Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2362524 A1 2000/08/17

(21) 2 362 524

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2000/02/15
(87) Date publication PCT/PCT Publication Date: 2000/08/17
(85) Entrée phase nationale/National Entry: 2001/08/14
(86) N° demande PCT/PCT Application No.: IB 00/00159
(87) N° publication PCT/PCT Publication No.: WO 00/47745
(30) Priorité/Priority: 1999/02/15 (99420041.8) EP

(51) Cl.Int.^{7}/Int.Cl.^{7} C12N 15/48, C12Q 1/68, C07K 16/10, C07K 14/15, C12N 15/11, G01N 33/569

(71) Demandeur/Applicant:
BIO MERIEUX, FR

(72) Inventeurs/Inventors:
KOMURIAN-PRADEL, FLORENCE, FR;
PERRON, HERVE, FR;
PARANHOS-BACCALA, GLAUCIA, FR

(74) Agent: SWABEY OGILVY RENAULT

(54) Titre : REGION LTR DU RETROVIRUS MSRV-1 ET PROTEINE QU'ELLE CODE, AINSI QUE SONDES ET PROCEDES POUR LA DETECTION DU RETROVIRUS MSRV-1

(54) Title: THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES AND METHODS FOR DETECTING MSRV-1 RETROVIRUS

(57) Abrégé/Abstract:
The invention relates to a nucleotide fragment of a LTR-RU5 region comprising a nucleotide sequence which encodes the expression of a protein, wherein said protein comprises a peptide sequence selected from SEQ ID NO: 2, SEQ ID NO: 3 and SEQ ID NO: 4, and a complementary nucleotide fragment; a probe and a primer hybridizing with said fragment; a protein encoded by said fragment and a polypeptide; an antibody directed against said protein or said polypeptide; and a process for detecting the presence of MSRV-1 retrovirus using a probe or an antibody of the invention.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/48, C07K 14/15, C12N 15/11, C07K 16/10, C12Q 1/68, G01N 33/569

A1

(11) International Publication Number: WO 00/47745

(43) International Publication Date: 17 August 2000 (17.08.00)

(21) International Application Number: PCT/IB00/00159

(22) International Filing Date: 15 February 2000 (15.02.00)

(30) Priority Data:
99420041.8 15 February 1999 (15.02.99) EP

(71) Applicant *(for all designated States except US)*: BIO MERIEUX [FR/FR]; Chemin de l'Orme, F–69280 Marcy L'Etoile (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: PARANHOS–BACCALA, Glaucia [CH/FR]; 75, Cours Gambetta, F–69003 Lyon (FR). PERRON, Hervé [FR/FR]; 15, rue de Boyer, F–69005 Lyon (FR). KOMURIAN–PRADEL, Florence [FR/FR]; 114, Chemin du Pavillon, F–69250 Poleymieux au Mont d'Or (FR).

(74) Agent: DIDIER, Mireille; Cabinet Germain & Maureau, Boîte postale 6153, F–69466 Lyon Cedex 06 (FR).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: THE LTR REGION OF MSRV–1 AND THE PROTEINS IT ENCODES, AND PROBES AND METHODS FOR DETECTING MSRV–1 RETROVIRUS

(57) Abstract

The invention relates to a nucleotide fragment of a LTR–RU5 region comprising a nucleotide sequence which encodes the expression of a protein, wherein said protein comprises a peptide sequence selected from SEQ ID NO: 2, SEQ ID NO: 3 and SEQ ID NO: 4, and a complementary nucleotide fragment; a probe and a primer hybridizing with said fragment; a protein encoded by said fragment and a polypeptide; an antibody directed against said protein or said polypeptide; and a process for detecting the presence of MSRV–1 retrovirus using a probe or an antibody of the invention.

THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES AND METHODS FOR DETECTING MSRV-1 RETROVIRUS

Multiple sclerosis (MS) is a demyelinating disease of the central nervous system (CNS) the cause of which remains as yet unknown.

Many studies have supported the hypothesis of a viral aetiology of the disease, but none of the known viruses tested has proved to be the causal agent sought: a review of the viruses sought for several years in MS has been compiled by E. Norrby and R.T. Johnson.

Recently, a retrovirus different from the known human retroviruses has been isolated in patients suffering from MS. The authors were also able to show that this retrovirus could be transmitted in vitro, that patients suffering from MS produced antibodies capable of recognizing proteins associated with the infection of leptomeningeal cells by this retrovirus, and that the expression of the latter could be strongly stimulated by the immediate-early genes of some herpesviruses.

All these results point to the role in MS of at least one unknown retrovirus or of a virus having reverse transcriptase activity which is detectable according to the method published by H. Perron and qualified as "LM7-like RT" activity.

The Applicant's studies have enabled two continuous cell lines infected with natural isolates originating from two different patients suffering from MS to be obtained by a culture method as described in the document WO-A-93/20188, the content of which is incorporated in the present description by reference. These two lines, derived from human choroid plexus cells, designated LM7PC and PLI-2, were deposited with the ECACC on 22nd July 1992 and 8th January 1993, respectively, under numbers 92072201 and 93010817, in accordance with the provisions of the Budapest Treaty. Moreover, the viral isolates possessing LM7-like RT activity were also deposited with the ECACC under the overall designation of "strains". The "strain" or isolate harboured by the PLI-2 line, designated POL-2, was deposited with the ECACC on 22nd July 1992 under No. V92072202. The "strain" or isolate harboured by the LM7PC line, designated MS7PG, was deposited with the ECACC on 8th January 1993 under No. V93010816.

CONFIRMATION COPY

Starting from the cultures and isolates mentioned above, characterized by biological and morphological criteria, the next step was to endeavour to characterize the nucleic acid material associated with the viral particles produced in these cultures.

The portions of the genome which have already been characterized have been used to develop tests for molecular detection of the viral genome and immunoserological tests, using the amino acid sequences encoded by the nucleotide sequences of the viral genome, in order to detect the immune response directed against epitopes associated with the infection and/or viral expression.

The viral system discovered by the Applicant is related to a complex retroviral system. In effect, the sequences to be found encapsidated in the extracellular viral particles produced by the different cultures of cells of patients suffering from MS show clearly that there is coencapsidation of retroviral genomes which are related but different from the "wild-type" retroviral genome which produces the infective viral particles. This phenomenon has been observed between replicative retroviruses and endogenous retroviruses belonging to the same family, or even heterologous retroviruses. The notion of endogenous retroviruses is very important. In the case of MSRV-1, it has been observed that endogenous retroviral sequences comprising sequences homologous to the MSRV-1 genome exist in normal human DNA. The existence of endogenous retroviral elements (ERV) related to MSRV-1 by all or part of their genome explains the fact that the expression of the MSRV-1 retrovirus in human cells is able to interact with closely related endogenous sequences.

Defective clones which only express proteins may be involved in the pathology.

The Applicant has made an unexpected discovery, according to which RU5 region of a retroviral LTR that is defined in the present invention, encodes the expression of at least one protein. This is unusual for LTRs, in particular in the RU5 region.

The present invention first relates to a nucleotide fragment of a LTR-RU5 region comprising a nucleotide sequence which encodes the expression of a protein, wherein said protein comprises at least six, preferably at least eight, and more preferably at least twelve, amino acids

of a peptide sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4, and a complementary nucleotide fragment.

Advantageously, a nucleotide fragment of the invention, or the complementary nucleotide fragment thereof, comprises a nucleotide sequence which encodes the expression of a protein, wherein said protein comprises a peptide sequence selected from SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4, and complementary nucleotide fragment. Preferably said protein consists of SEQ ID NO:3 or comprises or consists of SEQ ID NO:2 and SEQ ID NO:4.

The invention also relates to the following matter :

- a nucleic acid probe for the detection of MSRV-1 retrovirus, which comprises 10 to 1000 monomers and specifically hybridizes with the nucleotide fragment of the invention, in high stringency conditions ;

- a primer for the amplification by polymerization of a nucleic acid retroviral sequence of MSRV-1 virus, which comprises 10 to 30 monomers and hybridizes with the nucleotide fragment of the invention, in high stringency conditions ;

- a protein encoded by a nucleotide fragment of the invention ; preferably, said protein comprises a peptide sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4, or consists of a peptide sequence selected from the group consisting of SEQ ID NO:3, SEQ ID NO:2 and SEQ ID NO:4 ;

- a polypeptide comprising at least six, preferably at least eight, and more preferably at least twelve, amino acids of SEQ ID NO:4 ;

- a polyclonal or monoclonal antibody directed against a protein of the invention or a polypeptide of the invention ;

- a process for detecting, in a biological sample, the presence of MSRV-1 retrovirus comprising :

contacting a probe of the invention with said biological sample,

determining whether the probe binds to a nucleic acid in said biological sample, wherein binding indicates the presence of MSRV-1 virus ; said process may comprise an amplification step wherein said nucleic acid is amplified with a primer of the invention ;

- a process for detecting, in a biological sample, the presence of MSRV-1 retrovirus comprising :

contacting an antibody of the invention with said biological sample,

determining whether the antibody binds to a protein of the invention, in said biological sample, wherein binding indicates the presence of MSRV-1 virus ;

- a process for detecting, in a biological sample, the presence of MSRV-1 retrovirus comprising detecting the antigenic or biological properties of a protein of the invention or a fragment thereof ; advantageously, said protein fragment is a polypeptide of the invention.

Before describing the invention in detail, different terms used in the description and the claims are now defined :

- the term "MSRV" used in the present description denotes any pathogenic and/or infective agent associated with MS, in particular a viral species, the attenuated strains of the said viral species or the defective-interfering particles or particles containing coencapsidated genomes, or alternatively genomes recombined with a portion of the MSRV-1 genome, derived from this species. Viruses, and especially viruses containing RNA, are known to have a variability resulting, in particular, from relatively high rates of spontaneous mutation ;

- human virus is understood to mean a virus capable of infecting, or of being harboured by human beings,

- according to the invention, a nucleotide fragment or an oligonucleotide or polynucleotide is an arrangement of monomers, or a biopolymer, characterized by the informational sequence of the natural nucleic acids, which is capable of hybridizing with any other nucleotide fragment under predetermined conditions, it being possible for the arrangement to contain monomers of different chemical structures and to be obtained from a molecule of natural nucleic acid and/or by genetic recombination and/or by chemical synthesis; a nucleotide fragment may be identical to a genomic fragment of the MSRV-1 virus discussed in the present invention ;

-thus, a monomer can be a natural nucleotide of nucleic acid whose constituent elements are a sugar, a phosphate group and a nitrogenous base; in RNA the sugar is ribose, in DNA the sugar is 2-deoxyribose; depending on whether the nucleic acid is DNA or RNA, the nitrogenous base is chosen from adenine, guanine, uracil, cytosine and

thymine; or the nucleotide can be modified in at least one of the three constituent elements; as an example, the modification can occur in the bases, generating modified bases such as inosine, 5-methyldeoxycytidine, deoxyuridine, 5-(dimethylamino)deoxyuridine, 2,6-diaminopurine, 5-bromodeoxyuridine and any other modified base promoting hybridization; in the sugar, the modification can consist of the replacement of at least one deoxyribose by a polyamide, and in the phosphate group, the modification can consist of its replacement by esters chosen, in particular, from diphosphate, alkyl- and arylphosphonate and phosphorothioate esters ;

- « informational sequence » is understood to mean any ordered succession of monomers whose chemical nature and order in a reference direction constitute or otherwise an item of functional information of the same quality as that of the natural nucleic acids ;

- hybridization is understood to mean the process during which, under suitable working conditions, two nucleotide fragments having sufficiently complementary sequences pair to form a complex structure, in particular double or triple, preferably in the form of a helix in particular in high stringency conditions (see Maniatis et al., Molecular Cloning, Cold Spring Harbor, 1982) ; in general, depending on the length of the probes used, these conditions are the following : the temperature for the hybridization reaction is between approximately 20°C and 65°C, and especially between 35°C and 65°C, in a saline solution at a concentration of approximately 0.8 to 1 M ;

- a probe comprises a nucleotide fragment synthesized chemically or obtained by digestion or enzymatic cleavage of a longer nucleotide fragment, comprising at least six monomers, advantageously from 10 to 100 monomers and preferably 10 to 30 monomers, and possessing a specificity of hybridization under high stringency conditions; preferably, a probe possessing fewer than 10 monomers is not used alone, but is used in the presence of other probes of equally short size or otherwise; under certain special conditions, it may be useful to use probes of size greater than 100 monomers; a probe may be used, in particular, for diagnostic purposes, such molecules being, for example, capture and/or detection probes ;

- the capture probe may be immobilized on a solid support by any suitable means, that is to say directly or indirectly, for example by covalent bonding or passive adsorption ;

- the detection probe may be labelled by means of a label chosen, in particular, from radioactive isotopes, enzymes chosen, in particular, from peroxidase and alkaline phosphatase and those capable of hydrolysing a chromogenic, fluorogenic or luminescent substrate, chromophoric chemical compounds, chromogenic, fluorogenic or luminescent compounds, nucleotide base analogues and biotin ;

- the probes used for diagnostic purposes of the invention may be employed in all known hybridization techniques, and in particular the techniques termed « DOT-BLOT », « SOUTHERN BLOT », « NORTHERN BLOT », which is a technique identical to the « SOUTHERN BLOT » technique but which uses RNA as target, and the SANDWICH technique ; advantageously, the SANDWICH technique is used in the present invention, comprising a specific capture probe and/or a specific detection probe, on the understanding that the capture probe and the detection probe must possess an at least partially different nucleotide sequence,

- a primer is a probe comprising at least six monomers, and advantageously from 10 to 30 monomers, possessing a specificity of hybridization under high stringency conditions for the initiation of an enzymatic polymerization, for example in an amplification technique such as PCR (polymerase chain reaction), in an elongation process such as sequencing, in a method of reverse transcription or the like ;

In view of the fact that a virus possessing reverse transcriptase enzymatic activity may be genetically characterized equally well in RNA and in DNA form, both the viral DNA and RNA will be referred to for characterizing the sequences relating to a virus possessing such reverse transcriptase activity, termed MSRV-1 according to the present description.

Detection of a substance or agent is understood below to mean both an identification and a quantification, or a separation or isolation, of the said substance or said agent.

A better understanding of the invention will be gained on reading the detailed description which follows, prepared with reference to the attached figures, in which:

Figure 1 shows 5′ nucleotide sequence corresponding to regions encompassed by clones LB15 (represented in nt 1 to 623) in RU5 region and clones CL2 (represented in nt 624 to 1680) and CL17 (represented in nt 1681 to 2052) in *gag*. The aminoacid (AA) translation is shown under the *gag* nucleotidic sequences. The polyadenylation site in R is boxed and the polyadenylation downstream signal in U5 is underlined. (//) indicates a frameshift and (.) a stop codon. AA in bold characters correspond to the major homology region (MHR) in the capsid. Underlined AA correspond to the conserved positions in nucleocapsid.

Figure 2 shows 5' *env* region were obtained from clone C15 (nt 1 to 1479) and 3' *env* and LTR sequences were obtained from clone CL6 (nt 1480 to 2030) . Vertical arrows indicate the *env* region and the U3, R substructures. In the *env* region : the peptide signal and the putative immunosuppressive peptide are underlined, the N-linked glycosylation sites are boxed, and the two putative cleavage sites are indicated by vertical arrows. In the U3R region : the CAAT regulatory element, the TATA box are underlined ; the cap site and the polyA signal are also indicated.

Figure 3 shows the promotor activity of U3R clones obtained from MS plasma RNA (CL6), normal placenta RNA (PH74, accession number AF072506) and human cell DNA (5M6). U3R sequences were cloned into pCAT3 Enhancer*r* reporter vector. CAT activity was evaluated after 48 hours of incubation and represents the promotor efficiency of the corresponding sequences. The value presented correspond to the mean of 3 independant experiments.

Example 1 : Determination of MSRV-LTR regions

A RT-PCR amplification was performed with an antisense 3′ primers located in the *gag* region and with a 5′ sense primer defined from the R sequence previously obtained in the 3'end (clone 6, described below). The clone LB15 (SEQ ID NO:1) encompassing R (87 bp), U5 (456 bp), PBS and 5′*gag* regions, was thus obtained from concentrated culture particles. A polyadenylation site is compatible with the "CA" dinucleotide motif located at the junction of R and U5 regions and a putative poly(A) downstream signal is located at a distance of 24 nucleotides downstream of the poly(A) signal, with the consensus sequence " YGTGTTYY ". The putative primer-binding site (PBS), identified downstream the U5 region,

has proved to be related to the 3' end of avian $tRNA^{Trp}$ complementary sequence.

It is also noteworthy that 2 rather short in-frame *orfs* (SEQ ID NO:2 and SEQ ID NO:3 were found in the RU5 region.

The 3' LTR U3R region was identified in the CL6 clone (Fig. 2) obtained by amplification on DNase-treated RNA extracted from MS plasma. By comparison with the 5' LTR sequences obtained by 5' extension on MSRV RNA, a 299 bp-long U3 region and a 84 bp-long R region in CL6 LTR region, located downstream the *env orf* and ending with the polyA-tail has been identified. The typical regulatory element CAAT box was observed in two locations in the putative U3 region, but other putative binding sites for several transcription factors were also detected with Transcription Factor database (Heinemeyer T, Wingender, E , Reuter, I, Hermjakob, H, Kel, AE, Kel, OV, Ignatieva, EV, Ananko, EA, Podkolodnaya, OA , Kolpakov, FA, Podkolodny, NL, and Kolchanov, NA. (1998). Databases on transcriptional regulation : TRANSFAC, TRRD, and COMPEL. *Nucleic Acids Res.* **26**, 364-370).

A TATA box (TATAAA) was observed in the U3 region as described for numerous retroviruses. Finally, a poly(A) signal (AATAAA) was found 83 bp downstream of the TATA box.

In order to evaluate the promotor activity of LTR clones from different origins, CAT assays with sub-cloned U3R regions from LTR clones obtained from MS plasma RNA (CL6) and from related HERV-W copies in non-MS DNA (5M6) and in non-MS placenta RNA (PH74) were performed. As shown in Figure 3, a potent activity was associated with the LTR sequence in CL6 clone from MS plasma, whereas a weak and moderate promotor activity was found with 5M6 and PH74 clones respectively. Similar results were obtained with different cell types.

Example 2 : Material and methods used in example 1

Total RNA was extracted from MS plasma or purified particles by standard acidified guanidium thiocyanate procedure or by the « viral RNA extraction kit » (Boehringer Mannheim). After a DNase I treatment, RNA was reverse transcribed (Expand™ RT, Boehringer Mannheim) with either random hexanucleotide primers, specific MSRV primer or anchored oligodT primer, and was amplified by nested or semi-nested PCR (Long

Expand PCR kit - Boerhinger Mannheim). For each assay, a no-RT control was performed. 5' and 3' extension from the MSRV *pol* region was performed to obtain clones with primers indicated below.

LB15 :

Sense primer

5′ GCAACAGCAACCCCCTTTGGGT 3′

Antisense primers

5′ CTTGGAGGGTGCATAACCAGGGAAT 3′

5′ CTATGTCCTTTTGGACTGCTGTTTGGGT 3′

CL6 :

Sense primers

5′ GCCATCAAGCCACCCAAGAACTCTTAACTT 3′

5′ CCAATAGCCAGACCATTATATACACTAATT 3′

Antisense primer

5′ GACTCGCTGCAGATCGATTTTTTTTTTTTTTTTT 3′

PCR fragments were cloned into a pCR2.1 of the TA cloning™ kit (Invitrogen) and were sequenced in the both directions using the 'Prism ready reaction kit dye deoxyterminator cycle sequencing kit' (Applied Biosystems), with the Applied Biosystem 377 and 373A automated DNA sequencers.

U3R regions from CL6 (as shown in WO-99/02666 in the name of Applicant), PH74 (as shown in WO-99/02696 in the name of Applicant) and 5M6 (as shown in WO-99/02666 in the name of Applicant) clones (respectively obtained from MS plasma RNA, placenta RNA and Human DNA), were cloned into pCAT3 (Promega-Biotech, Madison, WI, USA) for CAT assays. Transfection experiments were performed on Hela, PG4, BeWo or Jurkat cells using the Superfect Transfection Kit (Qiagen GmbH, Germany) with 2 μg of purified recombinant plasmid. After 48h of incubation, the cells were harvested in order to evaluate CAT activity by the use of the CAT Enzyme Assay System (Promega-Biotech). For this purpose the Liquid Scintillation Counting (LSC) protocol was followed as recommended by the manufacturer. A positive control consisted of cells transfected by 2 μg of pCAT3 Contol^r (Promega-Biotech).

CLAIMS

1. A nucleotide fragment of a LTR-RU5 region selected from:

fragments comprising a nucleotide sequence which encodes the expression of a protein, wherein said protein comprises a peptide sequence selected from SEQ ID NO:2nd SEQ ID NO:4,

fragments consisting of a nucleotide sequence which encodes the expression of a protein consisting of SEQ ID NO:3, and

complementary nucleotide fragments thereof.

2. The nucleotide fragment of claim 1, wherein said protein comprises SEQ ID NO:2 and SEQ ID NO:4.

3. The nucleotide fragment of claim 1, wherein said protein consists of SEQ ID NO:2 and SEQ ID NO:4.

4. A nucleic acid probe for the detection of a LTR-RU5 region of MSRV-1 retrovirus wherein said probe comprises 10 to 1000 monomers and specifically hybridizes with a nucleotide sequence which encodes the expression of a protein, wherein said protein consists of SEQ ID NO:2, in high stringency conditions.

5. A primer for the amplification by polymerization of a nucleic acid retroviral sequence of a LTR-RU5 region of MSRV-1 virus, wherein said primer comprises 10 to 30 monomers and hybridizes with a nucleotide sequence which encodes the expression of a protein, wherein said protein consists of SEQ ID NO:2, in high stringency conditions.

6. A protein encoded by a nucleotide fragment as claimed in any one of claims 1 to 3.

7. The protein of claim 6 comprising a peptide sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4.

8. The protein of claim 6 consisting of peptide sequence SEQ ID NO:3.

9. The protein of claim 6 comprising a peptide sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4.

10. The protein of claim 6 consisting of a peptide sequence selected from SEQ ID NO:2 and SEQ ID NO:4.

11. A polypeptide comprising at least 6 amino acids of SEQ ID NO:2.

12. Polypeptide of claim 11, comprising at least 8 amino acids of SEQ ID NO:2.

Empfangszeit 31.Mai 19:29

13. Polypeptide of claim 12, comprising at least 12 amino acids of SEQ ID NO:2.

14. A polyclonal or monoclonal antibody directed against a protein as claimed in any one of claims 6 to 10 or a polypeptide as claimed in any one of claims 11 to 13.

15. A process for detecting, in a biological sample, the presence of MSRV-1 retrovirus comprising :

contacting a probe of claim 4 with said biological sample,

determining whether the probe binds to a nucleic acid in said biological sample, wherein binding indicates the presence of MSRV-1 virus.

16. A process for detecting, in a biological sample, the presence of MSRV-1 retrovirus comprising :

contacting an antibody of claim 14 with said biological sample,

determining whether the antibody binds to a protein in said biological sample, wherein binding indicates the presence of MSRV-1 virus.

17. A process for detecting, in a biological sample, the presence of MSRV-1 retrovirus comprising detecting the antigenic or biological properties of a protein as claimed in any one of claims 6 to 10 or a fragment thereof.

18. Process of claim 17, wherein said fragment is a polypeptide of claim 11, 12 or 13.

FIG. 1

R

CAGCAACCCC CTTTGGGTCC CCTCCCATTG TATGGGAGCT CTGTTTTCAC TCTATTTCAC TCTATTAAAT CATGCAACTG CACTCTTCTG GTCCGTGTTT 100

US

TTTATGGCTC AAGCTGAGCT TTTGTTCGCC ATCCACCACT GCTGTTTGCC ACCGTCACAG ACCCGCTGCT GACTTCCATC CCTTTGGATC CAGCAGAGTG 200

TCCGCTGTGC TCCTGATCCA GCACAGGCGC CCATTGCCTC TCCCAATTGG GCTAAAGGCT TGCCATTGTT CCTGCACAGC TAAGTGCCTG GGTTCATCCT 300

AATCGAGCTG AACACTAGTC ACTGGGTTCC ACGGTTCTCT TCCATGACCC ATGGCTTCTA ATAGAGCTAT AACACTCACT GCATGGTCCA AGATTCCATT 400

CCTTGGAATC CGTGAGACCA AGAACCCCAG GTCAGAGAAC ACAAGGCTTG CCACCATGTT GGAAGCAGCC CACCACCATT TTGGAAGCAG CCCGCCACTA 500

gag

TCTTGGGAGC TCTGGGAGCA AGGACCCCAG GTAACAATTT GGTGACCACG AAGGGACCTG AATCCGCAAC CATGAAGGGA TCTCCAAAGC AATGGGAAAC 600

PBS

M G N 3

GTTCCCCCCG AGGCAAAAAT GCCCCTAGAA CGTATTCTGG AGAATTGGGA CCAATGTGAC ACTCAGACGC TAAGAAAGAA ACGATTTATA TTCTTCTGCA 700

V P P E A K M P L E R I L E N W D Q C D T Q T L R K K R F I F F C S 37

GTACCGCCTG GCCACAATAT CCTCTTCAAG GGAGAGAAAC CTGGCTTCCT GAGGGAAGTA TAAATTATAA CATCATCTTA CAGCTAGACC TCTTCTGTAG 800

T A W P Q Y P L Q G R E T W L P E G S I N Y N I I L Q L D L F C R 70

AAAGGAGGGC AAATGGAGTG AAGTGCCATA TGTGCAAACT TTCTTTTCAT TAAGAGACAA CTCACAATTA TGTAAAAAGT GTGGTTTATG CCCTACAGGA 900

K E G K W S E V P Y V Q T F F S L R D N S Q L C K K C G L C P T G 103

AGCCCTCAGA GTCCACCTCC CTACCCCAGC GTCCCCTCCC CGACTCCTTC CTCAACTAAT AAGGACCCCC CTTTAACCCA AACGGTCCAA AAGGAGATAG 1000

S P Q S P P P Y P S V P S P T P S S T N K D P P L T Q T V Q K E I D 137

ACAAAGGGGT AAACAATGAA CCAAAGAGTG CCAATATTCC CCGATTATGC CCCCTCCAAG CAGTGAGAGG AGGAGAATTC GGCCCAGCCA GAGTGCCTGT 1100

K G V N N E P K S A N I P R L C P L Q A V R G G E F G P A R V P V 170

ACCTTTTTCT CTCTCAGACT TAAAGCAAAT TAAAATAGAC CTAGGTAAAT TCTCAGATAA CCCTGACGGC TATATTGATG TTTTACAAGG GTTAGGACAA 1200

P P S L S D L K Q I K I D L G K F S D N P D G Y I D V L Q G L G Q 203

TCCTTTGATC TGACATGGAG AGATATAATG TTACTACTAA ATCAGACACT AACCCCAAAT GAGAGAAGTG CCGCTGTAAC TGCAGCCCGA GAGTTTGGCG 1300

S F D L T W R D I M L L L N Q T L T P N E R S A A V T A A R E F G D 237

ATCTTTGGTA TCTCAGTCAG GCCAACAATA GGATGACAAC AGAGGAAAGA ACAACTCCCA CAGGCCAGCA GGCAGTTCCC AGTGTAGACC CTCATTGGGA 1400

L W Y L S Q A N N R M T T E E R T T P T G Q Q A V P S V D P H W D 270

CACAGAATCA GAACATGGAG ATTGGTGCCA CAAACATTTG CTAACTTGCG TGCTAGAAGG ACTGAGGAAA ACTAGGAAGA AGCCTATGAA TTACTCAATG 1500

T E S E H G D W C H K H L L T C V L E G L R K T R K K P M N Y S M 303

ATGTCCACTA TAACACAGGG AAAGGAAGAA AATCTTACTG CTTTTCTGGA CAGACTAAGG GAGGCATTGA GGAAGCATAC CTCCCTGTCA CCTGACTCTA 1600

M S T I T Q G K E E N L T A F L D R L R E A L R K H T S L S P D S I 337

TTGAAGGCCA ACTAATCTTA AAGGATAAGT TTATCACTCA GTCAGCTGCA GACATTAGAA AAAACTTCAA AAGTCTGCCT TTAGGCTCGG AACAAAACTT 1700

E G Q L I L K D K F I T Q S A A D I R K N F K S L P L G S E Q N L 370

AGAAACCCTA TTGAACTTGG CAACCTCGGT TTTTTATAAT AGAGATCAGG AGGAGCAGGC AGAAAGACAA ACGGGATTAA AAAAAAAGGC CACCGCTTTA 1800

E T L L N L A T S V F Y N R D Q E E Q A E G Q T G L K K K A T A L 403

GTCATGGCCC TCAGGCAAGC GGACTATGGA GGCTCTGGAA AAGGGAAAAG CTGAGAAAAT TGGATGCCTA ATAGGGTTTG CTTCCAGTGC GGTCTACAAG 1900

V M A L R Q A D Y G G S G K G K S . E N W M P N R V C F Q C G L Q G 437

GACACTTTAA AAAAGATTGT CCAAGTAGAA ATAAGCCGCC CCCTTGTCCA TGCCCCTTAC GTCAAGGGAA TCACTGGAAG GCCCACTGCC CCAGGGGATG 2000

H F K K D C P S R N K P P P C P C P L R Q G N H W K A H C P R G . 470

AAGATACTCT GAGTCAGAAG CCATTAACCA GATGATCCAG CAGCAGGACT GA 2052

R Y S E S E A I N Q M I Q Q Q D . 486

CA 02362524 2001-08-14

PCT/IB00/00159

FIG. 2

Env

```
ATGGCCCTCC CTTATCATAC TTTTCTCTTT ACTGTTCTCT TACCCCCTTT CGCTCTCACT GCACCCCCTC CATGCTGCTG TACAACCAGT AGCTCCCCTT      100
M  A  L  P  Y  H  T  F  L  F  T  V  L  L  P  P  F  A  L  T  A  P  P  P  C  C  C  T  T  S  S  S  P  Y            34
                                  Signal peptide
ACCAAGAGTT TCTATGAAGA ACGCGGCTTC CTGGAAATAT TGATGCCCCA TCATATAGGA GTTTATCTAA GGGAAACTCC ACCTTCACTG CCCACACCCA      200
  Q  E  F  L  .  R  T  R  L  P  G  N  I  D  A  P  S  Y  R  S  L  S  K  G  N  S  T  F  T  A  H  T  H              67

TATGCCCCGC AACTGCTATA ACTCTGCCAC TCTTTGCATG CATGCAAATA CTCATTATTG GACAGGGAAA ATGATTAATC CTAGTTGTCC TGGAGGACTT      300
 M  P  R  N  C  Y  N  S  A  T  L  C  M  H  A  N  T  H  Y  W  T  G  K  M  I  N  P  S  C  P  G  G  L              100

GGAGCCACTG TCTGTTGGAC TTACTTCACC CATACCAGTA TGTCTGATGG GGGTGGAATT CAAGGTCAGG CAAGAGAAAA ACAAGTAAAG GAAGCAATCT      400
G  A  T  V  C  W  T  Y  F  T  H  T  S  M  S  D  G  G  G  I  Q  G  Q  A  R  E  K  Q  V  K  E  A  I  S            134

CCCAACTGAC CGGGGGACAT AGCACCCCTA GCCCCTACAA AGGACTAGTT CTCTCAAAAC TACATGAAAC CCTCCGTACC CATACTCGCC TGGTGAGCCT      500
  Q  L  T  R  G  H  S  T  P  S  P  Y  K  G  L  V  L  S  K  L  H  E  T  L  R  T  H  T  R  L  V  S  L            167

ATTTAATACC ACCCTCACTC GGCTCCATGA GGTCTCAGCC CAAAACCCTA CTAACTGTTG GATGTGCCTC CCCCTGCACT TCAGGCCATA CATTTCAATC      600
F  N  T  T  L  T  R  L  H  E  V  S  A  Q  N  P  T  N  C  W  M  C  L  P  L  H  F  R  P  Y  I  S  I              200

CCTGTTCCTG AACAATGGAA CAACTTCAGC ACAGAAATAA ACACCACTTC CGTTTTAGTA GGACCTCTTG TTTCCAATCT GGAAATAACC CATACCTCAA      700
P  V  P  E  Q  W  N  N  F  S  T  E  I  N  T  T  S  V  L  V  G  P  L  V  S  N  L  E  I  T  H  T  S  N            234

ACCTCACCTG TGTAAAATTT AGCAATACTA TAGACACAAC CAGCTCCCAA TGCATCAGGT GGGTAACACC TCCCACACGA ATAGTCTGCC TACCCTCAGG      800
  L  T  C  V  K  F  S  N  T  I  D  T  T  S  S  Q  C  I  R  W  V  T  P  P  T  R  I  V  C  L  P  S  G            267

AATATTTTTT GTCTGTGGTA CCTCAGCCTA TCATTGTTTG AATGGCTCTT CAGAATCTAT GTGCTTCCTC TCATTCTTAG TGCCCCCTAT GACCATCTAC      900
I  F  F  V  C  G  T  S  A  Y  H  C  L  N  G  S  S  E  S  M  C  F  L  S  F  L  V  P  P  M  T  I  Y              300

ACTGAACAAG ATTTATACAA TCATGTCGTA CCTAAGCCCC ACAACAAAAG AGTACCCATT CTTCCTTTTG TTATCAGAGC AGGAGTGCTA GGCAGACTAG     1000
T  E  Q  D  L  Y  N  H  V  V  P  K  P  H  N  K  R  V  P  I  L  P  F  V  I  R  A  G  V  L  G  R  L  G            334

GTACTGGCAT TGGCAGTATC ACAACCTCTA CTCAGTTCTA CTACAAACTA TCTCAAGAAA TAAATGGTGA CATGGAACAG GTCACTGACT CCCTGGTCAC     1100
  T  G  I  G  S  I  T  T  S  T  Q  F  Y  Y  K  L  S  Q  E  I  N  G  D  M  E  Q  V  T  D  S  L  V  T            367

CTTGCAAGAT CAACTTAACT CCCTAGCAGC AGTAGTCCTT CAAAATCGAA GAGCTTTAGA CTTGCTAACC GCCAAAAGAG GGGGAACCTG TTTATTTTTA     1200
L  Q  D  Q  L  N  S  L  A  A  V  V  L  Q  N  R  R  A  L  D  L  L  T  A  K  R  G  G  T  C  L  F  L            400

GGAGAAGAAC GCTGTTATTA TGTTAATCAA TCCAGAATTG TCACTGAGAA AGTTAAAGAA ATTCGAGATC GAATACAATG TAGAGCAGAG GAGCTTCAAA     1300
G  E  E  R  C  Y  Y  V  N  Q  S  R  I  V  T  E  K  V  K  E  I  R  D  R  I  Q  C  R  A  E  E  L  Q  N            434

ACACCGAACG CTGGGGCCTC CTCAGCCAAT GGATGCCCTG GGTTCTCCCC TTCTTAGGAC CTCTAGCAGC TCTAATATTG TTACTCCTCT TTGGACCCTG     1400
  T  E  R  W  G  L  L  S  Q  W  M  P  W  V  L  P  F  L  G  P  L  A  A  L  I  L  L  L  L  F  G  P  C            467

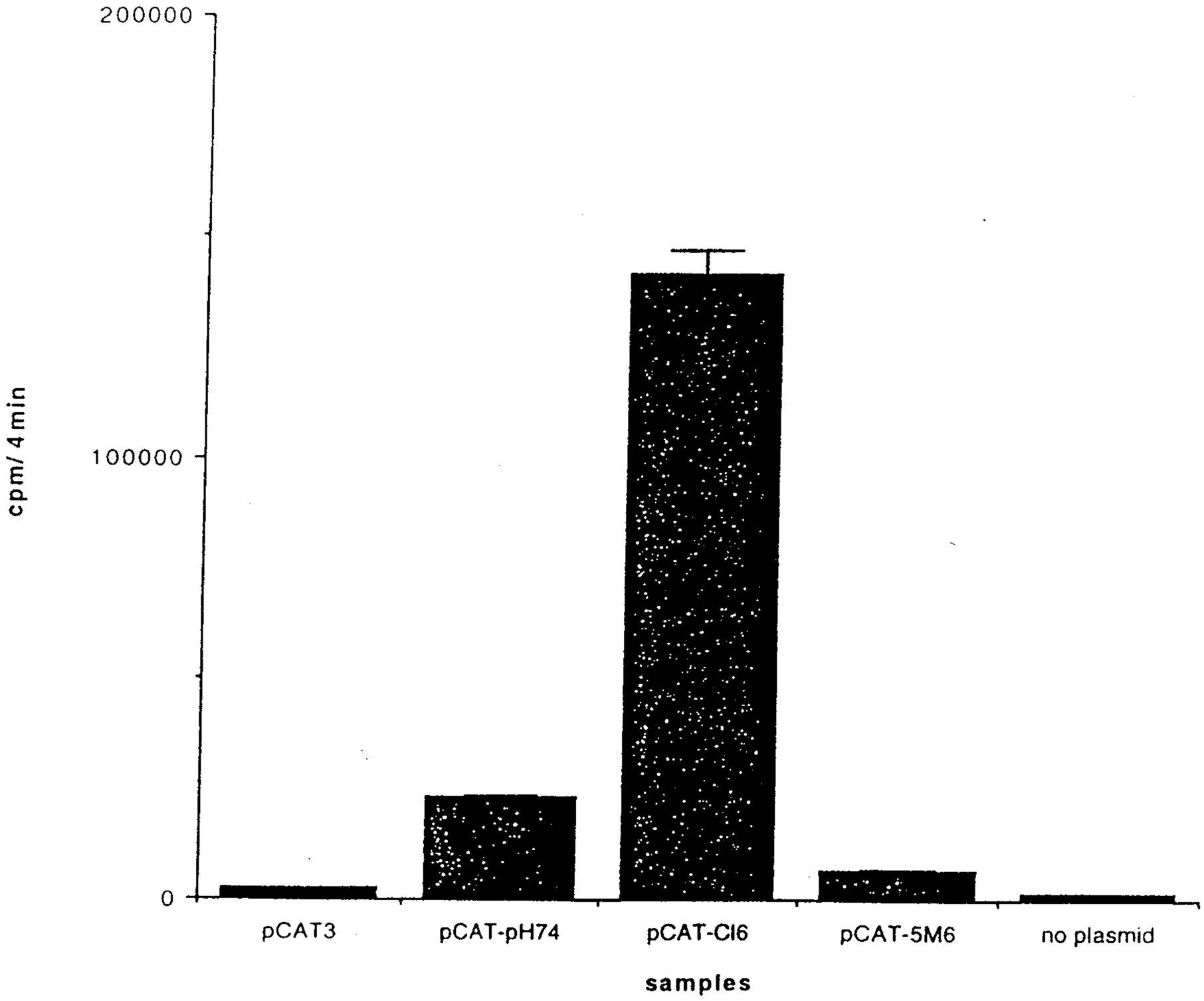
TATCTTTAAC CTCCTTGTTA AGTTTGTCTC TTCCAGAATT GAAGCTGTAA AGCTACAGAT GGTCTTACAA ATGGAACCCC AGATGGAGTC CATGACTAAG     1500
I  F  N  L  L  V  K  F  V  S  S  R  I  E  A  V  K  L  Q  M  V  L  Q  M  E  P  Q  M  E  S  M  T  K            500

ATCCACCGTG GACCCCTGGA CCGGCCTGCT AGCCCATGCT CCGATGTTAA TGACATTGAA GGCACCCCTC CCGAGGAAAT CTCAACTGCA CAACCCCTAC     1600
I  H  R  G  P  L  D  R  P  A  S  P  C  S  D  V  N  D  I  E  G  T  P  P  E  E  I  S  T  A  Q  P  L  L            534
                                  U3
TATGCCCCAA TTCAGCGGGA AGCAGTTAGA GCGGTCATCA GCCAACCTCC CCAACAGCAC TTGGGTTTTC CTGTTGAGAG GGGGGACTGA GAGACAGGAC     1700
  C  P  N  S  A  G  S  S  .                                                                                      542

TAGCTGGATT TCCTAGGCCA ACGAAGAATC CCTAAGCCTA GCTGGGAAGG TGACTGCATC CACCTCTAAA CATGGGGCTT GCAACTTAGC TCACACCCGA     1800

CCAATCAGAG AGCTCACTAA AATGCTAATT AGGCAAAAAT AGGAGGTAAA GAAATAGCCA ATCATCTATT GCCTGAGAGC ACAGCGGGAG GGACAAGGAT     1900
                                  R
CGGGATATAA ACCCAGGCAT TCGAGCCGGC AACGGCAACC CCCTTTGGGT CCCCTCCCTT TGTATGGGCG CTCTGTTTTC ACTCTATTTC ACTCTATTAA     2000
                                  Cap site
ATCTTGCAAC TGAAAAAAAA AAAAAAAAAA                                                                  Poly A signal  2030
```

FIG. 3


200000
100000
0
cpm/ 4 min
pCAT3
pCAT-pH74
pCAT-CI6
pCAT-5M6
no plasmid
samples

SEQUENCE LISTING

<110> BIO MERIEUX

<120> Nucleotide fragment, protein, probe, primer and process for detecting MSRV-1 retrovirus

<130> b05b3320-SEP20-MSRV1-LTR

<140>

<141>

<160> 4

<170> PatentIn Ver. 2.1

<210> 1
<211> 1003
<212> DNA
<213> MSRV-1 retrovirus

<400> 1
cagcaacccc ctttgggtcc cctcccattg tatgggagct ctgttttcac tctatttcac 60
tctattaaat catgcaactg cactcttctg gtccgtgttt tttatggctc aagctgagct 120
tttgttcgcc atccaccact gctgtttgcc accgtcacag acccgctgct gacttccatc 180
cctttggatc cagcagagtg tccgctgtgc tcctgatcca gcacaggcgc ccattgcctc 240
tcccaattgg gctaaaggct tgccattgtt cctgcacagc taagtgcctg ggttcatcct 300
aatcgagctg aacactagtc actgggttcc acggttctct tccatgaccc atggcttcta 360
atagagctat aacactcact gcatggtcca agattccatt ccttggaatc cgtgagacca 420
agaaccccag gtcagagaac acaaggcttg ccaccatgtt ggaagcagcc caccaccatt 480
ttggaagcag cccgccacta tcttgggagc tctgggagca aggaccccag gtaacaattt 540
ggtgaccacg aagggacctg aatccgcaac catgaaggga tctccaaagc aatgggaaac 600
gttccccccg aggcaaaaat gcccctagaa cgtattctgg agaattggga ccaatgtgac 660
actcagacgc taagaaagaa acgatttata ttcttctgca gtaccgcctg gccacaatat 720
cctcttcaag ggagagaaac ctggcttcct gagggaagta taaattataa catcatctta 780
cagctagacc tcttctgtag aaaggagggc aaatggagtg aagtgccata tgtgcaaact 840
ttcttttcat taagagacaa ctcacaatta tgtaaaaagt gtggtttatg ccctacagga 900
agccctcaga gtccacctcc ctaccccagc gtcccctccc cgactccttc ctcaactaat 960
aaggaccccc ctttaaccca aacaccagtc caaaaggaca tag 1003

<210> 2
<211> 76
<212> PRT
<213> MSRV-1 retrovirus

<400> 2

Met	Gln	Leu	His	Ser	Ser	Gly	Pro	Cys	Phe	Leu	Trp	Leu	Lys	Leu	Ser
1				5				10					15		
Phe	Cys	Ser	Pro	Ser	Thr	Thr	Ala	Val	Cys	His	Arg	His	Arg	Pro	Ala
			20					25					30		
Ala	Asp	Phe	His	Pro	Phe	Gly	Ser	Ser	Arg	Val	Ser	Ala	Val	Leu	Leu
		35					40					45			
Ile	Gln	His	Arg	Arg	Pro	Leu	Pro	Leu	Pro	Ile	Gly	Leu	Lys	Ala	Cys
	50					55					60				

His Cys Ser Cys Thr Ala Lys Cys Leu Gly Ser Ser
65 70 75

<210> 3
<211> 64
<212> PRT
<213> MSRV-1 retrovirus

<400> 3
Met Ala Ser Asn Arg Ala Ile Thr Leu Thr Ala Trp Ser Lys Ile Pro
1 5 10 15

Phe Leu Gly Ile Arg Glu Thr Lys Asn Pro Arg Ser Glu Asn Thr Arg
20 25 30

Leu Ala Thr Met Leu Glu Ala Ala His His His Phe Gly Ser Ser Pro
35 40 45

Pro Leu Ser Trp Glu Leu Trp Glu Gln Gly Pro Gln Val Thr Ile Trp
50 55 60

<210> 4
<211> 140
<212> PRT
<213> MSRV-1 retrovirus

<400> 4
Met Gln Leu His Ser Ser Gly Pro Cys Phe Leu Trp Leu Lys Leu Ser
1 5 10 15

Phe Cys Ser Pro Ser Thr Thr Ala Val Cys His Arg His Arg Pro Ala
20 25 30

Ala Asp Phe His Pro Phe Gly Ser Ser Arg Val Ser Ala Val Leu Leu
35 40 45

Ile Gln His Arg Arg Pro Leu Pro Leu Pro Ile Gly Leu Lys Ala Cys
50 55 60

His Cys Ser Cys Thr Ala Lys Cys Leu Gly Ser Ser Met Ala Ser Asn
65 70 75 80

Arg Ala Ile Thr Leu Thr Ala Trp Ser Lys Ile Pro Phe Leu Gly Ile
85 90 95

Arg Glu Thr Lys Asn Pro Arg Ser Glu Asn Thr Arg Leu Ala Thr Met
100 105 110

Leu Glu Ala Ala His His His Phe Gly Ser Ser Pro Pro Leu Ser Trp
115 120 125

Glu Leu Trp Glu Gln Gly Pro Gln Val Thr Ile Trp
130 135 140