATEKI
TRANSLATION OF ENV287-LINK	(1)	
TRANSLATION OF 120.22A-23	(1)	
PGD-E1	(1)	
PGD-E1 PGD-E2	(1)	
PGD-E2 PGD-E3	(1)	
CONSENSUS	(1)	
		61 120
		51 120
GI_4185940_EMB_CAA76880.1	(1)	
GI_4185944 EMB CAA76883.1	(1)	***************************************
GI 4185948 EMB CAA76886.1	(1)	
GI_5931706_EMB_CAB56604.1	(1)	
ENV OF AB047240	(1)	KIFQFLKLTTWILPKITRREPLENALTVFTDGSSNGKAAYTGPKERVIKTPYOSAORAEL
TRANSLATION OF E207TOP-LINK	(01)	KIFQFBKBIIWIDFKIIRKEPBENADIVFIDGSSNGKAAITGPKERVIKIPIQSAQKAKL
TRANSLATION OF E20710P-LINK	(1)	
TRANSLATION OF 120.22A-23	(1)	
PGD-E1	(1)	
PGD-E1 PGD-E2	(1)	***************************************
PGD-E2 PGD-E3	(1)	
	• - •	
CONSENSUS	(61)	
		121 180
GI_4185940_EMB_CAA76880.1	(1)	
GI_4185944 EMB CAA76883.1	(1)	
GI_4185948 EMB CAA76886.1	(1)	
GI_5931706 EMB CAB56604.1	(1)	
ENV OF AB047240	(121)	VAVITVLQDFDQPINIISDSAYVVQATRDVETALIKYSTDDHLNQLFNLLQQTVRKRNFP
TRANSLATION OF E207TOP-LINK	(121)	
TRANSLATION OF E20710P-LINK	(1)	
TRANSLATION OF ENV287-DINK TRANSLATION OF T20.22A-23	(1)	
PGD-E1	(1)	
PGD-E2	(1)	
PGD-E3	(1)	
CONSENSUS	(121)	

		181 240
GI_4185940_EMB_CAA76880.1_	(1)	181 240
GI_4185944_EMB_CAA76883.1_	(1)	
GI_4185948_EMB_CAA76886.1_ GI_5931706_EMB_CAB56604.1_	(1)	
ENV OF AB047240	(181)	FYITHIRAHTNLPGPLTKANEQADLLVSSAFIKAQELLALTHVNAAGLKNKFDVTWKQAK
TRANSLATION OF E207TOP-LINK	(1)	***************************************
TRANSLATION OF ENV287-LINK	(1)	
TRANSLATION OF T20.22A-23	(1)	
PGD-E1	(1)	
PGD-E2	(1)	
PGD-E3 CONSENSUS	(1) (181)	
CONSENDUS	(101)	
		241 300
GI_4185940_EMB_CAA76880.1_		
GI_4185944_EMB_CAA76883.1_	(1) (1)	
GI_4185948_EMB_CAA76886.1_ GI 5931706 EMB CAB56604.1	(1)	
ENV OF AB047240	(241)	DIVQHCTQCQVLHLSTQBAGVNPRGLCPNALWQMDGTHVPSFGRLSYVHVTVDTYSHFIW
TRANSLATION OF E207TOP-LINK	(1)	
TRANSLATION OF ENV287-LINK	(1)	·····
TRANSLATION OF T20.22A-23	(1)	***************************************
PGD-E1 PGD-E2	(1) (1)	
PGD-E3	(1)	
CONSENSUS	(241)	
CT 4195940 PMP CNA76900 1	(1)	301 360
GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_	(1)	
GI_4185948_EMB_CAA76886.1_	(1)	
GI_5931706_EMB_CAB56604.1	(1)	
ENV OF AB047240	(301)	ATCQTGESTSHVKKHLLSCFAVMGVPEKIKTDNGPGYCSKAFQKFLSQWKISHTTGIPYN
TRANSLATION OF E207TOP-LINK	(1)	
TRANSLATION OF ENV287-LINK	(1)	
TRANSLATION OF T20.22A-23 PGD-E1	(1)	
PGD-E2	(1)	
PGD-E3	(1)	
CONSENSUS	(301)	
		361 420
GI_4185940_EMB_CAA76880.1_	(1)	361 420
GI_4185944_EMB_CAA76883.1_	(1)	MQRKAPPRRRHRNRAPLTHKMNKMVTSEEQMKL
GI_4185948_EMB_CAA76886.1_	(1)	MQRKAPPRRRHRNRAPLTHKMNKMVTSEEQMKL
GI_5931706_EMB_CAB56604.1_	(1)	
ENV OF AB047240	(361)	SQGQAIVERTNRTLKTQLVKQKEGGDSKECTTPQMQLNLALYTLNFLNIYRNQTTTSAKQ
TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK	(1)	
TRANSLATION OF ENV287-BINK TRANSLATION OF T20.22A-23	(1)	MNPSEMQRKAPPRRRHRNRAPLTHKMNKMVTSEEQMKL
PGD-E1	(1)	
PGD-E2	(1)	
PGD-E3	(1)	
CONSENSUS	(361)	
		421 480
GI_4185940_EMB_CAA76880.1_	(35)	PSTKKAEPPTWAQLKKLTQLATKYLENTKVTQTPESMLLAALMIVSMVVSLPMPAGAAAA
GI_4185944_EMB_CAA76883.1_		PSTKKAEPPTWAQLKKLTQLATKYLENTKVTQTPESMLLAALMIVSMVVSLPMPAGAAAA
GI_4185948_EMB_CAA76886.1_		PSTKKAEPPTWAQLKKLTQLATKYLENTKVTQTPESMLLAALMIVSMVVSLPMPAGAAAA
GI_5931706_EMB_CAB56604.1_ ENV OF AB047240	•-•	HLTGKKHSPHEGKLIWWKDNKNKTWEIGKVITWGRGFACVSPGENQLPVWIPTRHLKFYN
TRANSLATION OF E207TOP-LINK		HEIGKKHSPHEGKBIWWKDNKNKIWEIGKVIIWGKGFACVSPGENQEPVWIPTRHEKFYN
TRANSLATION OF ENV287-LINK		
TRANSLATION OF T20.22A-23	(40)	PSTKKAEPPTWAQLKKLTQLATKYLENTKVTQTPESMLLAALMIVSMVVSLPMPAGAAAA
PGD-E1		
PGD-E2		
PGD-E3 CONSENSUS		
CONSENSOS	(421)	

FIGURE 9 CONTD...

		481 540
GI_4185940 EMB_CAA76880.1_	(95)	NYTYWAYVPFPP-LIRAVTWMDNPTEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGY
GI_4185944 EMB_CAA76883.1_		NYTYWAYVPFPP-LIRAVTWMDNPIEVYVNDSVWVPGPTDDHCPAKPEEEGMMINISIGY
GI_4185948_EMB_CAA76886.1_	(95)	NYTYWAYVPFPP-LIRAVTWMDNPTEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGY
GI_5931706 EMB_CAB56604.1	(1)	MVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIGY
ENV OF AB047240		EPIGDAKKRASTEMVTPVTWMDNPIEVYVNDSVWVPGPTDDRCPAKPEEEGMMINISIVY
TRANSLATION OF E207TOP-LINK	(1)	
TRANSLATION OF ENV287-LINK	(1)	
TRANSLATION OF T20.22A-23	(100)	NYTYWAYVPFPP-LTRAVTWMDNPTEVYVNDSVWVPGPIDDRCPAKPEEEGMMINISIGY
PGD-E1	(1)	
PGD-E2	(1)	
PGD-E3	(1)	
CONSENSUS	(481)	LI VTWMDNP EVYVNDSVWVPGP DD CPAKPEEEGMMINISI Y
		544
		541 600
GI_4185940_EMB_CAA76880.1_	(154)	HYPPICLGRAPGCLMPAVQNWLVEVPTVSPICRFTYHMVSGMSLRPRVNYLQDFSYQRSL
GI_4185944 EMB_CAA76883.1_	(154)	RYPPICLGRAPGCLMPAVQNWLVEVPTVSPISRFTYHMVSGMSLRPRVNYLQDFSYQRSF
GI_4185948_EMB_CAA76886.1_		HYPPICLGRAPGCIMPAVQNWLVEVPTVSPICRFTYHMVSGMSLRPRVNYLQDFSYQRSL
GI_5931706 EMB CAB56604.1	(48)	HYPPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNCLQDFSYQRSL
		RYPPICLGRAPGCLMPAVQNWLVEVPTVSPNSRFTYHMVSGMSLRPRVNYLQDFSYQRSL
TRANSLATION OF E207TOP-LINK	(1)	FSYQRSL
TRANSLATION OF ENV287-LINK	(1)	
TRANSLATION OF T20.22A-23		HYPPICLGRAPGCEMPAVONWLVEVPTVSPICRFTYHMVSGMSLRPRVNYLODFSYORSL
PGD-E1	(1)	
PGD-E2	(1)	
PGD-E3	(1)	**
CONSENSUS	(541)	YPPICLGRAPGCLMPAVQNWLVEVPTVSP RFTYHMVSGMSLRPRVN LQDFSYQRSL
		601 660
GI 4185940 EMB CAA76880.1	(214)	KFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEFGTIIDWAPRGQFYHNCSGO
GI 4185944 EMB CAA76883.1		KFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEFGTIIDWAPRGOFYHNCSGO
GI_4185948_EMB_CAA76886.1_	(214)	KFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEFGTIIDWAPRGQFYHNCSGQ
GI 5931706 EMB CAB56604.1	(108)	KFRPKGKTCPKEIPKGSKNTEVLVWEECVANSVVILONNEFGTIIDWAPRGOFYHNCSGO
ENV OF AB047240	. (601)	KFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEFGTIIDWAPRGQFYHNCSGQ
TRANSLATION OF E207TOP-LINK		KFRPKGKPCPKEIPKESKNTEVL
TRANSLATION OF ENV287-LINK	(1)	
TRANSLATION OF T20.22A-23		KFRPKGKPCPKEIPKESKNTEVLVWEECVANSAVILQNNEFGTIIDWAPRGQFYHNCSGO
PGD-E1	(1)	RPKGKPCPKEIPKESC
PGD-E2	(1)	
PGD-E3	(1)	
	• - •	
CONSENSUS	(601)	KFRPKGKPCPKEIPKESKNTEVLVWEECVANS VILQNNEFGTIIDWAPRGQFYHNCSGQ
		661 720
	(0.0.4.)	
GI_4185940_EMB_CAA76880.1_		TQSCQSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIVSPVSGPEHPE
GI 4185944 EMB CAA76883.1	(274)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKITSPVSGPEHPE
GI 4185948 EMB CAA76886.1	(274)	
GT 4103340 700 CMM/000001		
		TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYLWEWEEKGISTPRPKITSPVSGPEHPE
GI_5931706_EMB_CAB56604.1_	(168)	100ctong and and 10000110000110010110101010101010101010
ENV OF AB047240	(661)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEITSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK	(661) (31)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI
ENV OF AB047240	(661) (31)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEITSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK	(661) (31) (1)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGI
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23	(661) (31) (1) (279)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI SPVSGPEHPB
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1	(661) (31) (1) (279) (17)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIËSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23	(661) (31) (1) (279)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI SPVSGPEHPB
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2	(661) (31) (1) (279) (17) (1)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIËSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3	(661) (31) (1) (279) (17) (1) (1)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI SPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2	(661) (31) (1) (279) (17) (1) (1)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIËSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3	(661) (31) (1) (279) (17) (1) (1)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIËSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3	(661) (31) (1) (279) (17) (1) (1)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIËSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3	(661) (31) (1) (279) (17) (1) (1)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI SPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3	(661) (31) (1) (279) (17) (1) (1)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS	(661) (31) (1) (279) (17) (1) (1) (661)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGI SDLTESLDKHKHKKLQSFYPWEWGEKGI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKINSPVSGPEHPE TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE 721 780
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS	(661) (31) (279) (17) (1) (1) (661) (334)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIESPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS	(661) (31) (279) (17) (1) (1) (661) (334)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGI SDLTESLDKHKHKKLQSFYPWEWGEKGI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKINSPVSGPEHPE TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE 721 780
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS	(661) (31) (279) (17) (1) (1) (661) (334) (334)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEILSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_ GI_4185948_EMB_CAA76886.1_	(661) (31) (1) (279) (17) (1) (1) (661) (334) (334) (334)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEILSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_ GI_5931706_EMB_CAB56604.1_	(661) (31) (279) (17) (1) (1) (661) (334) (334) (334)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEIDSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_ GI_4185948_EMB_CAA76886.1_	(661) (31) (279) (17) (11) (11) (661) (334) (334) (334) (228) (721)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEILSPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76880.1_ GI_4185948_EMB_CAA76886.1_ GI_5931706_EMB_CAB56604.1_ ENV OF AB047240	(661) (31) (279) (17) (1) (1) (661) (334) (334) (228) (721) (31)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SPVSGPEHPE
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_ GI_4185948_EMB_CAA76886.1_ GI_5931706_EMB_CAB56604.1_ ENV OF AB047240 TRANSLATION OF E207TOP-LINK	(661) (31) (279) (17) (1) (1) (661) (334) (334) (228) (721) (31)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SPVSGPEHPE
GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76880.1_ GI_4185944_EMB_CAA76880.1_ GI_531706_EMB_CAA76886.1_ GI_531706_EMB_CAA76886.1_ GI_531706_EMB_CAA76886.1_ GI_531706_EMB_CAA76886.1_ GI_531706_EMB_CAA76886.1_ GI_531706_EMB_CAA76886.1_ GI_531706_EMB_CAA76886.1_ GI_531706_EMB_CAA76886.1_ GI_531706_EMB_CAA76886.1_ ENV OF AB047240 TRANSLATION OF E207TOP-LINK	(661) (31) (279) (17) (1) (1) (661) (334) (334) (228) (721) (31) (29)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE 721 780 LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTVDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSLTVPLQSCVKPPYMLVVGNIVIKP
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_ GI_4185948_EMB_CAA76886.1_ GI_5931706_EMB_CAB56604.1_ ENV OF AB047240 TRANSLATION OF E207TOP-LINK	(661) (31) (279) (17) (1) (1) (661) (334) (334) (334) (228) (721) (31) (29) (339)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE 721 780 LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP
GI_4185940_EMB_CAA76880.1_ GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76880.1_ GI_4185948_EMB_CAA76880.1_ GI_5931706_EMB_CAA76880.1_ GI_5931706_EMB_CAB56604.1_ ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF T20.22A-23	(661) (31) (12) (279) (17) (11) (11) (661) (334) (334) (334) (228) (721) (31) (29) (339) (37)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIÑSPVSGPEHPE SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIÑSPVSGPEHPE TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE 721 780 LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_ GI_5931706_EMB_CAB56604.1_ ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1	(661) (31) (12) (279) (17) (11) (11) (661) (334) (334) (334) (228) (721) (31) (29) (339) (37)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIÑSPVSGPEHPE SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIÑSPVSGPEHPE TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE 721 780 LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_ GI_4185944_EMB_CAA76883.1_ GI_5931706_EMB_CAB56604.1_ ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2	(661) (31) (12) (279) (17) (11) (11) (661) (334) (334) (334) (228) (721) (31) (29) (339) (37)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIÑSPVSGPEHPE SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKIÑSPVSGPEHPE TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE 721 780 LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRÍWSGNQTLETRDRKPFYTIDLNSSITVPLQSCVKPPYMLVVGNIVIKP
ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1 PGD-E2 PGD-E3 CONSENSUS GI_4185940_EMB_CAA76880.1_ GI_4185944_EMB_CAA76883.1_ GI_5931706_EMB_CAB56604.1_ ENV OF AB047240 TRANSLATION OF E207TOP-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23 PGD-E1	(661) (31) (12) (279) (17) (11) (11) (661) (334) (334) (334) (228) (721) (31) (29) (339) (37)	TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPEI SDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI TQSCPSAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRPKI TQSC SAQVSPAVDSDLTESLDKHKHKKLQSFYPWEWGEKGISTPRP IISPVSGPEHPE 721 780 LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP LWRLTVASHHIRTWSGNQTLETRDRKPFYTIDLNSSTVPLQSCVKPPYMLVVGNIVIKP

FIGURE 9 CONTD

•

CT 4105040 EMB C3376000 1	(204)	
GI_4185940_EMB_CAA76880.1_ GI 4185944 EMB CAA76883.1		DSQTITCENCRLLTCIDSTFNWOHRILLVRAREGVWIPVSMDRPWEASPSVHILTEVLKG DSQTITCENCRLLTCIDSTFNWOHRILLVRAREGVWIPVSMDRPWETSPSIHTLTEVLKG
GI 4185948 EMB CAA76886.1		DSQTITCENCRLLTCIDSTFNWQHRILLVRAREGVWIFVSMDRPWEISFSIHILTEVLKG
GI 5931706 EMB CAB56604.1		ASQTITCENCRLFTCIDSTFNWQHRILLVRAREGMWIPVSTDRPWEASPSIHILTEILKG
ENV OF AB047240	(727)	
TRANSLATION OF E207TOP-LINK	(31)	
TRANSLATION OF ENV287-LINK	(29)	
TRANSLATION OF T20.22A-23	(399)	DSQTITCENCRLLTCIDSTFNWQHRILLVRAREGVWIPVSMDRPWEASPSVHILTEVLKG
PGD-E1	(17)	
PGD-E2	(17)	
PGD-E3	(1)	
CONSENSUS	(781)	DST W IL
		841 900
GI_4185940 EMB_CAA76880.1	(454)	VLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTRLWNSQSSI
GI_4185944_EMB_CAA76883.1_		VLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTRLWNSQSSI
GI_4185948_EMB_CAA76886.1_		VLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVQSVNFVNDWQKNSTRLWNSQSSI
GI_5931706 EMB_CAB56604.1		VLNRSKRFIFTLIAVIMGLIAVIATAAVAGVALHSSVQSVNFVNJWQKNSTRLWNSQSSI
ENV OF AB047240	(739)	VIIMASAAR IF INIRVIMONIAVIRIARVAGVANASSVQSVAFVAIMQAASIRIIMASQSSI
TRANSLATION OF E207TOP-LINK	(31)	
TRANSLATION OF ENV287-LINK	(29)	
TRANSLATION OF T20.22A-23		VLNRSKRFIFTLIAVIMGLIAVTATAAVAGVALHSSVOSVNFVNDWOKNSTRLWNSOSSI
PGD-E1	(135)	VIINKSKKFIFIIIAVINGLIAVIAIAAVAGVALHSSVQSVNFVNDWQINSIRLWNSQSSI
PGD-E2	(17)	
PGD-B2 PGD-E3	(1)	
CONSENSUS	(841)	
	1	901 960
GI_4185940_EMB_CAA76880.1_		DQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCITPQIYNESEHHWDMVRRH
GI_4185944_EMB_CAA76883.1_		DQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFSITPQIYNESEHHWDMVRRH
GI_4185948_EMB_CAA76886.1_		DQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCITPQIYNESELHWDMVRRH
GI_5931706_EMB_CAB56604.1_		DQKLASQINDLRQTVIWMGDRLMTLEHHFQLQCDWNTSDFCITPQIYNESEHHWDMVRRH
ENV OF AB047240	(739)	
TRANSLATION OF E207TOP-LINK	(31)	
TRANSLATION OF ENV287-LINK	(29)	
TRANSLATION OF T20.22A-23		DQKLANQINDLRQTVIWMGDRLMSLEHRFQLQCDWNTSDFCITPQIYNESEHHWDMVRRH
PGD-E1	(17)	
PGD-E2	(17)	
PGD-E3 CONSENSUS	(1) (901)	
	(501)	
GI_4185940 EMB CAA76880.1	(574)	961 LQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWVKTIGSTTI
GI_4185944 EMB CAA76883.1		LQGREDNLTLDISKLKEQIFEASKAHLNLVPGIEAIAGVADGLANLNPVTWVKTIGSTTI
GI_4185948_EMB_CAA76886.1_		LQGREDNITEDISKEKEQIFEASKAHENEVFGIEAIAGVADGEANENFVTWVKTIGSTI
GI 5931706 EMB CAB56604.1		LQGREDNLTLDISKLKEOIFEASKAHLNLVPGTEAIAGVADGHAWLAVFVIWVKIIGSIII
ENV OF AB047240	(739)	Leonard Didning The Control of Co
TRANSLATION OF E207TOP-LINK	(31)	
TRANSLATION OF ENV287-LINK	(29)	
TRANSLATION OF T20.22A-23		LQGREDNLTLDISKLKEQIFEASKAHLNLVPGTEAIAGVADGLANLNPVTWVKTIGSTTI
PGD-E1	(17)	
PGD-E2	(17)	
PGD-E3	• •	
CONSENSUS	(961)	
		1001
GI_4185940_EMB_CAA76880.1_	1571	1021 1081
GI_4185944_EMB_CAA76883.1		INLILILVCLFCLLLVCRCTQQLRRDSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV
GI_4185948_EMB_CAA76886.1		INLILILVCLFCLLLVCRCTQQLRRDSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV
GI_5931706_EMB_CAB56604.1_		INLILILVCLFCLLLVCRCTQQLRRDSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV
ENV OF AB047240		INLILIVVCLFCLLLVCRCTQQLRRDSDIENGP
TRANSLATION OF E207TOP-LINK	• •	
TRANSLATION OF E20/TOP-LINK TRANSLATION OF ENV287-LINK		
TRANSLATION OF ENV287-LINK TRANSLATION OF T20.22A-23		
_		INLILILVCLFCLLLVCRCTQQLRRDSDHRERAMMTMAVLSKRKGGNVGKSKRDQIVTVSV
PGD-E1 PGD-E2		
PGD-E2 PGD-E3		RCTQQLRRDSDHRERA
CONSENSUS	(1021)	RCTQQLRRDSD
CONDENSOS	(2061)	VCTXXTVVTGT

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 0004149 A [0251]
- US 5707829 A [0252]
- EP 0509612 B **[0252]**
- EP 0505012 B [0252]
- US 4959314 A [0252]
- US 5530101 A [0252]
- US 5585089 A [0252]
- WO 9824893 A [0252]
- WO 9110741 A [0252]
- WO 9630498 A [0252]
- WO 9402602 A [0252]
- US 5939598 A [0252]
- WO 9633735 A [0252]
- WO 9314778 A [0252]
- WO 9007936 A [0252]
- WO 9007930 A [0252]
 WO 9403622 A [0252]
- WO 9403022 A [0252]
 WO 9325698 A [0252]
- WO 9325236 A [0252]
 WO 9325234 A [0252]
- US 5219740 A [0252]
- WO 9311230 A [0252]
- WO 9310218 A [0252]
- US 4777127 A [0252]
- GB 2200651 A [0252]
- EP 0345242 A [0252]
- WO 9102805 A [0252]
- WO 9412649 A [0252]
- WO 9303769 A [0252]
- WO 9319191 A [0252]
- WO 9428938 A [0252]
- WO 9511984 A [0252]
- WO 9500655 A [0252]
- US 5814482 A [0252]
- WO 9507994 A [0252]
- WO 9617072 A [0252]
- WO 9530763 A [0252]
- WO 9742338 A [0252]
- WO 9011092 A [0252]
- US 5580859 A [0252]
- US 5422120 A [0252]
- WO 9513796 A [0252]
- WO 9423697 A [0252]
- WO 9114445 A **[0252]**
- EP 0524968 A [0252]
- US 5206152 A [0252]
- WO 9211033 A [0252]

- US 5149655 A [0252] WO 9014837 A [0252] WO 0007621 A [0252] GB 2220221 A [0252] EP 0689454 A [0252] EP 0835318 A [0252] EP 0735898 A [0252] EP 0761231 A [0252] WO 9952549 A [0252] WO 0121207 A [0252] WO 0121152 A [0252] WO 0062800 A [0252] WO 0023105 A [0252] WO 9911241 A [0252] WO 9857659 A [0252] WO 9313202 A [0252] US 5858723 A [0252] US 5010175 A [0252] WO 9117823 A [0252] US 4816567 A [0252] US 154 A [0252] US 5807522 A [0252] EP 0799897 A [0252] WO 9729212 A [0252] WO 9727317 A [0252] EP 0785280 A [0252] WO 9702357 A [0252] US 5593839 A [0252] US 5578832 A [0252] EP 0728520 A [0252] US 5599695 A [0252] EP 0721016 A [0252] US 5556752 A [0252] WO 9522058 A [0252] US 5631734 A [0252] US 5134854 A [0252] US 5445934 A [0252] WO 9535505 A [0252] US 5800992 A [0252] WO 9202526 A [0252] US 5124246 A [0252]
- US 187 A [0252]
- US 4683195 A [0252]

•

• US 4683202 A [0252]

Non-patent literature cited in the description

- MAYER et al. Nat. Genet., 1999, vol. 21 (3), 257-258 [0252]
- FARRELL. RNA Methodologies. Academic Press, 1998 [0252]
- YANG et al. Proc Natl Acad Sci USA, 1999, vol. 96 (23), 13404-8 [0252]
- ROBBINS et al. Clin Lab Sci, 1997, vol. 10 (5), 265-71 [0252]
- YLIKOSKI et al. Clin Chem, 1999, vol. 45 (9), 1397-407 [0252]
- YLIKOSKI et al. Biotechniques, 2001, vol. 30, 832-840 [0252]
- SHIRAHATA ; PEGG. J. Biol. Chem., 1986, vol. 261 (29), 13833-7 [0252]
- SAMBROOK et al. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, 1989 [0252]
- Short protocols in molecular biology. 1999 [0252]
- Current Protocols in Molecular Biology. 1987, 30
 [0252]
- BERKHOUT et al. J. Virol., 1999, vol. 73, 2365-2375
 [0252]
- LÖWER et al. J. Virol., 1995, vol. 69, 141-149 [0252]
- MAGIN et al. J. Virol., 1999, vol. 73, 9496-9507 [0252]
- MAGIN-LACHMANN. J. Virol., 2001, vol. 75 (21), 10359-71 [0252]
- HASHIDO et al. Biochem. Biophys. Res. Comm., 1992, vol. 187, 1241-1248 [0252]
- GEYSEN et al. PNAS USA, 1984, vol. 81, 3998-4002 [0252]
- CARTER. Methods Mol Biol, 1994, vol. 36, 207-23
 [0252]
- JAMESON, BA et al. CABIOS, 1988, vol. 4 (1), 181-186 [0252]
- RADDRIZZANI ; HAMMER. Brief Bioinform, 2000, vol. 1 (2), 179-89 [0252]
- DE LALLA et al. J. Immunol., 1999, vol. 163, 1725-29
 [0252]
- BRUSIC et al. *Bioinformatics*, 1998, vol. 14 (2), 121-30 [0252]
- MEISTER et al. Vaccine, 1995, vol. 13 (6), 581-91 [0252]
- ROBERTS et al. AIDS Res Hum Retroviruses, 1996, vol. 12 (7), 593-610 [0252]
- MAKSYUTOV ; ZAGREBELNAYA. Comput Appl Biosci, 1993, vol. 9 (3), 291-7 [0252]
- FELLER; DE LA CRUZ. Nature, 1991, vol. 349 (6311), 720-1 [0252]
- HOPP. Peptide Research, 1993, vol. 6, 183-190 [0252]
- WELLING et al. FEBS Lett., 1985, vol. 188, 215-218
 [0252]
- DAVENPORT et al. Immunogenetics, 1995, vol. 42, 392-297 [0252]
- SMITH ; WATERMAN. Adv. Appl. Math., 1981, vol. 2, 482-489 [0252]

- GO et al. Int. J. Peptide Protein Res., 1980, vol. 15, 211 [0252]
- QUEROL et al. Prot. Eng., 1996, vol. 9, 265 [0252]
- OLSEN; THOMSEN. J. Gen. Microbiol., 1991, vol. 137, 579 [0252]
- CLARKE et al. Biochemistry, vol. 32, 4322 [0252]
- WAKARCHUK et al. Protein Eng., 1994, vol. 7, 1379 [0252]
- TOMA et al. Biochemistry, 1991, vol. 30, 97 [0252]
- HAEZERBROUCK et al. Protein Eng., 1993, vol. 6, 643 [0252]
- MASUL et al. Appl. Env. Microbiol., 1994, vol. 60, 3579 [0252]
- BREEDVELD. Lancet, 2000, vol. 355 (9205), 735-740 [0252]
- GORMAN ; CLARK. Semin. Immunol., 1990, vol. 2, 457-466 [0252]
- JONES et al. Nature, 1986, vol. 321, 522-525 [0252]
- MORRISON et al. Proc. Natl. Acad. Sci, US.A., 1984, vol. 81, 6851-6855 [0252]
- MORRISON ; OI. Adv. Immunol., 1988, vol. 44, 65-92 [0252]
- VERHOEYER et al. Science, 1988, vol. 239, 1534-1536 [0252]
- PADLAN. Molec. Immun., 1991, vol. 28, 489-498
 [0252]
- PADLAN. Molec. Immunol., 1994, vol. 31 (3), 169-217 [0252]
- KETTLEBOROUGH, C.A. et al. Protein Eng., 1991, vol. 4 (7), 773-83 [0252]
- CHOTHIA et al. J. Mol. Biol., 1987, vol. 196, 901-917
 [0252]
- FINDEIS et al. Trends Biotechnol., 1993, vol. 11, 202 [0252]
- CHIOU et al. Gene Therapeutics: Methods And Applications Of Direct Gene Transfer, 1994 [0252]
- WU et al. J. Biol. Chem., 1988, vol. 263, 621 [0252]
- WU et al. J. Biol. Chem., 1994, vol. 269, 542 [0252]
- ZENKE et al. Proc. Natl. Acad. Sci. (USA), 1990, vol. 87, 3655 [0252]
- WU et al. J. Biol. Chem., 1991, vol. 266, 338 [0252]
- JOLLY. Cancer Gene Therapy, 1994, vol. 1, 51 [0252]
- KIMURA. Human Gene Therapy, 1994, vol. 5, 845 [0252]
- CONNELLY. Human Gene Therapy, 1995, vol. 1, 185 [0252]
- KAPLITT. Nature Genetics, 1994, vol. 6, 148 [0252]
- CURIEL. Hum. Gene Ther., 1992, vol. 3, 147 [0252]
- WU. J. Biol. Chem., 1989, vol. 264, 16985 [0252]
- PHILIP. Mol. Cell Biol., 1994, vol. 14, 2411 [0252]
- WOFFENDIN. Proc. Natl. Acad. Sci., 1994, vol. 91, 11581 [0252]
- Vaccine Design the subunit and adjuvant approach. 1995 [0252]

- MCSHARRY. Antiviral Res, 1999, vol. 43 (1), 1-21 [0252]
- KUHELJ et al. J Biol Chem, 2001, vol. 276 (20), 16674-82 [0252]
- SCHOMMER et al. J Gen Virol, 1996, vol. 77, 375-379 [0252]
- MAGIN et al. Virology, 2000, vol. 274, 11-16 [0252]
- BOESE et al. FEBS Lett, 2001, vol. 493 (2-3), 117-21 [0252]
- LARSSON, E. et al. Current Topics in Microbiology and Immunology, 1989, vol. 148, 115 [0252]
- MARIANI-COSTANTINI et al. J. Virol., 1989, vol. 63, 4982 [0252]
- TÖNJES et al. J. AIDS Hum. Retrovir., 1996, vol. 13 (1), S261-S267 [0252]
- BARBULESCU et al. Curr. Biol., 1999, vol. 9, 861 [0252]
- ONO et al. J. Virol., 1986, vol. 58, 937 [0252]
- LÖWER et al. Proc. Natl. Acad. Sci USA, 1993, vol. 90, 4480 [0252]
- ONO et al. J. Virol., 1986, vol. 60, 589 [0252]
- BOLLER et al. Virol., 1993, vol. 196, 349 [0252]
- YANG et al. Proc. Natl. Acad. Sci USA, 1999, vol. 96, 13404 [0252]
- MUELLER- LANTZSCH et al. AIDS Research and Human Retroviruses, 1993, vol. 9, 343-350 [0252]
- HERBST et al. Amer. J. Pathol., 1996, vol. 149, 1727
 [0252]
- LÖWER et al. Proc. Natl. Acad. Sci USA, 1996, vol. 93, 5177 [0252]
- LÖWER et al. Virology, 1993, vol. 192, 501 [0252]
- ANDERSSON et al. J. Gen. Virol., 1999, vol. 80, 255-260 [0252]
- ZSÍROS et al. J. Gen. Virol., 1998, vol. 79, 61-70
 [0252]
- TÖNJES et al. J. Virol., 1999, vol. 73, 9187-9195 [0252]
- JOHNSTON et al. Ann Neurol, vol. 50 (4), 434-42 [0252]
- MEDSTRAND et al. J Virol, vol. 72 (12), 9782-7 [0252]
- MERRIFELD. J. Am. Chem. Soc., 1963, vol. 85, 2149 [0252]
- CAPRINO ; HAN. J. Org. Chem., 1972, vol. 37, 3404
 [0252]
- MILSTEIN; KOHLER. Nature, 1975, vol. 256, 495-497 [0252]
- GULFRE; MILSTEIN. Methods in Enzymology: Immunochemical Techniques, vol. 73, 1-46 [0252]
- ALTSCHUL et al. Nucleic Acids Res., 1997, vol. 25, 3389-3402 [0252]
- BRUTLAG et al. Comp. Chem., 1993, vol. 17, 203 [0252]

- SCHENA et al. Proc Natl Acad Sci U S A., 1996, vol. 93 (20), 10614-9 [0252]
- SCHENA et al. Science, 1995, vol. 270 (5235), 467-70 [0252]
- SHALON et al. Genome Res., 1996, vol. 6 (7), 639-45 [0252]
- PAPPALARADO et al. Sem. Radiation Oncol., vol. 8, 217 [0252]
- RAMSAY. Nature Biotechnol., 1998, vol. 16, 40 [0252]
- MULLIS et al. Meth. Enzymol., 1987, vol. 155, 335 [0252]
- SAIKI et al. Science, vol. 239, 487 [0252]
- HANAHAN et al. Cell, 2000, vol. 100, 57-70 [0252]
- WEISSMAN SM. Mol Biol. Med., 1987, vol. 4 (3), 133-143 [0252]
- PATANJALI et al. Proc. Natl. Acad. Sci. USA, 1991, vol. 88 [0252]
- SIMONE et al. Am J Pathol., 2000, vol. 156 (2), 445-52 [0252]
- CLAVERIE. Meth. Enzymol., 1996, vol. 266, 212-227
 [0252]
- Automated DNA Sequencing and Analysis Techniques. Academic, 1994, 267 [0252]
- CLAVERIE et al. Comput. Chem., 1993, vol. 17, 191 [0252]
- ALTSCHUL. J. Mol. Biol., 1990, vol. 215, 403-410 [0252]
- PEARSON; LIPMAN. PNAS, 1988, vol. 85, 2444
 [0252]
- LUO et al. Nature Med, 1999, vol. 5, 117-122 [0252]
- HIGGINS ; SHARP. CABIOS, 1989, vol. 5, 151-153 [0252]
- DELLI BOVI et al. Cancer Res., 1986, vol. 46, 6333-6338 [0252]
- CESARONE, C. et al. Anal Biochem, 1979, vol. 100, 188-197 [0252]
- SOUTHERN, E. M. J. Mol. Biol., 1975, vol. 98, 503-517 [0252]
- FEINBERG, A. P. et al. Anal. Biochem., 1983, vol. 132, 6-13 [0252]
- WRIGHT; MANOS et al. PCR Protocols, 1990, 153-158 [0252]
- KEOWN et al. Methods in Enzymology, 1990, vol. 185, 527-537 [0252]
- MARKS et al. Brit. J. Urol., 1995, vol. 75, 225 [0252]
- SKEA et al. J. Immunol., 1993, vol. 151, 3557 [0252]
- MATHER et al. J. Nucl. Med., 1990, vol. 31, 692 [0252]
- ZHANG et al. Nucl. Med. Biol., 1992, vol. 19, 607 [0252]

patsnap

专利名称(译)	内源性逆转录病毒在前列腺癌中上认	周	
公开(公告)号	EP1415005B1	公开(公告)日	2012-11-21
申请号	EP2001996222	申请日	2001-12-07
[标]申请(专利权)人(译)	希龙公司		
申请(专利权)人(译)	Chiron公司		
当前申请(专利权)人(译)	诺华疫苗与诊断,INC.		
[标]发明人	GARCIA PABLO HARDY STEPHEN F WILLIAMS LEWIS T ESCOBEDO JAIME		
发明人	GARCIA, PABLO HARDY, STEPHEN, F. WILLIAMS, LEWIS, T. ESCOBEDO, JAIME		
IPC分类号	C07K14/15 C12N7/00 C12Q1/68 C /00 A61K48/00 A61P3/10 A61P15/ C12Q1/02 C12Q1/70 G01N33/569	00 A61P25/00 A61P35/00 C07K1	00 A61K39/00 A61K39/395 A61K45 6/10 C12N15/09 C12P21/08
CPC分类号	A61K39/00 A61K2039/505 A61K20 C12N2740/10021 C12N2740/1002 G01N2333/15		
代理机构(译)	MARSHALL , CAMERON JOHN		
优先权	60/251830 2000-12-07 US 10/016604 2001-12-07 US		
其他公开文献	EP1415005A2		
外部链接	<u>Espacenet</u>		

摘要(译)

HML-2家族的人内源性逆转录病毒在前列腺肿瘤中表现出上调的表达。 该发现可用于前列腺癌筛查,诊断和治疗。

TABLE 1 - GAG protease (5') probes, isolate specific						
Isolate	Nucleotides SEQ ID Isolate Nucleotides SEQ ID					
K(CH)	1224-1238	161		1490-1510	188	
KII	2098-2114	162	1	1502-1520	189	
	874-890	163	1	1522-1538	190	
	894-908	164	1	1561-1576	191	
	910-927	165	1	1586-1605	192	
	927-944	166	1	1620-1635	193	
	989-1004	167	1	1653-1669	194	
	1019-1036	168	1	1698-1723	195	
	1046-1063	169	1	1722-1743	196	
	1063-1078	170	1	1748-1762	197	
	1084-1103	171	1	1773-1788	198	
	1131-1145	172	1	1820-1834	199	
	1148-1163	173	1	1872-1887	200	
	1164-1185	174	174 K10	1917-1935	201	
K10	1206-1223	175	1	1940-1955	202	
	1216-1235	176	1	1955-1969	203	
	1243-1260	177		1973-1995	204	
	1258-2375	178		2008-2042	205	
	1277-1295	179		2049-2064	206	
	1300-1329	180	1	2076-2093	207	
	1347-1361	181	1	2097-2113	208	
	1367-1382	182	1	2122-2139	209	
	1392-1410	183		2148-2118	210	
	1412-1428	184		2176-2196	211	
	1426-1442	185	1	2198-2212	212	
	1445-1461	186	1	2219-2235	213	
	1463-1477	187	1	2246-2261	214	