



(11)

**EP 1 694 829 B1**

(12)

**FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention  
de la délivrance du brevet:  
**04.08.2010 Bulletin 2010/31**

(51) Int Cl.:  
**C12N 7/00 (2006.01)**

(21) Numéro de dépôt: **04805625.3**

(86) Numéro de dépôt international:  
**PCT/FR2004/003106**

(22) Date de dépôt: **02.12.2004**

(87) Numéro de publication internationale:  
**WO 2005/056584 (23.06.2005 Gazette 2005/25)**

(54) **NOUVELLE SOUCHE DE CORONAVIRUS ASSOCIE AU SRAS ET SES APPLICATIONS.**  
NEUER MIT SARS VERBUNDEN CORONAVIRUS STAMM UND SEINE VERWENDUNGEN  
NOVEL STRAIN OF SARS-ASSOCIATED CORONAVIRUS AND APPLICATIONS THEREOF

(84) Etats contractants désignés:  
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR  
HU IE IS IT LI LT LU MC NL PL PT RO SE SI SK TR**

(30) Priorité: **02.12.2003 FR 0314151**  
**02.12.2003 FR 0314152**

(43) Date de publication de la demande:  
**30.08.2006 Bulletin 2006/35**

(60) Demande divisionnaire:  
**10005885.8**

(73) Titulaires:  
• **INSTITUT PASTEUR**  
**75724 Paris Cedex 15 (FR)**  
• **CENTRE NATIONAL DE**  
**LA RECHERCHE SCIENTIFIQUE (CNRS)**  
**75794 Paris Cedex 16 (FR)**  
• **UNIVERSITE PARIS VII**  
**75251 Paris Cedex 05 (FR)**

(72) Inventeurs:  
• **VAN DER WERF, Sylvie**  
**F-91190 Gif-sur-yvette (FR)**  
• **ESCRIOU, Nicolas**  
**F-75014 Paris (FR)**  
• **CRESCENZO-CHAIGNE, Bernadette**  
**F-92200 Neuilly-sur-seine (FR)**  
• **MANUGUERRA, Jean-Claude**  
**F-75018 Paris (FR)**  
• **KUNST, Frederik,**  
**Inst. Pasteur**  
**Bureau des Brevets et Inventions**  
**75724 Paris Cedex 15 (FR)**  
• **CALLENDRET, Benoît**  
**F- 92000 Nanterre (FR)**

- **BETTON, Jean-Michel**  
**75014 Paris (FR)**
- **LORIN, Valérie**  
**92120 Montrouge (FR)**
- **GERBAUD, Sylvie**  
**94100 Saint Maur Des Fosses (FR)**
- **BURGUIERE, Ana Maria**  
**92140 Clamart (FR)**
- **AZEBI, Saliha**  
**94400 Vitry-sur-seine (FR)**
- **CHARNEAU, Pierre**  
**75005 Paris (FR)**
- **TANGY, Frédéric**  
**93260 Les Lilas (FR)**
- **COMBREDET, Chantal**  
**75017 Paris (FR)**
- **DELAGNEAU, Jean-François**  
**78170 La Celle Saint Cloud (FR)**
- **MARTIN, Monique**  
**92290 Chatenay Malabry (FR)**

(74) Mandataire: **Marcadé, Véronique et al**  
**Cabinet Ores**  
**36, rue de St Pétersbourg**  
**75008 Paris (FR)**

(56) Documents cités:  
• **DATABASE EMBL 22 avril 2003 (2003-04-22),**  
**XP002294758 Database accession no. AY278489**  
• **DATABASE EMBL 10 juin 2003 (2003-06-10),**  
**XP002294760 Database accession no. AY290752**  
• **DATABASE UNIPROT 10 octobre 2003**  
**(2003-10-10), XP002294761 Database accession**  
**no. P59595**

Il est rappelé que: Dans un délai de neuf mois à compter de la publication de la mention de la délivrance du brevet européen au Bulletin européen des brevets, toute personne peut faire opposition à ce brevet auprès de l'Office européen des brevets, conformément au règlement d'exécution. L'opposition n'est réputée formée qu'après le paiement de la taxe d'opposition. (Art. 99(1) Convention sur le brevet européen).

**EP 1 694 829 B1**

- MARRA M A ET AL: "The genome sequence of the SARS-associated coronavirus" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 300, no. 5624, 30 mai 2003 (2003-05-30), pages 1399-1404, XP002269483 ISSN: 0036-8075
- CHE X-Y ET AL: "RAPID AND EFFICIENT PREPARATION OF MONOCLONAL ANTIBODIES AGAINST SARS-ASSOCIATED CORONAVIRUS NUCLEOCAPSID PROTEIN BY IMMUNIZING MICE" DI YI JUNYI DAXUE XUEBAO - ACADEMIC JOURNAL OF FIRST MEDICAL COLLEGE OF PLA, GAI KAN BIANJISHI, GUANGZHOU, CN, vol. 23, no. 7, juillet 2003 (2003-07), pages 640-642, XP008028243 ISSN: 1000-2588
- WANG J ET AL: "ASSESSMENT OF IMMUNOREACTIVE SYNTHETIC PEPTIDES FROM THE STRUCTURAL PROTEINS OF SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS" CLINICAL CHEMISTRY, AMERICAN ASSOCIATION FOR CLINICAL CHEMISTRY. WINSTON, US, vol. 49, no. 12, 13 novembre 2003 (2003-11-13), pages 1989-1996, XP001182489 ISSN: 0009-9147
- SHI Y ET AL: "DIAGNOSIS OF SEVERE ACUTE RESPIRATORY SYNDROME (SARS) BY DETECTION OF SARS CORONAVIRUS NUCLEOCAPSID ANTIBODIES IN AN ANTIGEN-CAPTURING ENZYME-LINKED IMMUNOSORBENT ASSAY" JOURNAL OF CLINICAL MICROBIOLOGY, WASHINGTON, DC, US, vol. 41, no. 12, décembre 2003 (2003-12), pages 5781-5782, XP008028263 ISSN: 0095-1137
- LIU G ET AL: "The C-Terminal Portion of the Nucleocapsid Protein Demonstrates SARS-CoV Antigenicity" GENOMICS, PROTEOMICS AND BIOINFORMATICS, vol. 1, no. 3, 2003, pages 193-197, XP001183377
- POON LEO L M ET AL: "Rapid diagnosis of a coronavirus associated with severe acute respiratory syndrome (SARS)" CLINICAL CHEMISTRY, AMERICAN ASSOCIATION FOR CLINICAL CHEMISTRY. WINSTON, US, vol. 49, no. 6 Pt 1, juin 2003 (2003-06), pages 953-955, XP002288942 ISSN: 0009-9147

## Description

**[0001]** La présente invention est relative à une nouvelle souche de coronavirus associé au syndrome respiratoire aigu sévère (SRAS), issue d'un prélèvement répertorié sous le n° 031589 et prélevé à Hanoi (Vietnam), à des molécules d'acide nucléique issues de son génome, aux protéines et peptides codés par lesdites molécules d'acide nucléique ainsi qu'à leurs applications, notamment en tant que réactifs de diagnostic et/ou comme vaccin.

**[0002]** Le coronavirus est un virus à ARN monocaténaire, de polarité positive, d'approximativement 30 kilobases qui se réplique dans le cytoplasme des cellules hôtes ; l'extrémité 5' du génome a une structure en coiffe et l'extrémité 3' comporte une queue polyA. Ce virus est enveloppé et comprend, à sa surface, des structures péplomériques dénommées spicules.

**[0003]** Le génome comprend les cadres ouverts de lecture ou ORF suivants, de son extrémité 5' vers son extrémité 3' : ORF1a et ORF1b correspondant aux protéines du complexe de transcription-réplication, et ORF-S, ORF-E, ORF-M et ORF-N correspondant aux protéines structurales S, E, M et N. Il comprend également des ORFs correspondant à des protéines de fonction inconnue codées par : la région située entre l'ORF-S et l'ORF-E et chevauchant cette dernière, la région située entre l'ORF-M et l'ORF-N, et la région incluse dans l'ORF-N.

**[0004]** La protéine S est une glycoprotéine membranaire (200-220 kDa) qui se présente sous la forme de spicules ou "Spike" émergeant de la surface de l'enveloppe virale. Elle est responsable de l'attachement du virus aux récepteurs de la cellule hôte et de l'induction de la fusion de l'enveloppe virale avec la membrane cellulaire.

**[0005]** La petite protéine d'enveloppe (E) également dénommée sM (*small membrane*) qui est une protéine transmembranaire non glycosylée d'environ 10 kDa, est la protéine présente en plus faible quantité dans le virion. Elle joue un rôle moteur dans le processus de bourgeonnement des coronavirus qui se produit au niveau du compartiment intermédiaire dans le réticulum endoplasmique et l'appareil de Golgi

**[0006]** La protéine M ou protéine de matrice (25-30 kDa) est une glycoprotéine membranaire plus abondante qui est intégrée dans la particule virale par une interaction M/E, tandis que l'incorporation de S dans les particules est dirigée par une interaction S/M. Elle semble être importante pour la maturation virale des coronavirus et pour la détermination du site au niveau duquel les particules virales sont assemblées.

**[0007]** La protéine N ou protéine de nucléocapside (45-50 kDa) qui est la plus conservée parmi les protéines structurales des coronavirus, est nécessaire pour encapsider l'ARN génomique puis pour diriger son incorporation dans le virion. Cette protéine est vraisemblablement également impliquée dans la réplication de l'ARN.

**[0008]** Lorsqu'une cellule hôte est infectée, le cadre de lecture (ORF) situé en 5' du génome viral est traduit en une polyprotéine qui est clivée par les protéases virales et libère alors plusieurs protéines non-structurales telles que l'ARN-polymérase ARN dépendante (Rep) et l'ATPase hélicase (Hel). Ces deux protéines sont impliquées dans la réplication du génome viral ainsi que dans la génération de transcrits qui sont utilisés dans la synthèse des protéines virales. Les mécanismes par lesquels ces ARNms sub-génomiques sont produits, ne sont pas complètement compris ; cependant des faits récents indiquent que les séquences de régulation de la transcription à l'extrémité 5' de chaque gène représentent des signaux qui régulent la transcription discontinue des ARNms sub-génomiques.

**[0009]** Les protéines de la membrane virale (protéines S, E et M) sont insérées dans le compartiment intermédiaire, alors que l'ARN répliqué (brin +) s'assemble avec la protéine N (nucléocapside). Ce complexe protéine-ARN s'associe ensuite avec la protéine M incluse dans les membranes du réticulum endoplasmique et les particules virales se forment lorsque le complexe de la nucléocapside bourgeonne dans le réticulum endoplasmique. Le virus migre ensuite à travers le complexe du Golgi et éventuellement sort de la cellule, par exemple par exocytose. Le site de l'attachement du virus à la cellule hôte se trouve au niveau de la protéine S.

**[0010]** Les coronavirus sont responsables de 15 à 30 % des rhumes chez l'Homme et d'infections respiratoires ou digestives chez les animaux, notamment le chat (FIPV : *Feline infectious peritonitis virus*), la volaille (IBV : *Avian Infectious bronchitis virus*), la souris (MHV : *Mouse Hepatitis virus*), le porc (TGEV : *Transmissible gastroenteritis virus*, PEDV : *Porcine Epidemic Diarrhea virus*, PRCoV : *Porcine Respiratory Coronavirus*, HEV : *Hemagglutinating encephalomyelitis Virus*) et les bovins (BcoV : *Bovine coronavirus*).

**[0011]** En général, chaque coronavirus n'affecte qu'une seule espèce ; chez les individus immunocompétents, l'infection induit des anticorps éventuellement neutralisants et une immunité cellulaire, capables de détruire les cellules infectées.

**[0012]** Une épidémie de pneumonie atypique, dénommée syndrome respiratoire aigu sévère (SARS ou *Severe acute respiratory syndrome*, SRAS en français) s'est propagée dans différents pays (Vietnam, Hong-Kong, Singapour, Thaïlande et Canada) au cours du premier trimestre 2003, à partir d'un foyer initial apparu en Chine dans le dernier trimestre de 2002. La sévérité de cette maladie est telle que son taux de mortalité est d'environ 3 à 6 %. La détermination de l'agent causatif de cette maladie a été entreprise par de nombreux laboratoires, à travers le monde.

**[0013]** En mars 2003, un nouveau coronavirus (SARS-CoV, SARS virus ou virus SRAS, en français) a été isolé, en association avec des cas de syndrome respiratoire aigu sévère (T.G.KSIAZEK et al., *The New England Journal of Medicine*, 2003, 348, 1319-1330 ; C. DROSTEN et al., *The New England Journal of Medicine*, 2003, 348, 1967-1976 ;

Peiris et al., Lancet, 2003, 361, 1319-).

**[0014]** Des séquences génomiques de ce nouveau coronavirus ont ainsi été obtenues, notamment celles de l'isolat Urbani (Genbank n° d'accès AY274119.3 et A. MARRA et al., Science, May 1, 2003, 300, 1399-1404) et de l'isolat de Toronto (Tor2, Genbank n° d'accès AY 278741 et A. ROTA et al., Science, 2003, 300, 1394-1399).

**[0015]** L'organisation du génome est comparable à celle des autres coronavirus connus permettant ainsi de confirmer l'appartenance du SARS-CoV à la famille des *Coronaviridae* ; les cadres ouverts de lecture ORF1a et 1b et les cadres ouverts de lecture correspondant aux protéines S, E, M, et N, ainsi qu'à des protéines codées par : la région située entre l'ORF-S et l'ORF-E (ORF3), la région située entre l'ORF-S et l'ORF-E et chevauchant l'ORF-E (ORF4), la région située entre l'ORF-M et l'ORF-N (ORF7 à ORF11) et la région correspondant à l'ORF-N (ORF13 et ORF14), ont notamment été identifiées.

**[0016]** Sept différences ont été mises en évidence entre les séquences des isolats Tor2 et Urbani ; 3 correspondent à des mutations silencieuses (c/t en position 16622 et a/g en position 19064 de l'ORF1b, t/c en position 24872 de l'ORF-S) et 4 modifient la séquence en acides aminés de respectivement : les protéines codées par l'ORF1a (c/t en position 7919 correspondant à la mutation A/V), la protéine S (g/t en position 23220 correspondant à la mutation A/S), la protéine codée par l'ORF3 (a/g en position 25298 correspondant à la mutation R/G) et de la protéine M (t/c en position 26857 correspondant à la mutation S/P).

**[0017]** En outre, l'analyse phylogénétique montre que le SARS-CoV est éloigné des autres coronavirus et qu'il est apparu, ni par mutation de coronavirus respiratoires humains, ni par recombinaison entre des coronavirus connus (pour une revue, voir Holmes, J.C.I., 2003, 111, 1605-1609).

**[0018]** La mise en évidence et la prise en compte de nouveaux variants sont importantes pour la mise au point de réactifs de détection et de diagnostic du SRAS suffisamment sensibles et spécifiques ainsi qu'à des compositions immunogènes aptes à protéger des populations contre des épidémies de SRAS.

**[0019]** Les Inventeurs ont maintenant mis en évidence une autre souche de coronavirus associé au SRAS, qui se distingue des isolats Tor2 et Urbani.

**[0020]** La présente invention a donc pour objet, une souche isolée ou purifiée de coronavirus humain associé au syndrome respiratoire aigu sévère, caractérisée en ce que son génome présente sous la forme d'ADN complémentaire un codon sérine en position 23220-23222 du gène de la protéine S ou un codon glycine en position 25298-25300 du gène de l'ORF3, et un codon alanine en position 7918-7920 de l'ORF1a ou un codon sérine en position 26857-26859 du gène de la protéine M, lesdites positions étant indiquées en référence à la séquence Genbank AY274119.3.

**[0021]** Selon un mode de réalisation avantageux de ladite souche, l'équivalent ADN de son génome présente une séquence correspondant à la séquence SEQ ID NO : 1 ; cette souche de coronavirus est issue du prélèvement de lavage bronchoalvéolaire d'un patient atteint de SRAS, répertorié sous le n° 031589 et effectué à l'hôpital français de Hanoi (Vietnam).

**[0022]** Conformément à l'invention, ladite séquence SEQ ID NO :1 est celle de l'acide désoxyribonucléique correspondant à la molécule d'acide ribonucléique du génome de la souche isolée de coronavirus telle que définie ci-dessus.

**[0023]** La séquence SEQ ID NO : 1 se distingue de la séquence Genbank AY274119.3 (isolat Tor2) en ce qu'elle possède les mutations suivantes :

- g/t en position 23220 ; le codon alanine (gct) en position 577 de la séquence en acides aminés de la protéine S de Tor2 est remplacé par un codon sérine (tct),
- a/g en position 25298 ; le codon arginine (aga) en position 11 de la séquence en acides aminés de la protéine codée par l'ORF3 de Tor 2 est remplacé par un codon glycine (gga).

**[0024]** En outre, la séquence SEQ ID NO : 1 se distingue de la séquence Genbank AY278741 (isolat Urbani) en ce qu'elle possède les mutations suivantes :

- t/c en position 7919 ; le codon valine (ggt) en position 2552 de la séquence en acides aminés de la protéine codée par l'ORF1a est remplacé par un codon alanine (gct),
- t/c en position 16622 : cette mutation ne modifie pas la séquence en acides aminés des protéines codées par l'ORF1b (mutation silencieuse),
- g/a en position 19064 : cette mutation ne modifie pas la séquence en acides aminés des protéines codées par l'ORF1b (mutation silencieuse),
- c/t en position 24872 : cette mutation ne modifie pas la séquence en acides aminés de la protéine S, et
- c/t en position 26857 : le codon proline (ccc) en position 154 de la séquence en acides aminés de la protéine M est remplacé par un codon sérine (tcc).

**[0025]** En l'absence de mention particulière, les positions des séquences nucléotidiques et peptidiques sont indiquées en référence à la séquence Genbank AY274119.3.



**[0026]** La présente invention a également pour objet un polynucléotide isolé ou purifié, caractérisé en ce que sa séquence est celle du génome de la souche isolée de coronavirus telle que définie ci-dessus.

**[0027]** Selon un mode de réalisation avantageux dudit polynucléotide il présente la séquence SEQ ID NO : 1.

**[0028]** Les Inventeurs décrivent également un polynucléotide isolé ou purifié, caractérisé en ce que sa séquence hybride dans des conditions de forte stringence avec la séquence du polynucléotide tel que défini ci-dessus.

**[0029]** Les termes « isolé ou purifié » signifient modifié « par la main de l'homme » à partir de l'état naturel ; autrement dit si un objet existe dans la nature, il est dit isolé ou purifié s'il a été modifié ou extrait de son environnement naturel ou les deux. Par exemple, un polynucléotide ou une protéine/un peptide naturellement présent dans un organisme vivant n'est ni isolé, ni purifié ; en revanche le même polynucléotide ou protéine /peptide séparé des molécules coexistantes dans son environnement naturel, obtenu par clonage, amplification et/ou synthèse chimique est isolé au sens de la présente invention. De plus, un polynucléotide ou une protéine/peptide qui est introduit dans un organisme par transformation, manipulation génétique ou par toute autre méthode, est « isolé » même s'il est présent dans ledit organisme. Le terme purifié tel qu'utilisé dans la présente invention, signifie que les protéines /peptides selon l'invention sont essentiellement libres d'association avec les autres protéines ou polypeptides, comme l'est par exemple le produit purifié de la culture de cellules hôtes recombinantes ou le produit purifié à partir d'une source non-recombinante.

**[0030]** On entend par conditions d'hybridation de forte stringence, des conditions de température et de force ionique choisies de telle manière qu'elles permettent le maintien de l'hybridation spécifique et sélective entre polynucléotides complémentaires.

**[0031]** A titre d'illustration, des conditions de forte stringence aux fins de définir les polynucléotides ci-dessus, sont avantageusement les suivantes : l'hybridation ADN-ADN ou ADN-ARN est réalisée en deux étapes : (1) préhybridation à 42°C pendant 3 heures en tampon phosphate (20 mM pH 7,5) contenant 5 x SSC (1 x SSC correspond à une solution 0,15 M NaCl + 0,015 M citrate de sodium), 50 % de formamide, 7 % de sodium dodécyl sulfate (SDS), 10 x Denhardt's, 5 % de dextran sulfate et 1 % d'ADN de sperme de saumon ; (2) hybridation pendant 20 heures à 42°C suivie de 2 lavages de 20 minutes à 20°C en 2 x SSC + 2 % SDS, 1 lavage de 20 minutes à 20°C en 0,1 x SSC + 0,1 % SDS. Le dernier lavage est pratiqué en 0,1 x SSC + 0,1 % SDS pendant 30 minutes à 60°C.

**[0032]** Les Inventeurs décrivent également un fragment représentatif du polynucléotide tel que défini ci-dessus, caractérisé en ce qu'il est susceptible d'être obtenu, soit par l'utilisation d'enzymes de restriction dont les sites de reconnaissance et de coupure sont présents dans ledit polynucléotide tel que défini ci-dessus, soit par amplification à l'aide d'amorces oligonucléotidiques spécifiques dudit polynucléotide tel que défini ci-dessus, soit par transcription *in vitro*, soit par synthèse chimique.

**[0033]** Selon un mode de réalisation avantageux dudit fragment, il est sélectionné dans le groupe constitué par : l'ADNc correspondant à au moins un cadre ouvert de lecture (ORF) choisi parmi : ORF1a, ORF1b, ORF-S, ORF-E, ORF-M, ORF-N, ORF3, ORF4, ORF7 à ORF11, ORF 13 et ORF 14, et l'ADNc correspondant aux extrémités 5' ou 3' non-codantes dudit polynucléotide.

**[0034]** Selon une disposition avantageuse de ce mode de réalisation, ledit fragment présente une séquence sélectionnée dans le groupe constitué par :

- les séquences SEQ ID NO : 2 et 4 représentant l'ADNc correspondant à l'ORF-S qui code pour la protéine S,
- les séquences SEQ ID NO : 13 et 15 représentant l'ADNc correspondant à l'ORF-E qui code pour la protéine E,
- les séquences séquence SEQ ID NO: 16 et 18 représentant l'ADNc correspondant à l'ORF-M qui code pour la protéine M,
- les séquences SEQ ID NO : 36 et 38 représentant l'ADNc correspondant à l'ORF-N qui code pour la protéine N,
- les séquences représentant les ADNc correspondant respectivement : aux ORF1a et ORF1b (ORF1ab, SEQ ID NO : 31), aux ORF3 et ORF4 (SEQ ID NO : 7, 8), aux ORF 7 à 11 (SEQ ID NO : 19, 20), à l'ORF13 (SEQ ID NO : 32) et à l'ORF 14 (SEQ ID NO : 34), et
- les séquences représentant les ADNc correspondant respectivement aux extrémités 5'(SEQ ID NO : 39 et 72) et 3' non-codantes (SEQ ID NO : 40, 73) dudit polynucléotide.

**[0035]** Les Inventeurs décrivent également un fragment de l'ADNc codant pour la protéine S, tel que défini ci-dessus, caractérisé en ce qu'il présente une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO : 5 et 6 (fragments Sa et Sb).

**[0036]** Les Inventeurs décrivent également un fragment de l'ADNc correspondant aux ORF1a et ORF1b tel que défini ci-dessus, caractérisé en ce qu'il présente une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO : 41 à 54 (fragments L0 à L12).

**[0037]** Les Inventeurs décrivent également un fragment du polynucléotide tel que défini ci dessus, caractérisé en ce qu'il présente au moins 15 bases ou paires de bases consécutives de la séquence du génome de ladite souche incluant au moins une de celles situées en position 7979, 16622, 19064, 23220, 24872, 25298 et 26857. De préférence, il s'agit d'un fragment de 20 à 2500 bases ou paires de bases, de manière préférée de 20 à 400.

**[0038]** La présente invention a également pour objet un fragment du polynucléotide tel que défini ci-dessus, caractérisé en ce qu'il comprend les ou est constitué des bases ou de paires de bases consécutives des positions 7919 à 23220, 7919 à 25298, 16622 à 23220, 19064 à 23220, 16622 à 25298, 19064 à 25298, 23220 à 24872, 23220 à 26857, 24872 à 25298, ou 25298 à 26857.

**[0039]** Les Inventeurs décrivent également des amorces d'au moins 18 bases aptes à amplifier un fragment du génome d'un coronavirus associé au SRAS ou de l'équivalent ADN de celui-ci.

**[0040]** Selon un mode de réalisation desdites amorces, elles sont sélectionnées dans le groupe constitué par :

- la paire d'amorces n° 1 correspondant respectivement aux positions 28507 à 28522 (amorce sens, SEQ ID NO : 60) et 28774 à 28759 (amorce anti-sens, SEQ ID NO : 61) de la séquence du polynucléotide tel que défini ci-dessus,
- la paire d'amorces n° 2 correspondant respectivement aux positions 28375 à 28390 (amorce sens, SEQ ID NO : 62) et 28702 à 28687 (amorce anti-sens, SEQ ID NO : 63) de la séquence du polynucléotide tel que défini ci-dessus., et
- la paire d'amorces constituée des amorces SEQ ID Nos 55 et 56.

**[0041]** Les Inventeurs décrivent également une sonde apte à détecter la présence du génome d'un coronavirus associé au SRAS ou d'un fragment de celui-ci, caractérisée en ce qu'elle est sélectionnée dans le groupe constitué par : les fragments tels que définis ci-dessus et les fragments correspondant aux positions suivantes de la séquence du polynucléotide tel que défini ci-dessus : 28561 à 28586, 28588 à 28608, 28541 à 28563 et 28565 à 28589 (SEQ ID NO : 64 à 67).

**[0042]** Les sondes et amorces telles que définies ci-dessus peuvent être marquées directement ou indirectement par un composé radioactif ou non radioactif par des méthodes bien connues de l'Homme du Métier, afin d'obtenir un signal détectable et/ou quantifiable. Parmi les isotopes radioactifs utilisés, on peut citer le <sup>32</sup>P, le <sup>33</sup>P, le <sup>35</sup>S, le <sup>3</sup>H ou l'<sup>125</sup>I. Les entités non radioactives sont sélectionnées parmi les ligands tels que la biotine, l'avidine, la streptavidine, la digoxigénine, les haptènes, les colorants, les agents luminescents tels que les agents radioluminescents, chémoluminescents, bioluminescents, fluorescents, phosphorescents.

**[0043]** Les Inventeurs décrivent des sondes et des amorces marquées dérivées des séquences précédentes.

**[0044]** De telles sondes et amorces sont utiles pour le diagnostic de l'infection par un coronavirus associé au SRAS.

**[0045]** Les Inventeurs décrivent également une méthode de détection d'un coronavirus associé au SRAS, à partir d'un échantillon biologique, laquelle méthode est caractérisée en ce qu'elle comprend au moins :

- (a) l'extraction d'acides nucléiques présents dans ledit échantillon biologique,
- (b) l'amplification d'un fragment de l'ORF-N par RT-PCR à l'aide d'une paire d'amorces telle que définie ci-dessus, et
- (c) la détection par tout moyen approprié des produits d'amplifications obtenus en (b).

**[0046]** Les produits d'amplifications (amplicons) en (b) sont de 268 pb pour la paire d'amorces n° 1 et de 328 pb pour la paire d'amorces n°2.

**[0047]** Selon un mode de mise en oeuvre avantageux dudit procédé, l'étape (c) de détection est réalisée à l'aide d'au moins une sonde correspondant aux positions 28561 à 28586, 28588 à 28608, 28541 à 28563 et 28565 à 28589 de la séquence du polynucléotide tel que défini ci-dessus.

**[0048]** De préférence, le génome du coronavirus associé au SRAS est détecté et éventuellement quantifié par PCR en temps réel, à l'aide de la paire d'amorces n°2 et des sondes correspondant aux positions 28541 à 28563 et 28565 à 28589 marquées avec des composés différents, notamment des agents fluorescents différents.

**[0049]** La RT-PCR en temps réel qui met en oeuvre cette paire d'amorces et cette sonde est très sensible puisqu'elle permet de détecter 10<sup>2</sup> copies d'ARN et jusqu'à 10 copies d'ARN, elle est en outre fiable et reproductible.

**[0050]** Les Inventeurs décrivent des polydésoxyribonucléotides et des poly-ribonucléotides simple-brin, double-brin et tripe-brin correspondant à la séquence du génome de la souche isolée de coronavirus et de ses fragments tels que définis ci-dessus, ainsi qu'à leurs séquences complémentaires, sens ou anti-sens, notamment les ARN et les ADNc correspondant à la séquence du génome et de ses fragments tels que définis ci-dessus.

**[0051]** Les Inventeurs décrivent également des fragments d'amplification obtenus à l'aide d'amorces spécifiques du génome de la souche purifiée ou isolée tel que défini ci-dessus, notamment à l'aide d'amorces et de paires d'amorces telles que définies ci-dessus, des fragments de restriction constitués par ou comprenant la séquence des fragments tels que définis ci-dessus, des fragments obtenus par transcription *in vitro* à partir d'un vecteur contenant la séquence SEQ ID NO : 1 ou un fragment tel que défini ci-dessus, ainsi que des fragments obtenus par synthèse chimique. Des exemples de fragments de restriction sont déduits de la carte de restriction de la séquence SEQ ID NO : 1 illustrée par la figure 13. Lesdits fragments sont, soit sous forme de fragments isolés, soit sous forme de mélanges de fragments. Les Inventeurs décrivent également des fragments modifiés, par rapport aux précédents, par enlèvement, ou addition de nucléotides dans une proportion d'environ 15 %, par rapport à la longueur des fragments ci-dessus et/ou modifiés au niveau de la nature des nucléotides, dès lors que les fragments nucléotidiques modifiés conservent une capacité

d'hybridation avec les séquences d'ARN génomiques ou antigénomiques de l'isolat tel que défini ci-dessus.

**[0052]** Les molécules d'acide nucléique telles que définies ci-dessus sont obtenues par les méthodes classiques, connues en elles-mêmes, en suivant les protocoles standards tels que ceux décrits dans *Current Protocols in Molecular Biology* (Frederick M. AUSUBEL, 2000, Wiley and son Inc, Library of Congress, USA). Par exemple, elles peuvent être

obtenues par amplification d'une séquence nucléique par PCR ou RT-PCR ou bien par synthèse chimique totale ou partielle.

**[0053]** Les Inventeurs décrivent également une puce ou filtre à ADN ou à ARN, caractérisé en ce qu'il comprend au moins un polynucléotide ou l'un de ses fragments tels que définis ci-dessus.

**[0054]** Les puces ou filtres à ADN ou à ARN tels que définis ci-dessus sont préparés par les méthodes classiques, connues en elles-mêmes, comme par exemple greffage chimique ou électrochimique d'oligonucléotides sur support de verre ou de nylon.

**[0055]** Les Inventeurs décrivent également un vecteur de clonage et/ou d'expression recombinant, notamment un plasmide, un virus, un vecteur viral ou un phage comprenant un fragment d'acide nucléique tel que défini ci-dessus. De préférence, ledit vecteur recombinant est un vecteur d'expression dans lequel ledit fragment d'acide nucléique est placé sous le contrôle d'éléments régulateurs de la transcription et de la traduction appropriés. En outre, ledit vecteur peut comprendre des séquences (étiquettes ou tag) fusionnées en phase avec l'extrémité 5' et/ou 3' dudit insert, utiles pour l'immobilisation, et/ou la détection et/ou la purification de la protéine exprimée à partir dudit vecteur.

**[0056]** Ces vecteurs sont construits et introduits dans des cellules hôtes par les méthodes classiques d'ADN recombinant et de génie génétique, qui sont connues en elles-mêmes. De nombreux vecteurs dans lesquels on peut insérer une molécule d'acide nucléique d'intérêt afin de l'introduire et de la maintenir dans une cellule hôte, sont connus en eux-mêmes ; le choix d'un vecteur approprié dépend de l'utilisation envisagée pour ce vecteur (par exemple répllication de la séquence d'intérêt, expression de cette séquence, maintien de la séquence sous forme extrachromosomique ou bien intégration dans le matériel chromosomique de l'hôte), ainsi que de la nature de la cellule hôte.

**[0057]** Ledit plasmide est notamment sélectionné parmi les plasmides suivants :

- le plasmide, dénommé SARS-S, compris dans la souche bactérienne déposée sous le n° I-3059, le 20 juin 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient la séquence d'ADNc codant pour la protéine S de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, laquelle séquence correspondant aux nucléotides des positions 21406 à 25348 (SEQ ID NO : 4), en référence à la séquence Genbank AY274119.3,
- le plasmide, dénommé SARS-S1, compris dans la souche bactérienne déposée sous le n° I-3020, le 12 mai 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient un fragment 5' de la séquence d'ADNc codant pour la protéine S de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, lequel fragment correspondant aux nucléotides des positions 21406 à 23454 (SEQ ID NO :5), en référence à la séquence Genbank AY274119.3 Tor2,
- le plasmide, dénommé SARS-S2, compris dans la souche bactérienne déposée sous le n° I-3019, le 12 mai 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient un fragment 3' de la séquence d'ADNc codant pour la protéine S de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, lequel fragment correspondant aux nucléotides des positions 23322 à 25348 (SEQ ID NO :6), en référence à la séquence Genbank n° d'accès AY274119.3,
- le plasmide, dénommé SARS-SE, compris dans la souche bactérienne déposée sous le n° I-3126, le 13 novembre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient l'ADNc correspondant à la région située entre l'ORF-S et l'ORF-E et chevauchant l'ORF-E de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle région correspondant aux nucléotides des positions 25110 à 26244 (SEQ ID NO :8), en référence à la séquence Genbank n° d'accès AY274119.3,
- le plasmide, dénommé SARS-E, compris dans la souche bactérienne déposée sous le n° I-3046, le 28 mai 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient la séquence d'ADNc codant pour la protéine E de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle séquence correspondant aux nucléotides des positions 26082 à 26413 (SEQ ID NO :15), en référence à la séquence Genbank n° d'accès AY274119.3,
- le plasmide, dénommé SARS-M ; compris dans la souche bactérienne déposée sous le n° I-3047, le 28 mai 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient la séquence d'ADNc codant pour la protéine M de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus ; laquelle séquence correspondant aux nucléotides des positions 26330 à 27098 (SEQ ID NO :18), en référence à la séquence Genbank n° d'accès AY274119.3,
- le plasmide dénommé SARS-MN, compris dans la souche bactérienne déposée sous le n° I-3125, le 13 novembre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris

Cedex 15 ; il contient la séquence d'ADNc correspondant à la région située entre l'ORF-M et l'ORF-N de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589 et prélevée à Hanoi, telle que définie ci-dessus, laquelle séquence correspondant aux nucléotides des positions 26977 à 28218 (SEQ ID NO :20), en référence à la séquence Genbank n° d'accès AY274119.3,

- le plasmide dénommé SARS-N, compris dans la souche bactérienne déposée sous le n° I-3048, le 5 juin 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient l'ADNc codant pour la protéine N de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle séquence correspondant aux nucléotides des positions 28054 à 29430 (SEQ ID NO :38), en référence à la séquence Genbank n° d'accès AY274119.3 ; ainsi ce plasmide comprend un insert de séquence SEQ ID NO :38 et est compris dans une souche bactérienne qui et qu'il a été déposée sous le n° I-3048, le 5 juin 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15,
- le plasmide dénommé SARS-5'NC, compris dans la souche bactérienne déposée sous le n° I- 3124, le 7 novembre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient l'ADNc correspondant à l'extrémité 5'non codante du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle séquence correspondant aux nucléotides des positions 1 à 204 (SEQ ID NO :39), en référence à la séquence Genbank n° d'accès AY274119.3,
- le plasmide dénommé SARS-3'NC, compris dans la souche bactérienne déposée sous le n° I-3123 le 7 novembre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15. ; il contient la séquence d'ADNc correspondant à l'extrémité 3'non codante du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle séquence correspondant à celle située entre le nucléotide en position 28933 à 29727 (SEQ ID NO :40), en référence à la séquence Genbank n° d'accès AY274119.3, se termine par une série de nucléotides a.,
- le plasmide d'expression dénommé pIV2.3N, contenant un fragment d'ADNc codant pour une fusion C-terminale de la protéine N (SEQ ID NO : 37) avec une étiquette polyhistidine,
- le plasmide d'expression dénommé pIV2.3S<sub>C</sub>, contenant un fragment d'ADNc codant pour une fusion C-terminale du fragment correspondant aux positions 475 à 1193 de la séquence en acides aminés de la protéine S (SEQ ID NO : 3) avec une étiquette polyhistidine,
- le plasmide d'expression pIV2.3S<sub>L</sub>, contenant un fragment d'ADNc codant pour une fusion C-terminale du fragment correspondant aux positions 14 à 1193 de la séquence en acides aminés de la protéine S (SEQ ID NO : 3) avec une étiquette polyhistidine,
- le plasmide d'expression dénommé pIV2.4N, contenant un fragment d'ADNc codant pour une fusion N-terminale de la protéine N (SEQ ID NO : 3) avec une étiquette polyhistidine,
- le plasmide d'expression dénommé pIV2.4S<sub>C</sub> ou pIV2.4S<sub>1</sub>, contenant un insert codant pour une fusion N-terminale du fragment correspondant aux positions 475 à 1193 de la séquence en acides aminés de la protéine S (SEQ ID NO : 3) avec une étiquette polyhistidine, et
- le plasmide d'expression dénommé pIV2.4S<sub>L</sub> contenant un fragment d'ADNc codant pour une fusion N-terminale du fragment correspondant aux positions 14 à 1193 de la séquence en acides aminés de la protéine S (SEQ ID NO : 3) avec une étiquette polyhistidine.

**[0058]** Selon une disposition avantageuse du plasmide d'expression tel que défini ci-dessus, il est compris dans une souche bactérienne qui a été déposée sous le n° I- 3117, le 23 octobre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15.

**[0059]** Selon une autre disposition avantageuse du plasmide d'expression tel que défini ci-dessus, il est compris dans une souche bactérienne qui a été déposée sous le n° I- 3118, le 23 octobre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15.

**[0060]** Selon une autre disposition du plasmide d'expression tel que défini ci-dessus, il est compris dans une souche bactérienne qui a été déposée à la CNCM, 25 rue du Docteur Roux, 75724 Paris Cedex 15 sous les numéros suivants :

- a) souche n° I- 3118, déposée le 23 octobre 2003,
- b) souche n° I- 3019, déposée le 12 mai 2003,
- c) souche n° I- 3020, déposée le 12 mai 2003,
- d) souche n° I-3059, déposée le 20 juin 2003,
- e) souche n° I-3323, déposée le 22 novembre 2004,
- f) souche n° I-3324, déposée le 22 novembre 2004,
- g) souche n° I-3326, déposée le 1<sup>er</sup> décembre 2004,
- h) souche n° I-3327, déposée le 1<sup>er</sup> décembre 2004,
- i) souche n° I-3332, déposée le 1<sup>er</sup> décembre 2004,

- j) souche n° I-3333, déposée le 1<sup>er</sup> décembre 2004,
- k) souche n° I-3334, déposée le 1<sup>er</sup> décembre 2004,
- l) souche n° I-3335, déposée le 1<sup>er</sup> décembre 2004,
- m) souche n° I-3336, déposée le 1<sup>er</sup> décembre 2004,
- n) souche n° I-3337, déposée le 1<sup>er</sup> décembre 2004,
- o) souche n° I-3338, déposée le 2 décembre 2004,
- p) souche n° I-3339, déposée le 2 décembre 2004,
- q) souche n° I-3340, déposée le 2 décembre 2004,
- r) souche n° I-3341, déposée le 2 décembre 2004.

**[0061]** Les Inventeurs décrivent également un insérat d'acide nucléique d'origine virale, caractérisé en ce qu'il est contenu dans l'une quelconque des souches telles que définies ci-dessus en a)-r).

**[0062]** Les Inventeurs décrivent également un acide nucléique comportant un gène synthétique permettant une expression optimisée de la protéine S dans des cellules eucaryotes, caractérisé en ce qu'il possède la séquence SEQ ID NO: 140.

**[0063]** Les Inventeurs décrivent également un vecteur d'expression comportant un acide nucléique comportant un gène synthétique permettant une expression optimisée de la protéine S, lequel vecteur est contenu dans la souche bactérienne déposée à la CNCM, le 1<sup>er</sup> décembre 2004, sous le n° I-3333.

**[0064]** Selon un mode de réalisation dudit vecteur d'expression, il s'agit d'un vecteur viral, sous forme de particule virale ou sous forme de génome recombinant.

**[0065]** Selon une disposition avantageuse de ce mode de réalisation, il s'agit d'une particule virale recombinante ou d'un génome viral recombinant susceptible d'être obtenu par transfection d'un plasmide selon les alinéas g), h) et k) à r) tels que définis ci-dessus, dans un système cellulaire approprié, c'est-à-dire, par exemple, des cellules transfectées avec un ou plusieurs autres plasmide(s), destinés à trancomplémenter certaines fonctions du virus délétées dans le vecteur et nécessaires à la formation des particules virales.

**[0066]** On entend ici par « famille de la protéine S » la protéine S complète, son ectodomaine et des fragments de cet ectodomaine qui sont de préférence produits dans un système eucaryote.

**[0067]** Les Inventeurs décrivent également un vecteur lentiviral codant pour un polypeptide de la famille de la protéine S, telle que définie ci-dessus.

**[0068]** Les Inventeurs décrivent également un virus rougeole recombinant codant pour un polypeptide de la famille de la protéine S, telle que définie ci-dessus.

**[0069]** Les Inventeurs décrivent également un virus vaccine recombinant codant pour un polypeptide de la famille de la protéine S, telle que définie ci-dessus.

**[0070]** Les Inventeurs décrivent également l'utilisation d'un vecteur selon les alinéas e) à r) tels que définis ci-dessus, ou d'un vecteur comportant un gène synthétique de la protéine S, tel que défini ci-dessus, pour la production, en système eucaryote, de la protéine S du coronavirus associé au SRAS, ou d'un fragment de cette protéine.

**[0071]** Les Inventeurs décrivent également une méthode de production de la protéine S en système eucaryote, comportant une étape de transfection de cellules eucaryotes en culture par un vecteur choisi parmi les vecteurs contenus dans les souches bactériennes mentionnées aux alinéas e) à r) ci-dessus ou un vecteur comportant un gène synthétique permettant une expression optimisée de la protéine S.

**[0072]** Les Inventeurs décrivent également une banque d'ADNc caractérisée en ce qu'elle comprend des fragments tels que définis ci-dessus, en particulier des fragments d'amplification ou des fragments de restriction, clonés dans un vecteur recombinant, notamment un vecteur d'expression (banque d'expression).

**[0073]** Les Inventeurs décrivent également des cellules, notamment des cellules procaryotes, modifiées par un vecteur recombinant tel que défini ci-dessus.

**[0074]** Les Inventeurs décrivent également une cellule eucaryote génétiquement modifiée exprimant une protéine ou un polypeptide tels que définis ci-dessus. Bien évidemment, les termes "cellule eucaryote génétiquement modifiée" ne désignent pas une cellule modifiée par un virus sauvage.

**[0075]** Selon un mode de réalisation avantageux de ladite cellule, elle est susceptible d'être obtenue par transfection par l'un quelconque des vecteurs mentionnés aux alinéas k) à n) ci-dessus.

**[0076]** Selon une disposition avantageuse de ce mode de réalisation, il s'agit de la cellule FRhK4-Ssol-30, déposée à la CNCM le 22 novembre 2004, sous le n° I-3325.

**[0077]** Les vecteurs recombinants tels que définis ci-dessus et les cellules transformées par lesdits vecteurs d'expression sont avantageusement utilisés pour la production des protéines et des peptides correspondants. Les banques d'expression dérivées desdits vecteurs, ainsi que les cellules transformées par lesdites banques d'expression sont avantageusement utilisées pour identifier les épitopes immunogènes (épitopes B et T) des protéines du coronavirus associé au SRAS.

**[0078]** Les Inventeurs décrivent également des protéines et des peptides purifiées ou isolées, caractérisés en ce qu'ils

sont codés par le polynucléotide ou l'un de ses fragments tels que définis ci-dessus.

**[0079]** Selon un mode de réalisation avantageux, ladite protéine est sélectionnée dans le groupe constitué par :

- la protéine S de séquence SEQ ID NO :3 ou son ectodomaine
- la protéine E de séquence SEQ ID NO :14
- la protéine M de séquence SEQ ID NO :17
- la protéine N de séquence SEQ ID NO : 37
- les protéines codées par les ORFs : ORF1a, ORF1b, ORF3, ORF4 et ORF7 à ORF11, ORF 13 et ORF 14 de séquence respectivement, SEQ ID NO :74, 75, 10, 12, 22, 24, 26, 28, 30, 33 et 35.

**[0080]** On utilisera ci-après indifféremment les termes « ectodomaine de la protéine S » et « forme soluble de la protéine S ».

**[0081]** Selon un mode de réalisation avantageux, ledit polypeptide est constitué des acides aminés correspondant aux positions 1 à 1193 de la séquence en acides aminés de la protéine S

**[0082]** Selon un autre mode de réalisation avantageux, ledit peptide est sélectionné dans le groupe constitué par :

- a) les peptides correspondant aux positions 14 à 1193 et 475 à 1193 de la séquence en acides aminés de la protéine S,
- b) les peptides correspondant aux positions 2 à 14 (SEQ ID NO : 69) et 100 à 221 de la séquence en acides aminés de la protéine M ; ces peptides correspondent respectivement à l'ectodomaine et à l'endodomaine de la protéine M, et
- c) les peptides correspondant aux positions 1 à 12 (SEQ ID NO : 70) et 53 à 76 (SEQ ID NO : 71) de la séquence en acides aminés de la protéine E ; ces peptides correspondent respectivement à l'ectodomaine et à l'extrémité C-terminale de la protéine E, et
- d) les peptides de 5 à 50 acides aminés consécutifs, de préférence de 10 à 30 acides aminés, inclus ou chevauchant partiellement ou totalement la séquence des peptides tels que définis en a), b) ou c).

**[0083]** Les Inventeurs décrivent également un peptide caractérisé en ce qu'il présente une séquence de 7 à 50 acides aminés incluant un résidu d'acide aminé sélectionné dans le groupe constitué par :

- l'alanine située en position 2552 de la séquence en acides aminés de la protéine codée par l'ORF1a.
- la sérine située en position 577 de la séquence en acides aminés de la protéine S de la souche de SARS-CoV telle que définie ci-dessus,
- la glycine en position 11 de la séquence en acides aminés de la protéine codée par l'ORF3 de la souche de SARS-CoV telle que définie ci-dessus,
- la sérine en position 154 de la séquence en acides aminés de la protéine M de la souche de SARS-CoV telle que définie ci-dessus.

**[0084]** Les Inventeurs décrivent également un anticorps ou un fragment d'anticorps polyclonal ou monoclonal, susceptible d'être obtenu par immunisation d'un animal avec un vecteur recombinant tel que défini ci-dessus, une banque d'ADNc telle que définie ci-dessus ou bien une protéine ou un peptide tels que définis ci-dessus, caractérisé en ce qu'il se lie avec l'une au moins des protéines codées par le SARS-CoV telles que définies ci-dessus.

**[0085]** Les Inventeurs décrivent des anticorps polyclonaux, des anticorps monoclonaux, les anticorps chimériques tels que les anticorps humanisés, ainsi que leurs fragments (Fab, Fv, scFv).

**[0086]** Les Inventeurs décrivent également un hybridome produisant un anticorps monoclonal contre la protéine N, caractérisé en ce qu'il est choisi parmi les hybridomes suivants :

- l'hybridome produisant l'anticorps monoclonal 87, déposé à la CNCM le 1<sup>er</sup> décembre 2004 sous le numéro I-3328,
- l'hybridome produisant l'anticorps monoclonal 86, déposé à la CNCM le 1<sup>er</sup> décembre 2004 sous le numéro I-3329,
- l'hybridome produisant l'anticorps monoclonal 57, déposé à la CNCM le 1<sup>er</sup> décembre 2004 sous le numéro I-3330, et
- l'hybridome produisant l'anticorps monoclonal 156, déposé à la CNCM le 1<sup>er</sup> décembre 2004 sous le numéro I-3331.

**[0087]** Les Inventeurs décrivent également un anticorps ou fragment d'anticorps polyclonal ou monoclonal dirigé contre la protéine N, caractérisé en ce qu'il est produit par un hybridome tel que défini ci-dessus.

**[0088]** On entend par anticorps chimérique, relativement à un anticorps d'une espèce animale particulière ou d'une classe particulière d'anticorps, un anticorps comprenant tout ou partie d'une chaîne lourde et/ou d'une chaîne légère d'un anticorps d'une autre espèce animale ou d'une autre classe d'anticorps.

**[0089]** On entend par anticorps humanisé une immunoglobuline humaine dans laquelle les résidus des CDRs (*Complementary-Determining Regions*) qui forment le site de liaison à l'antigène sont remplacés par ceux d'un anticorps monoclonal non-humain possédant la spécificité, l'affinité ou l'activité recherchées. Par comparaison avec les anticorps

non-humains, les anticorps humanisés sont moins immunogènes et possèdent une demi-vie prolongée chez l'Homme car ils ne possèdent qu'une faible proportion de séquences non-humaines étant donné que la quasi-totalité des résidus des régions FR (Framework) et de la région constante (Fc) de ces anticorps sont ceux d'une séquence consensus d'immunoglobulines humaines.

**[0090]** Les Inventeurs décrivent également une puce ou filtre à protéine, caractérisé en ce qu'il comprend une protéine, un peptide ou bien un anticorps tels que définis ci-dessus.

**[0091]** Les puces à protéine sont préparées par les méthodes classiques, connues en elles-mêmes. Parmi les supports appropriés sur lesquels peuvent être immobilisés des protéines, on peut citer ceux en matière plastique ou en verre, notamment sous la forme de microplaques.

**[0092]** La présente invention a également pour objet des réactifs dérivés de la souche isolée de coronavirus associé au SRAS, issue du prélèvement répertorié sous le n° 031589, utiles pour l'étude et le diagnostic de l'infection provoquée par un coronavirus associé au SRAS, lesquels réactifs sont sélectionnés dans le groupe constitué par une souche isolée de coronavirus et un polynucléotide tels que définis ci-dessus.

**[0093]** Les Inventeurs décrivent également des réactifs dérivés de la souche isolée de coronavirus associé au SRAS, issue du prélèvement répertorié sous le n° 031589, utiles pour l'étude et le diagnostic de l'infection provoquée par un coronavirus associé au SRAS, lesquels réactifs sont sélectionnés dans le groupe constitué par :

- (a) une paire d'amorces, une sonde ou une puce à ADN telles que définies ci-dessus,
- (b) un vecteur recombinant ou une cellule modifiée tels que définis ci-dessus,
- (c) une protéine ou un peptide tel que défini ci-dessus,
- (d) un anticorps ou fragment d'anticorps tels que définis ci-dessus, et
- (e) une puce à protéine telle que définie ci-dessus.

**[0094]** Ces différents réactifs sont préparés et utilisés selon les techniques classiques de biologie moléculaire et d'immunologie, en suivant les protocoles standards tels que ceux décrits dans Current Protocols in Molecular Biology (Frederick M. AUSUBEL, 2000, Wiley and Son Inc., Library of Congress, USA), dans Current Protocols in Immunology (John E. Coligan, 2000, Wiley and Son Inc. Library of Congress, USA) et dans Antibodies : A Laboratory Manual (E. Howell and D Lane, Cold Spring Harbor Laboratory, 1988).

**[0095]** Les fragments d'acide nucléique selon l'invention sont préparés et utilisés selon les techniques classiques telles que définies ci-dessus. Les peptides et les protéines sont préparés par les techniques d'ADN recombinant, connues de l'Homme du métier, notamment à l'aide des vecteurs recombinants tels que définis ci-dessus. Alternativement, les peptides peuvent être préparés par les techniques classiques de synthèse en phase solide ou liquide, connues de l'Homme du métier.

**[0096]** Les anticorps polyclonaux sont préparés par immunisation d'un animal approprié avec une protéine ou un peptide tels que définis ci-dessus, éventuellement couplé à la KLH ou à l'albumine et/ou associé à un adjuvant approprié tel que l'adjuvant de Freund (complet ou incomplet) ou l'hydroxyde d'alumine ; après obtention d'un titre en anticorps satisfaisant, les anticorps sont récoltés par prélèvement du sérum des animaux immunisés et enrichis en IgG par précipitation, selon les techniques classiques, puis les IgG spécifiques des protéines du SARS-CoV sont éventuellement purifiées par chromatographie d'affinité sur une colonne appropriée sur laquelle sont fixés ledit peptide ou ladite protéine, tels que définis ci-dessus, de façon à obtenir une préparation d'IgG monospécifiques.

**[0097]** Les anticorps monoclonaux sont produits à partir d'hybridomes obtenus par fusion de lymphocytes B d'un animal immunisé par une protéine ou un peptide tels que définis ci-dessus avec des myélomes, selon la technique de Köhler et Milstein (Nature, 1975, 256, 495-497) ; les hybridomes sont cultivés *in vitro*, notamment dans des fermenteurs ou produits *in vivo*, sous forme d'ascite ; alternativement lesdits anticorps monoclonaux sont produits par génie génétique comme décrit dans le brevet américain US 4,816,567.

**[0098]** Les anticorps humanisés sont produits par des méthodes générales comme celles décrites dans la Demande Internationale WO 98/45332.

**[0099]** Les fragments d'anticorps sont produits à partir des régions V<sub>H</sub> et V<sub>L</sub> clonées, à partir des ARNm d'hybridomes ou de lymphocytes spléniques d'une souris immunisée ; par exemple, les fragments Fv, scFv ou Fab sont exprimés à la surface de phages filamenteux selon la technique de Winter et Milstein (Nature, 1991, 349, 293-299) ; après plusieurs étapes de sélection, les fragments d'anticorps spécifiques de l'antigène sont isolés et exprimés dans un système d'expression approprié, par les techniques classiques de clonage et d'expression d'ADN recombinant.

**[0100]** Les anticorps ou leur fragments tels que définis ci-dessus, sont purifiés par les techniques classiques connues de l'Homme du métier, telles que la chromatographie d'affinité.

**[0101]** La présente invention a en outre pour objet l'utilisation d'un produit sélectionné dans le groupe constitué par : une souche isolée de coronavirus et un polynucléotide tels que définis ci-dessus, pour la préparation d'un réactif de détection et éventuellement de génotypage/sérotypage, d'un coronavirus associé au SRAS,

**[0102]** Les Inventeurs décrivent en outre l'utilisation d'un produit sélectionné dans le groupe constitué par : une paire

d'amorces, une sonde, une puce à ADN, un vecteur recombinant, une cellule modifiée, une protéine ou un peptide, un anticorps ou un fragment d'anticorps et une puce à protéine tels que définis ci-dessus, pour la préparation d'un réactif de détection et éventuellement de génotypage/sérotypage, d'un coronavirus associé au SRAS.

**[0103]** Les protéines et les peptides, qui sont aptes à être reconnus et/ou à induire la production d'anticorps spécifiques du coronavirus associé au SRAS, sont utiles pour le diagnostic de l'infection par un tel coronavirus ; l'infection est détectée, par une technique appropriée- notamment EIA, ELISA, RIA, immunofluorescence-, à partir d'un échantillon biologique prélevé chez un individu susceptible d'être infecté.

**[0104]** Selon une disposition avantageuse de ladite utilisation, lesdites protéines sont sélectionnées dans le groupe constitué par les protéines S, E, M et/ou N et les peptides tels que définis ci-dessus.

**[0105]** Les protéines S, E, M et/ou N et les peptides dérivés de ces protéines tels que définis ci-dessus, par exemple la protéine N, sont utilisées pour le diagnostic indirect d'une infection à coronavirus associé au SRAS (diagnostic sérologique ; détection d'anticorps spécifiques du SARS-CoV), notamment par une méthode immunoenzymatique (ELISA).

**[0106]** Les anticorps et les fragments d'anticorps tels que définis ci-dessus, notamment ceux dirigés contre les protéines S, E, M et/ou N et les peptides dérivés tels que définis ci-dessus, sont utiles pour le diagnostic direct d'une infection à coronavirus associé au SRAS ; la détection de protéine(s) du SARS-CoV est réalisée par une technique appropriée, notamment EIA, ELISA, RIA, immunofluorescence à partir d'un échantillon biologique prélevé chez un individu susceptible d'être infecté.

**[0107]** Les Inventeurs décrivent également une méthode de détection d'un coronavirus associé au SRAS, à partir d'un échantillon biologique, laquelle méthode est caractérisée en ce qu'elle comprend au moins :

(a) la mise en contact dudit échantillon biologique avec au moins un anticorps ou un fragment d'anticorps, une protéine, un peptide ou bien une puce ou un filtre à protéine ou à peptide tels que définis ci-dessus, et

(b) la révélation par tout moyen approprié des complexes antigène-anticorps formés en (a), par exemple par EIA, ELISA, RIA, ou par immunofluorescence.

**[0108]** Selon un mode de mise en oeuvre avantageux dudit procédé l'étape (a) comprend :

(a<sub>1</sub>) la mise en contact dudit échantillon biologique avec au moins un premier anticorps ou un fragment d'anticorps qui est fixé sur un support approprié, notamment une microplaque,

(a<sub>2</sub>) le lavage de la phase solide, et

(a<sub>3</sub>) l'addition d'au moins un second anticorps ou un fragment d'anticorps, différent du premier, ledit anticorps ou fragment d'anticorps étant éventuellement marqué de façon appropriée.

**[0109]** Ce procédé qui permet de capturer les particules virales présentes dans l'échantillon biologique est également dénommé procédé d'immunocapture.

**[0110]** Par exemple :

- l'étape (a<sub>1</sub>) est réalisée avec au moins un premier anticorps monoclonal ou polyclonal ou un fragment de ceux-ci, dirigé contre la protéine S, M, et/ou E, et/ou un peptide correspondant à l'ectodomaine de l'une de ces protéines (peptides M2-14 ou E1-12)

- l'étape (a<sub>3</sub>) est réalisée avec au moins un anticorps ou un fragment d'anticorps dirigé contre un autre épitope de la même protéine ou de préférence contre une autre protéine, de manière préférée contre une protéine interne telle que la nucléoprotéine N ou l'endodomaine de la protéine E ou M, de manière encore plus préférée il s'agit d'anticorps ou de fragments d'anticorps dirigés contre la protéine N qui est très abondante dans la particule virale ; lorsqu'un anticorps ou un fragment d'anticorps dirigé contre une protéine interne (N) ou contre l'endodomaine des protéines E ou M est utilisé, le dit anticorps est incubé en présence de détergent, comme le Tween 20 par exemple, à des concentrations de l'ordre de 0,1 %.

- l'étape (b) de révélation des complexes antigène-anticorps formés est réalisée, soit directement à l'aide d'un second anticorps marqué par exemple avec de la biotine ou une enzyme appropriée telle que la peroxydase ou la phosphatase alcaline, soit indirectement à l'aide d'un sérum anti-immunoglobulines marqué comme ci-dessus. Les complexes ainsi formés sont révélés à l'aide d'un substrat approprié.

**[0111]** Selon une mise en oeuvre préférée de cet aspect de l'invention, l'échantillon biologique est mélangé à l'anticorps monoclonal de révélation préalablement à sa mise en contact avec les anticorps monoclonaux de capture. Le cas échéant, le mélange sérum-anticorps de révélation est incubé pendant au moins 10 minutes à température ambiante avant d'être appliqué sur la plaque.

**[0112]** La présente invention a également pour objet un test d'immunocapture destiné à détecter une infection par le



coronavirus associé au SRAS par détection de la nucléoprotéine native (protéine N), en particulier caractérisé en ce que l'anticorps utilisé pour la capture de la nucléoprotéine virale native est un anticorps monoclonal spécifique de la région centrale et/ou d'un épitope conformationnel.

**[0113]** Selon un mode de réalisation dudit test, l'anticorps utilisé pour la capture de la protéine N est l'anticorps monoclonal Acm87, produit par l'hybridome déposé à la CNM le 1<sup>er</sup> décembre 2004 sous le numéro I-3328.

**[0114]** Selon un autre mode de réalisation dudit test d'immunocapture, l'anticorps utilisé pour la capture de la protéine N est l'anticorps monoclonal Acm86, produit par l'hybridome déposé à la CNM le 1<sup>er</sup> décembre 2004 sous le numéro I-3329.

**[0115]** Selon un autre mode de réalisation dudit test d'immunocapture, les anticorps monoclonaux Acm86 et Acm87 sont utilisés pour la capture de la protéine N.

**[0116]** Dans les tests d'immunocapture selon l'invention, on peut utiliser pour la révélation de la protéine N, l'anticorps monoclonal Acm57, produit par l'hybridome déposé à la CNM le 1<sup>er</sup> décembre 2004 sous le numéro I-3330, ledit anticorps étant conjugué à une molécule ou une particule révélatrice.

**[0117]** Conformément audit test d'immunocapture, une combinaison des anticorps Acm57 et Acm87, conjugués à une molécule ou une particule révélatrice, est utilisée pour la révélation de la protéine N.

**[0118]** Une molécule révélatrice peut être un atome radioactif, un colorant, une molécule fluorescente, un fluorophore, une enzyme ; une particule révélatrice peut être, par exemple : de l'or colloïdal, une particule magnétique ou une bille de latex.

**[0119]** La présente invention a également pour objet un réactif de détection d'un coronavirus associé au SRAS, caractérisé en ce qu'il est sélectionné dans le groupe constitué par une souche isolée de coronavirus telle que définie ci-dessus et un polynucléotide tel que défini ci-dessus.

**[0120]** Les Inventeurs décrivent également un réactif de détection d'un coronavirus associé au SRAS, caractérisé en ce qu'il est sélectionné dans le groupe constitué par :

- (a) une paire d'amorces ou une sonde telles que définies ci-dessus,
- (b) un vecteur recombinant tel que défini ci-dessus ou une cellule modifiée telle que définie ci-dessus,
- (c) un anticorps ou un fragment d'anticorps tel que défini ci-dessus,
- (d) une combinaison d'anticorps comprenant les anticorps monoclonaux Acm86 et/ou Acm87, et l'anticorps monoclonal Acm57, telle que définie ci-dessus,
- (e) une puce ou un filtre tels que définis ci-dessus.

**[0121]** Les Inventeurs décrivent également une méthode de détection d'une infection par un coronavirus associé au SRAS, à partir d'un échantillon biologique, par ELISA IgG indirect utilisant la protéine N, laquelle méthode est caractérisée en ce que les plaques sont sensibilisées par une solution de protéine N à une concentration comprise entre 0,5 et 4 µg/mL, de préférence 2 µg/mL, dans un tampon PBS 10mM pH 7,2, rouge de phénol à 0,25mL /L.

**[0122]** Les Inventeurs décrivent, en outre, une méthode de détection d'une infection par un coronavirus associé au SRAS, à partir d'un échantillon biologique, par ELISA double épitope, caractérisée en ce que le sérum à tester est mélangé à l'antigène de révélation, ledit mélange étant ensuite mis au contact de l'antigène fixé sur un support solide.

**[0123]** Selon une variante des tests de détection des coronavirus associé au SRAS, ces tests combinent un ELSA utilisant la protéine N, et un autre ELSA utilisant la protéine S, tel que décrit plus bas.

**[0124]** Les Inventeurs décrivent aussi un complexe immunitaire formé d'un anticorps ou d'un fragment d'anticorps polyclonal ou monoclonal tel que défini ci-dessus, et d'une protéine ou d'un peptide du coronavirus associé au SRAS.

**[0125]** La présente invention a en outre pour objet un kit de détection d'un coronavirus associé au SRAS, caractérisé en ce qu'il comprend au moins un réactif sélectionné dans le groupe constitué par une souche isolée de coronavirus et un polynucléotide.

**[0126]** Les Inventeurs décrivent en outre un kit de détection d'un coronavirus associé au SRAS, caractérisé en ce qu'il comprend au moins un réactif sélectionné dans le groupe constitué par : une paire d'amorces, une sonde, une puce à ADN ou à ARN, un vecteur recombinant, une cellule modifiée, une protéine ou un peptide, un anticorps, et une puce à protéine tels que définis ci-dessus.

**[0127]** Les Inventeurs décrivent en outre, une composition immunogène, caractérisée en ce qu'elle comprend au moins un produit sélectionné dans le groupe constitué par :

- a) une protéine ou un peptide tels que définis ci-dessus,
- b) un polynucléotide de type ADN ou ARN ou l'un de ses fragments représentatifs tels que définis ci-dessus, de séquence choisie parmi :
  - (i) la séquence SEQ ID NO : 1 ou son équivalent ARN
  - (ii) la séquence hybridant dans des conditions de forte stringence avec la séquence SEQ ID NO : 1,

- (iii) la séquence complémentaire de la séquence SEQ ID NO : 1 ou de la séquence hybridant dans des conditions de forte stringence avec la séquence SEQ ID NO : 1,
- (iv) la séquence nucléotidique d'un fragment représentatif du polynucléotide tel que défini en (i), (ii) ou (iii),
- (v) la séquence telle que définie en (i), (ii), (iii) ou (iv), modifiée, et

5

- c) un vecteur d'expression recombinant comprenant un polynucléotide tel que défini en b), et
- d) une banque d'ADNc telle que définie ci-dessus,

10

ladite composition immunogène étant capable d'induire une immunité humorale ou cellulaire protectrice spécifique du coronavirus associé au SRAS, notamment la production d'un anticorps dirigé contre un épitope spécifique du coronavirus associé au SRAS.

15

**[0128]** Les protéines et les peptides tels que définis ci-dessus, notamment les protéines S, M, E et/ou N et les peptides dérivés, ainsi que les molécules d'acide nucléique (ADN ou ARN) codant lesdites protéines ou lesdits peptides, sont de bons candidats vaccin et peuvent être utilisées dans des compositions immunogènes pour la production d'un vaccin contre le coronavirus associé au SRAS.

20

**[0129]** Selon un mode de réalisation avantageux des compositions définies ci-dessus, elles contiennent en outre, au moins un véhicule pharmaceutiquement acceptable et éventuellement des substances porteuses et/ou des adjuvants.

**[0130]** Les véhicules pharmaceutiquement acceptables, les substances porteuses et les adjuvants sont ceux classiquement utilisés.

25

**[0131]** Les adjuvants sont avantageusement choisis dans le groupe constitué par des émulsions huileuses, de la saponine, des substances minérales, des extraits bactériens, de l'hydroxyde d'alumine et le squalène.

**[0132]** Les substances porteuses sont avantageusement sélectionnées dans le groupe constitué par des liposomes unilamellaires, des liposomes multilamellaires, des micelles de saponine ou des microsphères solides de nature saccharidique ou aurifère.

30

**[0133]** Les compositions telles que définies ci-dessus, sont administrées par voie générale, notamment intramusculaire ou sous-cutanée ou bien par voie locale notamment nasale (aérosol).

**[0134]** Les Inventeurs décrivent également l'utilisation d'une protéine ou d'un peptide isolé ou purifié présentant une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO : 3, 10, 12, 14, 17, 22, 24, 26, 28, 30, 33, 35, 37, 69, 70, 71, 74 et 75 pour former un complexe immun avec un anticorps dirigé spécifiquement contre un épitope du coronavirus associé au SRAS.

35

**[0135]** Les Inventeurs décrivent également un complexe immun formé d'une protéine ou d'un peptide isolé ou purifié présentant une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO : 3, 10, 12, 14, 17, 22, 24, 26, 28, 30, 33, 35, 37, 69, 70, 71, 74 et 75, et d'un anticorps dirigé spécifiquement contre un épitope du coronavirus associé au SRAS.

**[0136]** Les Inventeurs décrivent également l'utilisation d'une protéine ou d'un peptide isolé ou purifié présentant une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO : 3, 10, 12, 14, 17, 22, 24, 26, 28, 30, 33, 35, 37, 69, 70, 71, 74 et 75 pour induire la production d'un anticorps capable de reconnaître spécifiquement un épitope du coronavirus associé au SRAS.

40

**[0137]** Les Inventeurs décrivent également l'utilisation d'un polynucléotide isolé ou purifié présentant une séquence sélectionnée dans le groupe constitué par les séquences SEQ ID NO : 1, 2, 4, 7, 8, 13, 15, 16, 18, 19, 20, 31, 36 et 38 pour induire la production d'un anticorps dirigé contre la protéine codée par ledit polynucléotide et capable de reconnaître spécifiquement un épitope du coronavirus associé au SRAS

**[0138]** Les Inventeurs décrivent également des anticorps monoclonaux reconnaissant la protéine S native d'un coronavirus associé au SRAS.

45

**[0139]** Les Inventeurs décrivent également l'utilisation d'une protéine ou d'un polypeptide de la famille de la protéine S, telle que définie ci-dessus, ou d'un anticorps reconnaissant la protéine S native, tel que défini ci-dessus, pour détecter une infection par un coronavirus associé au SRAS, à partir d'un échantillon biologique.

**[0140]** Les Inventeurs décrivent également une méthode de détection d'une infection par un coronavirus associé au SRAS, à partir d'un échantillon biologique, caractérisée en ce que la détection est effectuée par ELISA utilisant la protéine

50

S recombinante, exprimée dans un système eucaryote.

**[0141]** Selon un mode de mise en oeuvre avantageux de ladite méthode, il s'agit d'une méthode par ELISA double épitope, et le sérum à tester est mélangé à l'antigène de révélation, ledit mélange étant ensuite mis au contact de l'antigène fixé sur un support solide.

55

**[0142]** Les Inventeurs décrivent aussi un complexe immun formé d'un anticorps ou d'un fragment d'anticorps monoclonal reconnaissant la protéine S native, et d'une protéine ou d'un peptide du coronavirus associé au SRAS.

**[0143]** Les Inventeurs décrivent également un complexe immun formé d'une protéine ou d'un polypeptide de la famille de la protéine S, telle que définie ci-dessus, et d'un anticorps dirigé spécifiquement contre un épitope du coronavirus associé au SRAS.

**[0144]** Les Inventeurs décrivent en outre un kit ou coffret de détection d'un coronavirus associé au SRAS, caractérisé en ce qu'il comprend au moins un réactif sélectionné dans le groupe constitué par : une protéine ou polypeptide de la famille de la protéine S, telle que définie ci-dessus, un acide nucléique codant pour une protéine ou polypeptide de la famille de la protéine S, telle que définie ci-dessus, une cellule exprimant une protéine ou polypeptide de la famille de la protéine S, telle que définie ci-dessus, ou un anticorps reconnaissant la protéine S native d'un coronavirus associé au SRAS.

**[0145]** Les Inventeurs décrivent une composition immunogène et/ou vaccinale, caractérisée en ce qu'elle comprend un polypeptide ou une protéine recombinante de la famille de la protéine S, telle que définie ci-dessus, obtenu dans un système d'expression eucaryote.

**[0146]** Les Inventeurs décrivent également une composition immunogène et/ou vaccinale, caractérisée en ce qu'elle comprend un vecteur ou virus recombinant, exprimant une protéine ou un polypeptide de la famille de la protéine S, telle que définie ci-dessus.

**[0147]** Outre les dispositions qui précèdent, l'invention comprend encore d'autres dispositions, qui ressortiront de la description qui va suivre, qui se réfère à des exemples de mise en oeuvre du polynucléotide représentant le génome de la souche de SARS-CoV issue du prélèvement répertorié sous le numéro 031589, et des fragments d'ADNc dérivés objets de la présente invention, ainsi qu'au Tableau I présentant la liste des séquences :

**Tableau I : Liste des séquences**

numéro d'identification	Séquence	Position de l'ADNc en référence à Genbank AY274119.3	Numéro de dépôt à la CNCM du plasmide correspondant
SEQ ID NO : 1	génom de la souche issue du prélèvement 031589	-	-
SEQ ID NO : 2	ORF-S*	21406-25348	-
SEQ ID NO : 3	Protéine S	-	-
SEQ ID NO : 4	ORF-S**	21406-25348	I-3059
SEQ ID NO : 5	fragment Sa	21406-23454	I-3020
SEQ ID NO : 6	fragment Sb	23322-25348	I-3019
SEQ ID NO : 7	ORF-3+ORF-4*	25110-26244	-
SEQ ID NO : 8	ORF-3+ORF-4**	25110-26244	I-3126
SEQ ID NO : 9	ORF3	-	-
SEQ ID NO : 10	Protéine ORF-3	-	-
SEQ ID NO : 11	ORF4	-	-
SEQ ID NO : 12	Protéine ORF-4	-	-
SEQ ID NO : 13	ORF-E*	26082-26413	-
SEQ ID NO : 14	Protéine E	-	-
SEQ ID NO : 15	ORF-E**	26082-26413	I-3046
SEQ ID NO : 16	ORF-M*	26330-27098	-
SEQ ID NO : 17	Protéine M	-	-
SEQ ID NO : 18	ORF-M**	26330-27098	I-3047
SEQ ID NO : 19	ORF7 à 11*	26977-28218	-
SEQ ID NO : 20	ORF7 à 11**	26977-28218	I-3125
SEQ ID NO : 21	ORF7	-	-
SEQ ID NO : 22	Protéine ORF7	-	-
SEQ ID NO : 23	ORF8	-	-

# EP 1 694 829 B1

(suite)

	numéro d'identification	Séquence	Position de l'ADNc en référence à Genbank AY274119.3	Numéro de dépôt à la CNCM du plasmide correspondant
5	SEQ ID NO : 24	Protéine ORF8	-	-
	SEQ ID NO : 25	ORF9	-	-
	SEQ ID NO : 26	Protéine ORF9	-	-
10	SEQ ID NO : 27	ORF10	-	-
	SEQ ID NO : 28	Protéine ORF10	-	-
	SEQ ID NO : 29	ORF11	-	-
15	SEQ ID NO : 30	Protéine ORF11	-	-
	SEQ ID NO : 31	OrF1ab	265-21485	-
	SEQ ID NO : 32	ORF13	28130-28426	-
	SEQ ID NO : 33	Protéine ORF13	-	-
20	SEQ ID NO : 34	ORF14	-	-
	SEQ ID NO : 35	Protéine ORF14	28583-28795	-
	SEQ ID NO : 36	ORF-N*	28054-29430	-
25	SEQ ID NO : 37	Protéine N	-	-
	SEQ ID NO : 38	ORF-N**	28054-29430	I-3048
	SEQ ID NO : 39	5'non-codante**	1-204	I-3124
	SEQ ID NO : 40	3'non-codante**	28933-29727	I-3123
30	SEQ ID NO : 41	<b>ORF1ab</b> Fragment L0	30-500	-
	SEQ ID NO : 42	Fragment L1	211-2260	-
35	SEQ ID NO : 43	Fragment L2	2136-4187	-
	SEQ ID NO : 44	Fragment L3	3892-5344	-
	SEQ ID NO : 45	Fragment L4b	4932-6043	-
	SEQ ID NO : 46	Fragment L4	5305-7318	-
40	SEQ ID NO : 47	Fragment L5	7275-9176	-
	SEQ ID NO : 48	Fragment L6	9032-11086	-
	SEQ ID NO : 49	Fragment L7	10298-12982	-
45	SEQ ID NO : 50	Fragment L8	12815-14854	-
	SEQ ID NO : 51	Fragment L9	14745-16646	-
	SEQ ID NO : 52	Fragment L10	16514-18590	-
	SEQ ID NO : 53	Fragment L11	18500-20602	-
50	SEQ ID NO : 54	Fragment L12	20319-22224	-
	SEQ ID NO : 55	Amorce N sens	-	-
	SEQ ID NO : 56	Amorce N antisens	-	-
55	SEQ ID NO : 57	Amorce S <sub>C</sub> sens	-	-
	SEQ ID NO : 58	Amorce S <sub>L</sub> sens	-	-
	SEQ ID NO : 59	Amorce S <sub>C</sub> e t S <sub>L</sub> antisens	-	-

# EP 1 694 829 B1

(suite)

numéro d'identification	Séquence	Position de l'ADNc en référence à Genbank AY274119.3	Numéro de dépôt à la CNCM du plasmide correspondant
SEQ ID NO : 60	Amorce sens série 1	28507-28522	-
SEQ ID NO : 61	Amorce antisens série 1	28774-28759	
SEQ ID NO : 62	Amorce sens série 2	28375-28390	-
SEQ ID NO : 63	Amorce antisens série 2	28702-28687	-
SEQ ID NO : 64	Sonde 1/série 1	28561-28586	-
SEQ ID NO : 65	Sonde 2/série 1	28588-28608	-
SEQ ID NO : 66	Sonde 1/série 2	28541-28563	-
SEQ ID NO : 67	Sonde 2/série 2	28565-28589	-
SEQ ID NO : 68	Amorce ancre 14T		
SEQ ID NO : 69	Peptide M2-14	-	-
SEQ ID NO : 70	Peptide E1-12	-	-
SEQ ID NO : 71	Peptide E53-76	-	-
SEQ ID NO : 72	5'non-codante*	1-204	-
SEQ ID NO: 73	3'non-codante*	28933-29727	-
SEQ ID NO : 74	Protéine ORF1a	-	-
SEQ ID NO : 75	Protéine ORF1b	-	-
SEQ ID NO:76-139	Amorces		
SEQ ID NO:140	Pseudogène de S		
SEQ ID NO:141-148	amorces		
SEQ ID NO:149	Aa1-13 de S		
SEQ ID NO:150	polypeptide		
SEQ ID NO:151-158	amorces		
* produit d'amplification PCR (amplicon) ** insert cloné dans le plasmide déposé à la CNCM ainsi qu'aux dessins annexés dans lesquels :			

- la figure 1 illustre l'analyse par Western-blot de l'expression *in vitro* des protéines recombinantes N, S<sub>C</sub> et S<sub>L</sub> à partir des vecteurs d'expression pIVEX. Piste 1 : pIV2.3N. Piste 2 : pIV2.3S<sub>C</sub>. Piste 3 : pIV2.3S<sub>L</sub>. Piste 4 : pIV2.4N. Piste 5 : pIV2.4S<sub>1</sub> ou pIV2.4S<sub>C</sub>. Piste 6 : pIV2.4S<sub>L</sub>. L'expression de la protéine GFP exprimée à partir du même vecteur est utilisée comme contrôle.
- la figure 2 illustre l'analyse par électrophorèse en gel de polyacrylamide en conditions dénaturantes (SDS-PAGE) et coloration au bleu de Coomassie, de l'expression *in vivo* de la protéine N à partir des vecteurs d'expression pIVEX. La souche d'*E.coli* BL21(DE3)pDIA17 transformée par les vecteurs pIVEX recombinants est cultivée à 30°C dans du milieu LB, en présence ou en l'absence d'inducteur (IPTG 1mM). Piste 1 : pIV2.3N Piste 2 : pIV2.4N.
- la figure 3 illustre l'analyse par électrophorèse en gel de polyacrylamide en conditions dénaturantes (SDS-PAGE) et coloration au bleu de Coomassie, de l'expression *in vivo* des polypeptides S<sub>L</sub> et S<sub>C</sub> à partir des vecteurs d'expression pIVEX. La souche d'*E.coli* BL21(DE3)pDIA17 transformée par les vecteurs pIVEX recombinants est cultivée

à 30°C dans du milieu LB, en présence ou en l'absence d'inducteur (IPTG 1mM). Piste 1 : pIV2.3S<sub>C</sub> Piste 2 : pIV2.3S<sub>L</sub>. Piste 3 : pIV2.4S<sub>1</sub> Piste 4 : pIV2.4S<sub>L</sub>.

- la figure 4 illustre l'activité antigénique des protéines N, S<sub>L</sub> et S<sub>C</sub> recombinantes produites dans la souche *E. coli* BL21(DE3)pDIA17 transformée par les vecteurs pIVEX recombinants. A : électrophorèse (SDS-PAGE) des lysats bactériens. B et C : Western-blot avec les sérums, provenant d'un même patient infecté par le SARS-CoV, prélevés respectivement 8 jours (B : sérum M12) et 29 jours-(C : sérum M13) après le début des symptômes du SRAS. Piste 1 : pIV2.3N. Piste 2 : pIV2.4N. Piste 3 : pIV2.3S<sub>C</sub>. Piste 4 : pIV2.4 S<sub>1</sub>. Piste 5 : pIV2.3S<sub>L</sub>. Piste 6 : pIV2.4S<sub>L</sub>
- la figure 5 illustre la purification sur colonne Ni-NTA agarose de la protéine N recombinante produite dans la souche *E. coli* BL21(DE3)pDIA17 à partir du vecteur pIV2.3N. Piste 1 : Extrait bactérien total. Piste 2 : Extrait soluble. Piste 3 : Extrait insoluble. Piste 4 : Extrait déposé sur la colonne Ni-NTA. Piste 5 : protéines non-retenues. Piste 6 : Fractions du pic 1. Piste 7 : Fractions du pic 2.
- la figure 6 illustre la purification de la protéine S<sub>C</sub> recombinante à partir des corps d'inclusions produits dans la souche *E. coli* BL21(DE3)pDIA17 transformée par le pIV2.4S<sub>1</sub>.A. Traitement au Triton X-100 (2%) : Piste 1 : Extrait bactérien total. Piste 2 : Extrait soluble. Piste 3 : Extrait insoluble. Piste 4 : Surnageant après traitement au Triton X-100 (2 %). Pistes 5 et 6 : Culot après traitement au Triton X-100 (2 %).B : Traitement à l'urée 4M, 5M, 6M et 7M des extraits solubles et insolubles.
- la figure 7 représente l'immunoempreinte réalisée à l'aide d'un lysat de cellules infectées par le SARS-CoV et d'un sérum de patient atteint de pneumopathie atypique.
- la figure 8 représente des immunoempreintes réalisées à l'aide d'un lysat de cellules infectées par le SARS-CoV et d'immunsérums de lapins spécifiques de la nucléoprotéine N (A) et de la protéine de spicule S (B). I.S. : sérum immun. p.i. : sérum pré-immun. L'immunsérum anti-N a été utilisé au 1/50000 et l'immunsérum anti-S au 1/10000.
- la figure 9 illustre la réactivité en ELISA des sérums polyclonaux monospécifiques de lapin dirigés contre la protéine N ou le fragment court de la protéine S (S<sub>C</sub>), vis-à-vis des protéines recombinantes correspondantes utilisées pour l'immunisation. A : lapins P13097, P13081, et P13031 immunisés avec la protéine N recombinante purifiée. B : lapins P11135, P13042, et P14001 immunisés avec une préparation de corps d'inclusions correspondants au fragment court de la protéine S (S<sub>C</sub>). I.S. : sérum immun. p.i. : sérum pré-immun.
- la figure 10 illustre la réactivité en ELISA de la protéine N recombinante purifiée, vis-à-vis de sérum de patients atteints de pneumonie atypique causée par le SARS-CoV. Figure 10a : plaques ELISA préparées avec la protéine N à la concentration de 4 µg/ml et 2 µg/ml. Figure 10b : plaque ELISA préparée avec la protéine N à la concentration de 1 µg/ml. Les sérums désignés A, B, D, E, F, G, H correspondent à ceux du Tableau IV.
- la figure 11 illustre l'amplification par RT-PCR de quantités décroissantes d'ARN synthétique du gène N du SARS-CoV (10<sup>7</sup> à 1 copie), à l'aide des couples d'amorces n° 1 (N/+28507,N/-28774) (A) et n° 2 (N/+28375,N/-28702) (B). T : amplification réalisée en l'absence d'ARN. MW : marqueur d'ADN.
- la figure 12 illustre l'amplification par RT-PCR en temps réel d'ARN synthétique du gène N du SARS-CoV : des quantités décroissantes d'ARN synthétique en répliquat (repli. ; pistes 16 à 29) ainsi que de l'ARN viral dilué au 1/20x10<sup>-4</sup> (piste 32) ont été amplifiés par RT-PCR en temps réel à l'aide du kit "Light Cycler RNA Amplification Kit Hybridization Probes" et des couples d'amorces et de sondes de la série n° 2, dans les conditions décrites à l'exemple 8.
- la figure 13.1 à 13.70 (figure 13.1 à 13.70) représente la carte de restriction de la séquence SEQ ID NO : 1 correspondant à l'équivalent ADN du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le numéro 031589.
- la figure 14 montre le résultat du test de sérologie SRAS par ELISA N indirect (1ère série de sérums testés)
- la figure 15 montre le résultat du test de sérologie SRAS par ELISA N indirect (2ème série de sérums testés)
- la figure 16 présente le résultat du test de sérologie SRAS par ELISA N double épitope (1 ère série de sérums testés)
- la figure 17 montre le résultat du test de sérologie SRAS par ELISA N double épitope (2ème série de sérums testés)
- la figure 18 illustre le test de réactivité des anticorps monoclonaux anti-N par ELISA sur la nucléoprotéine N native du SRAS-CoV. Les anticorps ont été testés sous la forme de surnageants de culture d'hybridomes par un ELISA indirect utilisant un lysat irradié de cellules VeroE6 infectées par le SRAS-CoV comme antigène (courbes lysat SRAS). Un contrôle négatif de réactivité est réalisé pour chaque anticorps sur un lysat de cellules VeroE6 non infectées (courbes lysat négatif). Plusieurs anticorps monoclonaux de spécificité connue ont été utilisés comme anticorps témoins négatifs : para1-3 dirigé contre les antigènes des virus parainfluenza de type 1-3 (Bio-Rad) et grippe B dirigé contre les antigènes du virus de la grippe de type B (Bio-Rad).
- la figure 19 illustre le test de réactivité des anticorps monoclonaux anti-N du SRAS-CoV par ELISA sur les antigènes natifs du coronavirus humain 229E (HCoV-229E). Les anticorps ont été testés sous la forme de surnageants de culture d'hybridomes par un test ELISA indirect utilisant un lysat de cellules MRC-5 infectées par le coronavirus humain 229E comme antigène (courbes lysat 229E). Un contrôle négatif d'immunoréactivité est réalisé pour chaque anticorps sur un lysat de cellules MRC-5 non infectées (courbes lysat négatif). L'anticorps monoclonal 5-11H.6 dirigé contre la protéine S du coronavirus humain 229E (Sizun et al. 1998, J. Virol. Met. 72 : 145-152) est utilisé

comme anticorps témoin positif. Les anticorps para-3 dirigé contre les antigènes des virus parainfluenza de type 1-3 (Bio-Rad) et grippe B dirigé contre les antigènes du virus de la grippe de type B (Bio-Rad) ont été ajoutés au panel des anticorps monoclonaux testés.

- la figure 20 montre un test de réactivité des anticorps monoclonaux anti-N du SRAS-CoV par western blot sur la nucléoprotéine N native du SRAS-CoV dénaturée. Un lysat de cellules VeroE6 infectées par le SRAS-CoV a été préparé dans le tampon de dépôt selon Laemmli et mis à migrer dans un gel SDS à 12% de polyacrylamide puis les protéines ont été transférées sur membrane de PVDF. Les anticorps monoclonaux anti-N testés ont été utilisés pour l'immunoessai à la concentration de 0.05 µg/ml. La révélation est faite avec des anticorps anti-IgG(H+L) de souris couplés à la peroxydase (NA93IV, Amersham) et le système ECL+. Deux anticorps monoclonaux ont été utilisés comme témoins négatifs de réactivité : grippe B dirigé contre les antigènes du virus de la grippe de type B (Bio-Rad) et para-1-3 dirigé contre les antigènes des virus parainfluenza de type 1-3 (Bio-Rad).
- la figure 21 présente les plasmides d'expression en cellules de mammifères de la protéine S du SRAS-CoV. Le cDNA de la S du SRAS-CoV a été inséré entre les sites BamH1 et Xho1 du plasmide d'expression pcDNA3.1(+) (Clontech) pour obtenir le plasmide pcDNA-S et entre les sites Nhe1 et Xho1 du plasmide d'expression pCI (Promega) pour obtenir le plasmide pCI-S. Les séquences WPRE et CTE ont été insérées dans chacun des deux plasmides pcDNA-S et pCI-S entre les sites Xho1 et Xba1 pour obtenir respectivement les plasmides pcDNA-S-CTE, pcDNA-S-WPRE, pCI-S-CTE et pCI-S-WPRE.

SP : peptide signal prédit (aa 1-13) avec le logiciel signalP v2.0 (Nielsen et al., 1997, Protein Engineering, 10 : 1-6)  
 TM : région transmembranaire prédite (aa 1196-1218) avec le logiciel TMHMM v2.0 (Sonnhammer et al., 1998, Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182, AAAI Press). Il faut noter que les acides aminés W1194 et P1195 font possiblement partie de la région transmembranaire avec des probabilités respectives de 0,13 et 0,42

P-CMV : promoteur immédiat/précoce du cytomégalovirus. BGH pA : signal de polyadénylation du gène de l'hormone de croissance bovine

SV40 late pA : signal de polyadénylation tardif du virus SV40

SD/SA : sites donneur et accepteur d'épissage

WPRE : séquences du "Woodchuck Hepatitis Virus posttranscriptional regulatory element" du virus de l'hépatite de la marmotte

CTE : séquences du "constitutive transport element" du rétrovirus simien de Mason-Pfizer

- la figure 22 illustre l'expression de la protéine S après transfection de cellules VeroE6. Des extraits cellulaires ont été préparés 48 heures après transfection de cellules VeroE6 par les plasmides pcDNA, pcDNA-S, pCI et pCI-S. Des extraits cellulaires ont également été préparés 18 heures après infection par le virus recombinant de la vaccine VV-TF7.3 et transfection par les plasmides pcDNA ou pcDNA-S. A titre de contrôle, des extraits de cellules VeroE6 ont été préparés 8 heures après infection par le SRAS-CoV à une multiplicité d'infection de 3. Ils ont été séparés sur un gel SDS à 8% d'acrylamide et analysés par western blot à l'aide d'un anticorps polyclonal de lapin anti-S et d'un anticorps polyclonal anti-IgG(H+L) de lapin couplé à la peroxydase (NA934V, Amersham). Une échelle de masse moléculaire (kDa) est portée sur la figure.

SRAS-CoV : extrait de cellules VeroE6 infectées par le SRAS-CoV

Mock : extrait contrôle de cellules non infectées

- la figure 23 illustre l'effet des séquences CTE et WPRE sur l'expression de la protéine S après transfection de cellules VeroE6 et 293T. Des extraits cellulaires ont été préparés 48 heures après transfection de cellules VeroE6 (A) ou 293T (B) par les plasmides pcDNA, pcDNA-S, pcDNA-S-CTE, pcDNA-S-WPRE, pCI-S, pCI-S-CTE et pCI-S-WPRE, séparés sur un gel SDS à 8% d'acrylamide et analysés par western blot à l'aide d'un anticorps polyclonal de lapin anti-S et d'un anticorps polyclonal anti-IgG(H+L) de lapin couplé à la peroxydase (NA934V, Amersham). Une échelle de masse moléculaire (kDa) est portée sur la figure.

SRAS-CoV : extrait de cellules VeroE6 préparés 8 heures après infection par le SRAS-CoV à une multiplicité d'infection de 3.

Mock : extrait contrôle de cellules VeroE6 non infectées

- la figure 24 présente des vecteurs lentiviraux défectifs à DNA flap central pour l'expression de la S du SRAS-CoV. Le cADN de la protéine S du SRAS-CoV a été cloné sous la forme d'un fragment BamH1-Xho1 dans le plasmide pTRIPΔU3-CMV contenant un vecteur lentiviral défectif TRIP à DNA flap central (Sirven et al, 2001, Mol. Ther., 3 : 438-448) pour obtenir le plasmide pTRIP-S. Les cassettes d'expression optimales constituées du promoteur immé-

diat/précoce du virus CMV, d'un signal d'épissage, du cDNA de la S et de l'un ou l'autre des signaux post-transcriptionnels CTE ou WPRE ont été substituées à la cassette EF1 $\alpha$ -EGFP du vecteur d'expression lentiviral déficient à DNA FLAP central TRIP $\Delta$ U3-EF1  $\alpha$  (Sirven et al, 2001, Mol. Ther., 3 :438-448) pour obtenir les plasmides pTRIP-SD/SA-S-CTE et pTRIP-SD/SA-S-WPRE.

SP : peptide signal

TM : région transmembranaire

P-CMV : promoteur immédiat/précoce du cytomégalo virus

P-EF1 $\alpha$  : promoteur du gène EF1 $\alpha$

SD/SA : sites donneur et accepteur d'épissage

WPRE : séquences du "Woodchuck Hepatitis Virus posttranscriptional regulatory element" du virus de l'hépatite de la marmotte

CTE : séquences du "constitutive transport element" du rétrovirus simien de Mason-Pfizer

LTR : « Long terminal repeat »

$\Delta$ U3 : LTR délété des séquences « promoter/enhancer »

cPPT : « polypurine tract cis-active sequence »

CTS : « central termination sequence »

- la figure 25 montre l'analyse par western blot de l'expression de la S du SRAS-CoV par des lignées cellulaires transduites par les vecteurs lentiviraux TRIP-SD/SA-S-WPRE et TRIP-SD/SA-S-CTE. Des extraits cellulaires ont été préparés à partir de lignées FrhK4-S-CTE et FrhK4-S-WPRE établies après transduction par les vecteurs lentiviraux TRIP-SD/SA-S-CTE et TRIP-SD/SA-S-WPRE respectivement. Ils ont été séparés sur un gel SDS à 8% d'acrylamide et analysés par western blot à l'aide d'un anticorps polyclonal de lapin anti-S et d'un conjugué anti-IgG(H+L) de lapin couplé à la peroxidase. Une échelle de masse moléculaire (kDa) est portée sur la figure.

T- : extrait contrôle de cellules FrhK-4

T+ : extrait de cellules FrhK-4 préparées 24h après infection par le SRAS-CoV à une multiplicité d'infection de 3.

- la figure 26 concerne l'analyse de l'expression de polypeptide Ssol par des lignées cellulaires transduites par les vecteurs lentiviraux TRIP-SD/SA-Ssol-WPRE et TRIP-SD/SA-Ssol-CTE. La sécrétion du polypeptide Ssol a été recherchée dans le surnageant d'une série de clones cellulaires isolés après transduction de cellules FRhK-4 par les vecteurs lentiviraux TRIP-SD/SA-Ssol-WPRE et TRIP-SD/SA-Ssol-CTE. 5  $\mu$ l de surnageant dilués au 1/2 dans du tampon de dépôt selon Laemmli ont été analysés par western blot révélé par un anticorps monoclonal anti-FLAG (M2, Sigma) et un conjugué anti-IgG(H+L) de souris couplé à la peroxidase. T- : surnageant de la lignée FRhK-4 parentale. T+ : surnageant de cellules BHK infectées par un virus recombinant de la vaccine exprimant le polypeptide Ssol. La flèche pleine indique le polypeptide Ssol, tandis que la flèche creuse indique une réaction croisée avec une protéine d'origine cellulaire.
- la figure 27 montre les résultats relatifs à l'analyse du polypeptide Ssol purifié

A. 8, 2, 0.5 et 0.125  $\mu$ g de polypeptide recombinant Ssol purifié par chromatographie d'affinité anti-FLAG et gel filtration (G75) ont été séparés sur gel SDS à 8% de polyacrylamide. Le polypeptide Ssol ainsi que des quantités variables de marqueurs de masse moléculaire (MM) ont été révélés par coloration au nitrate d'argent (Gelcode SilverSNAP stain kit II, Pierce).

B. Marqueurs étalons pour l'analyse par spectrométrie de masse SELDI-TOF

IgG : IgG bovine de MM 147300

ConA : conalbumine de MM 77490

HRP : peroxydase de raifort analysée à titre de contrôle et de MM 43240

C. Analyse par spectrométrie de masse (SELDI-TOF) du polypeptide recombinant Ssol.

Les pics A et B correspondent au polypeptide Ssol simplement et doublement chargé.

D. Séquençage de l'extrémité N-terminale du polypeptide recombinant Ssol. 5 cycles de dégradation d'Edman en phase liquide ont été réalisés sur un séquenceur ABI494 (Applied Biosystems).

- la figure 28 illustre l'influence d'un signal d'épissage et des séquences CTE et WPRE sur l'efficacité de l'immunisation génique à l'aide d'ADN plasmidique codant la S du SRAS-CoV



A. Des groupes de 7 souris BALB/c ont été immunisés à deux reprises à 4 semaines d'intervalle à l'aide de 50 µg d'ADN plasmidique de pCI, pcDNA-S, pCI-S, pcDNA-N et pCI-HA.

B. Des groupes de 6 souris BALB/c ont été immunisés à deux reprises à 4 semaines d'intervalle à l'aide de 2 µg, 10 µg ou 50 µg d'ADN plasmidique de pCI, pCI-S, pCI-S-CTE et pCI-S-WPRE.

Les sérums immuns prélevés 3 semaines après la deuxième immunisation ont été analysés par ELISA indirect en utilisant un lysat de cellules VeroE6 infectées par le SRAS-CoV comme antigène. Les titres en anticorps anti-SRAS-CoV sont calculés comme l'inverse de la dilution produisant une DO spécifique de 0,5 après révélation par un anticorps polyclonal anti-IgG de souris couplé à la peroxidase (NA931V, Amersham) et du TMB (KPL).

- la figure 29 montre la séroneutralisation de l'infektivité du SRAS-CoV par les anticorps induits chez la souris après immunisation génique à l'aide d'ADN plasmidique codant la S du SRAS-CoV. Des pools de sérums immuns prélevés 3 semaines après la deuxième immunisation ont été réalisés pour chacun des groupes des expériences décrites dans la figure 28 et évalués pour leur capacité à séroneutraliser l'infektivité de 100 TCID<sub>50</sub> du SRAS-CoV sur cellules FRhK-4. 4 points sont réalisés pour chacune des dilutions de raison 2 testées à partir du 1/20. Le titre séroneutralisant est calculé selon la méthode de Reed et Munsch comme l'inverse de la dilution neutralisant l'infektivité de 2 cupules sur 4.

A. Groupes de souris BALB/c immunisés à deux reprises à 4 semaines d'intervalle à l'aide de 50 µg d'ADN plasmidique de pCI, pcDNA-S, pCI-S, pcDNA-N et pCI-HA. □ : sérum préimmun. ■ : sérum immun.

B. Groupes de souris BALB/c immunisés à deux reprises à 4 semaines d'intervalle à l'aide de 2 µg, 10 µg ou 50 µg d'ADN plasmidique de pCI, pCI-S, pCI-S-CTE et pCI-S-WPRE.

- la figure 30 illustre l'immunoréactivité du polypeptide recombinant Ssol vis à vis de sérums de patients atteints de SRAS. La réactivité de sérums de patients a été analysée par test ELISA indirect contre des phases solides préparées à l'aide du polypeptide recombinant Ssol purifié. Les anticorps de patients réagissant avec la phase solide à une dilution de 1/400 sont révélés par un anticorps polyclonal anti-IgG(H+L) humain couplé à la peroxidase (Amersham NA933V) et du TMB plus H<sub>2</sub>O<sub>2</sub> (KPL). Les sérums de cas probables de SRAS sont identifiés par un numéro d'ordre du Centre National de Référence des Virus Influenzae ainsi que par les initiales du patient et le nombre de jours écoulés depuis le début des symptômes, le cas échéant. Les sérums TV sont des sérums témoins de sujets qui ont été prélevés en France avant l'épidémie de SRAS survenue en 2003.
- La figure 31 montre l'induction d'anticorps dirigés contre le SRAS-CoV après immunisation avec le polypeptide recombinant Ssol. Deux groupe de 6 souris ont été immunisés à 3 semaines d'intervalle avec 10 µg de polypeptide recombinant Ssol (groupe Ssol) adjuvée avec de l'hydroxyde d'aluminium ou, à titre de contrôle, de l'adjuvant seul (groupe mock). Trois immunisations successives ont été réalisées et les sérums immuns ont été prélevés 3 semaines après chacune des trois immunisations (IS1, IS2, IS3). Les sérums immuns ont été analysés par pool pour chacun des 2 groupes par ELISA indirect en utilisant un lysat de cellules VeroE6 infectées par le SRAS-CoV comme antigène. Les titres en anticorps anti-SRAS-CoV sont calculés comme l'inverse de la dilution produisant une DO spécifique de 0,5 après révélation par un anticorps polyclonal anti-IgG de souris couplé à la peroxidase (Amersham) et du TMB (KPL).
- La figure 32 présente l'alignement nucléotidique des séquences du gène synthétique 040530 avec la séquence du gène sauvage de l'isolat 031589 du SRAS-CoV. I-3059 correspond aux nucléotides 21406-25348 de l'isolat 031589 du SRAS-CoV déposé à la C.N.C.M. sous le numéro I-3059 (SEQ ID NO : 4, plasmide pSRAS-S). S-040530 est la séquence du gène synthétique 040530.
- la figure 33 illustre l'utilisation d'un gène synthétique pour l'expression de la S du SRAS-CoV. Des extraits cellulaires préparés 48 heures après transfection de cellules VeroE6 (A) ou 293T (B) par les plasmides pCI, pCI-S, pCI-S-CTE, pCI-S-WPRE et pCI-Ssynth ont été séparés sur un gel SDS à 8% d'acrylamide et analysés par western blot à l'aide d'un anticorps polyclonal de lapin anti-S et d'un anticorps polyclonal anti-IgG(H+L) de lapin couplé à la peroxidase (NA934V, Amersham). Le western blot est révélé par luminescence (ECL+, Amersham) et acquisition sur un dispositif d'imagerie numérique (FluorS, BioRad). Les niveaux d'expression de la protéine S ont été mesurés en quantifiant les 2 bandes majoritaires repérées sur l'image.
- la figure 34 présente un schéma de la construction des virus vaccine recombinants VV-TG-S, VV-TG-Ssol, VV-TN-S et VV-TN-Ssol

A. Les cDNA de la protéine S et du polypeptide Ssol du SRAS-CoV ont été inséré entre les sites BamH1 et Sma1 du plasmide de transfert pTG186 pour obtenir les plasmides pTG-S et pTG-Ssol.

B. Les séquences du promoteur synthétique 480 ont ensuite été substituées à celles du promoteur 7.5 par échange du fragment Nde1-Pst1 des plasmides pTG186poly, pTG-S et pTG-Ssol pour obtenir les plasmides de transfert pTN480, pTN-S et pTN-Ssol.

C. Séquence du promoteur synthétique 480 tel que contenu entre les sites Nde1 et Pst1 des plasmides de transfert de la série pTN. Un site Asc1 a été inséré pour faciliter les manipulations ultérieures. Les sites de restriction ainsi que la séquence du promoteur sont soulignés

D. Les virus recombinants de la vaccine sont obtenus par double recombinaison homologue *in vivo* entre la cassette TK des plasmides de transfert des séries pTG et pTN et le gène TK de la souche Copenhague du virus de la vaccine.

SP : peptide signal prédit (aa 1-13) avec le logiciel signalP v2.0 (Nielsen et al., 1997, Protein Engineering, 10 : 1-6)

TM : région transmembranaire prédite (aa 1196-1218) avec le logiciel TMHMM v2.0 (Sonnhammer et al., 1998, Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182, AAAI Press). Il faut noter que les acides aminés W1194 et P1195 font possiblement partie de la région transmembranaire avec des probabilités respectives de 0,13 et 0,42.

TK-L, TK-R : parties gauche et droite du gène de la thymidine kinase du virus de la vaccine

MCS : site multiple de clonage

PE : promoteur précoce

PL : promoteur tardif

PL synth : promoteur tardif synthétique 480

- la figure 35 illustre l'expression de la protéine S par les virus vaccine recombinants, analysée par western blot. Des extraits cellulaires ont été préparés 18 heures après infection de cellules CV1 par les virus vaccine recombinants VV-TG, VV-TG-S et VV-TN-S à une M.O.I. de 2 (A). A titre de contrôle, des extraits de cellules VeroE6 ont été préparés 8 heures après infection par le SRAS-CoV à une multiplicité d'infection de 2. Des extraits cellulaires ont également été préparés 18 heures après infection de cellules CV1 par les virus vaccine recombinants VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S et VV-TN-Ssol (B). Ils ont été séparés sur des gels SDS à 8% d'acrylamide et analysés par western blot à l'aide d'un anticorps polyclonal de lapin anti-S et d'un anticorps polyclonal anti-IgG(H+L) de lapin couplé à la peroxidase (NA934V, Amersham). « 1µl » et « 10µl » indique les quantités d'extraits cellulaires déposées sur le gel. Une échelle de masse moléculaire (kDa) est portée sur la figure.

SRAS-CoV : extrait de cellules VeroE6 infectées par le SRAS-CoV

Mock : extrait contrôle de cellules non infectées

- la figure 36 montre le résultat d'une analyse par western blot de la sécrétion du polypeptide Ssol par les virus vaccine recombinants.

A. Des surnageants de cellules CV1 infectées par le virus vaccine recombinant VV-TN, différents clones du virus VV-TN-Ssol et par les virus VV-TG-Ssol ou VV-TN-Sflag ont été récoltés 18 heures après infection de cellules CV1 à une M.O.I. de 2.

B. Des surnageants de cellules 293T, FRhK-4, BHK-21 et CV1 infectées en dupliqués (1,2) par le virus vaccine recombinant VV-TN-Ssol à une M.O.I. de 2 ont été récoltés 18 heures après infection. Le surnageant de cellules CV1 infectées par le virus VV-TN a également été récolté à titre de contrôle (M).

Tous les surnageants ont été séparés sur un gel SDS à 8% d'acrylamide selon Laemmli et analysés par western blot à l'aide d'un anticorps monoclonal de souris anti-FLAG et d'un anticorps polyclonal anti-IgG(H+L) de souris couplé à la peroxidase (NA931V, Amersham) (A) ou à l'aide d'un anticorps polyclonal de lapin anti-S et d'un anticorps polyclonal anti-IgG(H+L) de lapin couplé à la peroxidase (NA934V, Amersham) (B).

Une échelle de masse moléculaire (kDa) est portée sur la figure.

- la figure 37 montre l'analyse du polypeptide Ssol, purifié par gel SDS de polyacrylamide 10, 5 et 2 µl de polypeptide recombinant Ssol purifié par chromatographie d'affinité anti-FLAG ont été séparés sur gel SDS en gradient de 4 à 15 % de polyacrylamide. Le polypeptide Ssol ainsi que des quantités variables de marqueurs de masse moléculaire (MM) ont été révélés par coloration au nitrate d'argent (Gelcode SilverSNAP stain kit II, Pierce).
- la figure 38 illustre l'immunoréactivité du polypeptide recombinant Ssol produit par le virus vaccine recombinant VV-TN-Ssol vis-à-vis de sérums de patients atteints de SRAS. La réactivité de sérums de patients a été analysée par test ELISA indirect contre des phases solides préparées à l'aide du polypeptide recombinant Ssol purifié. Les anticorps de patients réagissant avec la phase solide à une dilution de 1/100 et 1/400 sont révélés par un anticorps polyclonal anti-IgG(H+L) humain couplé à la peroxidase (Amersham NA933V) et du TMB plus H2O2 (KPL). Les

sérums de cas probables de SRAS sont identifiés par un numéro d'ordre du Centre National de Référence des Virus Influenzae ainsi que par les initiales du patient et le nombre de jours écoulés depuis le début des symptômes, le cas échéant. Les sérums TV sont des sérums témoins de sujets qui ont été prélevés en France avant l'épidémie de SRAS survenue en 2003.

- la figure 39 montre la réponse en anticorps anti-SRAS-CoV chez la souris après immunisation par les virus vaccine recombinants. Des groupes de 7 souris BALB/c ont été immunisés par voie i.v. à deux reprises à 4 semaines d'intervalle par 106 u.f.p. de virus vaccine recombinants VV-TG, VV-TG-HA, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S, VV-TN-Ssol.

A. Des pools de sérums immuns prélevés 3 semaines après chacune des deux immunisations ont été réalisés pour chacun des groupes et ont été analysés par ELISA indirect en utilisant un lysat de cellules VeroE6 infectées par le SRAS-CoV comme antigène. Les titres en anticorps anti-SRAS-CoV sont calculés comme l'inverse de la dilution produisant une DO spécifique de 0,5 après révélation par un anticorps polyclonal anti-IgG de souris couplé à la peroxidase (NA931V, Amersham) et du TMB (KPL).

B. Les pools de sérums immuns ont été évalués pour leur capacité à séroneutraliser l'infektivité de 100 TCID<sub>50</sub> du SRAS-CoV sur cellules FRhK-4. 4 points sont réalisés pour chacune des dilutions de raison 2 testées à partir du 1/20. Le titre séroneutralisant est calculé selon la méthode de Reed et Munsch comme l'inverse de la dilution neutralisant l'infektivité de 2 cupules sur 4.

- la figure 40 décrit la construction des virus recombinants MVSchw2-SARS-S et MVSchw2-SARS-Ssol.

A. Le vecteur rougeole est un génome complet de la souche vaccinale Schwarz du virus de la rougeole (MV) dans lequel une unité de transcription supplémentaire a été introduite (Combredet, 2003, Journal of Virology, 77 : 11546-11554). L'expression des phases ouvertes de lecture (ORF) supplémentaires est contrôlée par les éléments agissant en cis nécessaires à la transcription, à la formation de la coiffe et à la polyadénylation du transgène, qui ont été copiés des éléments présents à la jonction N/P. 2 vecteurs différents permettent l'insertion entre les gènes P (phosphoprotéine) et M (matrice) d'une part et H (hémagglutinine) et L (polymérase) d'autre part.

B. Les génomes recombinants MVSchw2-SARS-S et MVSchw2-SARS-Ssol du virus de la rougeole ont été construits par l'insertion des ORFs de la protéine S et du polypeptide Ssol au sein d'une unité de transcription supplémentaire localisée entre les gènes P et M du vecteur.

Les différents gènes du virus de la rougeole (MV) sont indiqués : N (nucléoprotéine), PVC (phosphoprotéine et protéine V/C), M (matrice), F (fusion), H (hémagglutinine), L (polymérase). T7 = promoteur de l'ARN polymérase T7, hh = ribozyme hammerhead, T7t = séquence terminatrice de l'ARN polymérase du phage T7, ∂ = ribozyme du virus de l'hépatite ∂, (2), (3) = unités de transcription supplémentaires (ATU).

Taille du génome du MV : 15894 nt.

SP : peptide signal

TM : région transmembranaire

FLAG : étiquette FLAG

- la figure 41 illustre l'expression de la protéine S par les virus rougeole recombinants, analysée par western blot.

Des extraits cytoplasmiques ont été préparés après infection de cellules Vero par différents passages des virus MVSchw2-SARS-S et MVSchw2-SARS-Ssol et le virus sauvage MWSchw à titre de contrôle. Des extraits cellulaires en tampon de dépôt selon Laemmli ont également été préparés 8 heures après infection de cellules VeroE6 par le SRAS-CoV à une multiplicité d'infection de 3. Ils ont été séparés sur un gel SDS à 8% d'acrylamide et analysés par western blot à l'aide d'un anticorps polyclonal de lapin anti-S et d'un anticorps polyclonal anti-IgG(H+L) de lapin couplé à la peroxidase (NA934V, Amersham).

Une échelle de masse moléculaire (kDa) est portée sur la figure.

Pn : nième passage du virus après coculture de cellules 293-3-46 et Vero.

SRAS-CoV : extrait de cellules VeroE6 infectées par le SRAS-CoV

Mock : extrait contrôle de cellules VeroE6 non infectées

- la figure 42 montre l'expression de la protéine S par les virus rougeole recombinants, analysée par immunofluorescence

Des cellules Vero en monocouches sur lamelles de verre ont été infectées par le virus sauvage MWSchw (A) ou les virus MVSchw2-SARS-S (B) et MVSchw2-SARS-Ssol (C). Quand les syncytia ont atteint 30 à 40% de confluence

(A., B.) ou 90-100% (C), les cellules ont été fixées, perméabilisées et marquées par des anticorps polyclonaux de lapins anti-SRAS-CoV et un conjugué anti-IgG(H+L) de lapin couplé au FITC (Jackson).

- la figure 43 illustre l'analyse par western blot de l'immunoréactivité de sérums de lapins dirigés contre les peptides E1-12, E53-76 et M2-14. Le lapin 20047 a été immunisé avec le peptide E1-12 couplé à la KLH. Les lapins 22234 et 22240 ont été immunisés avec le peptide E53-76 couplé à la KLH. Les lapins 20013 et 20080 ont été immunisés avec le peptide M2-14 couplé à la KLH. Les immunosérums ont été analysés par western blot à l'aide d'extraits de cellules infectées par le SRAS-CoV (B) ou à l'aide d'extraits de cellules infectées par un virus recombinant de la vaccine exprimant la protéine E (A) ou M (C) de l'isolat 031589 du SRAS-CoV. Les immunoempreintes ont été révélées à l'aide d'un conjugué anti-IgG(H+L) de lapin couplé à la peroxidase (NA934V, Amersham).

**[0148]** La position des protéines E et M est indiquée par une flèche.

**[0149]** Une échelle de masse moléculaire (kDa) est portée sur la figure.

**[0150]** Il doit être bien entendu, toutefois, que ces exemples sont donnés uniquement à titre d'illustration de l'objet de l'invention, dont ils ne constituent en aucune manière une limitation.

### **Exemple 1 : Clonage et séquençage du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le numéro 031589**

**[0151]** L'ARN de la souche de SARS-CoV a été extrait à partir du prélèvement de lavage bronchoalvéolaire répertorié sous le numéro 031589, effectué sur un patient de l'hôpital français de Hanoi (Vietnam) atteint de SRAS.

**[0152]** L'ARN isolé a été utilisé comme matrice pour amplifier les ADNc correspondant aux différents cadres ouverts de lecture du génome (ORF1a, ORF1b, ORF-S, ORF-E, ORF-M, ORF-N (incluant les ORF-13 et ORF-14), ORF3, ORF4, ORF7 à ORF11), et aux extrémités 5' et 3' non-codantes. Les séquences des amorces et des sondes utilisées pour l'amplification/détection ont été définies d'après la séquence nucléotidique disponible du SARS-CoV.

**[0153]** Dans ce qui suit les amorces et les sondes sont identifiées par : la lettre S, suivie d'une lettre qui indique la région correspondante du génome (L pour l'extrémité 5' incluant ORF1a et ORF1b ; S, M et N pour les ORF-S, ORF-M, ORF-N, SE et MN pour les régions intergéniques correspondantes), puis éventuellement de Fn, Rn, avec n inclus entre 1 et 6 correspondant aux amorces utilisées pour la PCR nichée ou imbriquée (paire F1 + R1 pour la première amplification, paire F2 + R2 pour la deuxième amplification, etc...), puis de +/- ou -/- correspondant à une amorce sens ou antisens et enfin des positions des amorces en référence à la séquence Genbank AY27411.3 ; pour les amorces S et N sens et antisens et les autres amorces sens uniquement, lorsqu'une seule position est indiquée elle correspond à celle de l'extrémité 5' d'une sonde ou d'une amorce d'environ 20 bases ; pour les amorces antisens autres que les amorces S et N, lorsqu'une seule position est indiquée elle correspond à celle de l'extrémité 3' d'une sonde ou d'une amorce d'environ 20 bases.

**[0154]** Les produits d'amplifications ainsi générés ont été séquencés à l'aide d'amorces spécifiques afin de déterminer la séquence complète du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le numéro 031589. Ces produits d'amplification, à l'exception de ceux correspondant aux ORF1a et ORF1b, ont ensuite été clonés dans des vecteurs d'expression afin de produire les protéines virales correspondantes et les anticorps dirigés contre ces protéines, notamment par immunisation à base d'ADN.

### **1. Extraction des ARN**

**[0155]** Les ARN ont été extraits à l'aide du kit *Qlamp viral RNA extraction mini* (QIAGEN) en suivant les recommandations du fabricant. De manière plus précise : 140 µl du prélèvement et 560 µl de tampon AVL ont été mélangés vigoureusement pendant 15 secondes, incubés 10 min à température ambiante puis centrifugés brièvement à vitesse maximale. 560 µl d'éthanol à 100% ont été ajoutés au surnageant et le mélange ainsi obtenu a été agité très vigoureusement pendant 15 sec. 630 µl du mélange ont ensuite été déposés sur la colonne.

**[0156]** La colonne a été placée sur un tube de 2 ml, centrifugée 1 min à 8000 rpm, puis le reste du mélange précédent a été déposé sur la même colonne, centrifugé à nouveau, 1 min à 8000 rpm et la colonne a été transférée sur un tube de 2 ml propre. Ensuite, 500 µl de tampon AW1 ont été ajoutés sur la colonne, puis la colonne a été centrifugée 1 min à 8000 rpm et l'éluat a été éliminé. 500 µl de tampon AW2 ont été ajoutés sur la colonne qui a ensuite été centrifugée 3 min à 14000 rpm et transférée sur un tube de 1,5 ml. Enfin, 60 µl de tampon AVE ont été ajoutés sur la colonne qui a été incubée 1 à 2 min à température ambiante puis centrifugée 1 min à 8000 rpm. L'éluat correspondant à l'ARN purifié a été récupéré et congelé à -20°C.

## 2. Amplification, séquençage et clonage des ADNc

### 2.1) ADNc codant pour la protéine S

**[0157]** Les ARN extraits à partir du prélèvement ont été soumis à une transcription inverse à l'aide d'oligonucléotides hexamériques de séquence aléatoire (pdN6), afin de produire des fragments d'ADNc.

**[0158]** La séquence codant pour la glycoprotéine S du SARS-CoV a été amplifiée sous la forme de deux fragments d'ADN chevauchants : fragment 5' (SRAS-Sa, SEQ ID NO:5) et fragment 3' (SRAS-Sb, SEQ ID NO:6), en réalisant deux amplifications successives à l'aide d'amorces imbriquées. Les amplicons ainsi obtenus ont été séquencés, clonés dans le vecteur plasmidique PCR 2.1-TOPO™ (IN VITROGEN), puis la séquence des ADNc clonés a été déterminée.

#### a) clonage et séquençage des fragments Sa et Sb

##### a<sub>1</sub>) synthèse de l'ADNc

**[0159]** Le mélange réactionnel contenant : ARN (5 µl), H<sub>2</sub>O ppi (3,5 µl), tampon de transcriptase inverse 5X (4 µl), dNTP 5 mM (2 µl), pdN6 100 ug/ml (4 µl), RNasin 40 UI/ul (0,5 µl) et transcriptase inverse AMV-RT, 10 UI/ul, PROMEGA (1 µl) a été incubé dans un thermocycleur dans les conditions suivantes : 45 min à 42°C, 15 min à 55°C, 5 min à 95°C, puis l'ADNc obtenu a été maintenu à +4°C.

##### a<sub>2</sub>) première amplification PCR

**[0160]** Les extrémités 5' et 3' du gène S ont été amplifiées respectivement avec les paires d'amorces S/F1/+ 21350-21372 et S/R1/-/ 23518-23498, S/F3/+ 23258-23277 et S/R3/-/25382-25363. Le mélange réactionnel de 50 µl contenant : ADNc (2 µl), amorces 50 µM (0,5 µl), tampon 10 X (5 µl), dNTP 5 mM (2 µl), Taq Expand High Fidelity, Roche (0,75 µl) et H<sub>2</sub>O (39,75 µl) a été amplifié dans un thermocycleur, dans les conditions suivantes : une étape initiale de dénaturation à 94°C pendant 2 min a été suivie de 40 cycles comprenant : une étape de dénaturation à 94°C pendant 30 sec, une étape d'hybridation à 55°C pendant 30 sec puis une étape d'élongation à 72°C pendant 2 min 30 sec, avec 10 sec d'élongation supplémentaire à chaque cycle, puis d'une étape finale d'élongation à 72°C pendant 5 min.

##### a<sub>3</sub>) deuxième amplification PCR

**[0161]** Les produits de la première amplification PCR (amplicons 5' et 3') ont subi une seconde étape d'amplification PCR (PCR nichée) dans des conditions identiques à celles de la première amplification, avec les paires d'amorces S/F2/+21406-21426 et S/R2/-/23454-23435, et S/F4/+23322-23341 et S/R4/-/25348-25329, respectivement pour l'amplicon 5' et l'amplicon 3'.

##### a<sub>4</sub>) clonage et séquençage des fragments Sa et Sb

**[0162]** Les amplicons Sa (extrémité 5') et Sb (extrémité 3') ainsi obtenus ont été purifiés à l'aide du kit *QIAquick PCR purification* (QIAGEN), en suivant les recommandations du fabricant, puis ils ont été clonés dans le vecteur PCR2.1-TOPO (kit Invitrogen), pour donner les plasmides dénommés SRAS-S1 et SRAS-S2.

**[0163]** L'ADN des clones Sa et Sb a été isolé puis l'insert correspondant a été séquencé à l'aide du Kit Big Dye, Applied Biosystem® et des amorces universelles M13 forward et M13 reverse, ainsi que des amorces : S/S/+21867, S/S/+22353, S/S/+22811, S/S/+23754, S/S/+24207, S/S/+24699, S/S/+24348, S/S/-24209, S/S/-23630, S/S/-23038, S/S/-22454, S/S/-21815, S/S/-24784, S/S/+21556, S/S/+23130 et S/S/+24465, en suivant les instructions du fabricant ; les séquences des fragments Sa et Sb ainsi obtenues correspondent aux séquences SEQ ID NO :5 et SEQ ID NO :6 dans la liste de séquences jointe en annexe.

**[0164]** Le plasmide, dénommé SARS-S1 a été déposé sous le n° I-3020, le 12 mai 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient un fragment 5' de la séquence du gène S de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, lequel fragment dénommé Sa correspondant aux nucléotides des positions 21406 à 23454 (SEQ ID NO :5), en référence à la séquence Genbank AY274119.3 Tor2.

**[0165]** Le plasmide, dénommé TOP10F'-SARS-S2 a été déposé sous le n° I-3019, le 12 mai 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient un fragment 3' de la séquence du gène S de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, lequel fragment dénommé Sb correspondant aux nucléotides des positions 23322 à 25348 (SEQ ID NO :6), en référence à la séquence Genbank n° d'accès AY274119.3.

b) clonage et séquençage de l'ADNc complet (clone SRAS-S de 4 kb)

**[0166]** L'ADNc S complet a été obtenu à partir des clones SARS-S1 et SARS-S2 précités, de la façon suivante :

- 1) une réaction d'amplification PCR a été réalisée sur un clone SARS-S2 en présence de l'amorce S/R4/-/25348-25329 précitée et de l'amorce S/S/+ /24696-24715: un amplicon de 633 pb a été obtenu,
  - 2) une autre réaction d'amplification PCR a été réalisée sur un autre clone SARS-S2, en présence des amorces S/F4/+ /23322-23341 précitée et S/S/- /24803-24784: un amplicon de 1481 pb a été obtenu,
- La réaction d'amplification a été réalisée dans les conditions telles que définies ci-dessus pour l'amplification des fragments Sa et Sb, à l'exception que 30 cycles d'amplifications comprenant une étape de dénaturation à 94° C pendant 20 sec et une étape d'élongation à 72° C pendant 2 min 30 sec ont été effectués.
- 3) les 2 amplicons (633 pb et 1481 pb) ont été purifiés dans les conditions telles que définies ci-dessus pour les fragments Sa et Sb.
  - 4) une autre réaction d'amplification PCR à l'aide des amorces S/F4/+ /23322-23341 et S/R4/-/25348-25329 précitées, a été réalisée sur les amplicons purifiés obtenus en 3). La réaction d'amplification a été réalisée dans les conditions telles que définies ci-dessus pour l'amplification des fragments Sa et Sb, à l'exception que 30 cycles d'amplifications ont été effectués.
- L'amplicon de 2026 pb ainsi obtenu a été purifié, cloné dans le vecteur PCR2.1-TOPO puis séquencé comme ci-dessus, à l'aide des amorces telles que définies ci-dessus pour les fragments Sa et Sb. Le clone ainsi obtenu a été dénommé clone 3'.
- 5) Le clone SARS-S1 précédemment obtenu et le clone 3' ont été digérés par *EcoR I*, les bandes d'environ 2kb ainsi obtenues ont été purifiées sur gel puis amplifiées par PCR avec les amorces S/F2/+ /21406-21426 et S/R4/-/25348-25329 précitées. La réaction d'amplification a été réalisée dans les conditions telles que définies ci-dessus pour l'amplification des fragments Sa et Sb, à l'exception que 30 cycles d'amplifications ont été effectués. L'amplicon d'environ 4 kb a été purifié et séquencé. Il a ensuite été cloné dans le vecteur PCR2.1-TOPO pour donner le plasmide, dénommé SARS-S, et l'insert contenu dans ce plasmide a été séquencé comme ci-dessus, à l'aide des amorces telles que définies ci-dessus pour les fragments Sa et Sb. Les séquences d'ADNc de l'insert et de l'amplicon codant pour la protéine S, correspondent respectivement aux séquences SEQ ID NO : 4 et SEQ ID NO : 2 dans la liste de séquences jointe en annexe, elles codent pour la protéine S (SEQ ID NO : 3).

**[0167]** La séquence de l'amplicon correspondant à l'ADNc codant pour la protéine S de la souche de SARS-CoV issue du prélèvement n°031589 présente les deux mutations suivantes par rapport aux séquences correspondantes de respectivement les isolats Tor2 et Urbani, les positions des mutations étant indiquées en référence à la séquence complète du génome de l'isolat Tor2 (Genbank AY274119.3) :

- g/t en position 23220 ; le codon alanine (gct) en position 577 de la séquence en acides aminés de la protéine S de Tor2 est remplacé par un codon sérine (tct),
- c/t en position 24872 : cette mutation ne modifie pas la séquence en acides aminés de la protéine S, et

**[0168]** Le plasmide, dénommé SARS-S, a été déposé sous le n° I-3059, le 20 juin 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient la séquence d'ADNc codant pour la protéine S de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, laquelle séquence correspondant aux nucléotides des positions 21406 à 25348 (SEQ ID NO :4), en référence à la séquence Genbank AY274119.3.

## 2.2) ADNc codant pour les protéines M et E

**[0169]** Les ARN issus du prélèvement 031589, extraits comme ci-dessus, ont été soumis à une transcription inverse, associée, lors de la même étape (kit *Titan One Step RT-PCR*®, Roche), à une réaction d'amplification par PCR, à l'aide des couples d'amorces :

- S/E/F1/+ /26051-26070 et S/E/RI/- /26455-26436 pour amplifier l'ORF-E, et
- S/M/F1/+ /26225-26244 et S/M/RI/- /27148-27129 pour amplifier l'ORF-M.

**[0170]** Un premier mélange réactionnel contenant : 8,6 µl d'H<sub>2</sub>O<sub>pp</sub>i, 1 µl de dNTP (5mM), 0,2 µl de chacune des amorces (50µM), 1,25 µl de DTT (100mM) et 0,25 µl de RNAsin (40UI/µl) a été combiné avec un deuxième mélange réactionnel contenant : 1 µl d'ARN, 7 µl d'H<sub>2</sub>O<sub>pp</sub>i, 5 µl de tampon de RT-PCR 5X et 0,5 µl de mélange d'enzyme et les mélanges combinés ont été incubés dans un thermocycleur dans les conditions suivantes : 30 min à 42°C, 10 min à

55°C, 2 min à 94°C suivi de 40 cycles comprenant une étape de dénaturation à 94°C pendant 10 sec, une étape d'hybridation à 55°C pendant 30 sec et une étape d'élongation à 68°C pendant 45 sec, avec 3 sec d'incrément par cycle et enfin une étape d'élongation terminale à 68°C pendant 7 min.

**[0171]** Les produits d'amplification ainsi obtenus (amplicons M et E) ont subi une deuxième amplification PCR (PCR nichée) en utilisant le kit Expand High-Fi®, Roche), à l'aide des couples d'amorces :

- S/E/F2/+26082-26101 et S/E/R2/-26413-26394 pour l'amplicon E, et
- S/M/F2/+26330-26350 et S/M/R2/-27098-27078 pour l'amplicon M.

**[0172]** Le mélange réactionnel contenant : 2 µl du produit de la première PCR, 39,25 µl d'H<sub>2</sub>O<sub>pp</sub>i, 5 µl de tampon 10X contenant du MgCl<sub>2</sub>, 2 µl de dNTP (5mM), 0,5µl de chacune des amorces (50 µM) et 0,75µl de mélange d'enzyme a été incubé dans un thermocycleur dans les conditions suivantes : une étape de dénaturation à 94°C pendant 2 min a été suivie de 30 cycles comprenant une étape de dénaturation à 94°C pendant 15 sec, une étape d'hybridation à 60°C pendant 30 sec et une étape d'élongation à 72°C pendant 45 sec, avec 3 sec d'incrément par cycle, et enfin une étape d'élongation terminale à 72°C pendant 7 min. Les produits d'amplification obtenus correspondant aux ADNc codant pour les protéines E et M ont été séquencés comme ci-dessus, à l'aide des amorces : S/E/F2/+26082 et S/E/R2/-26394, S/M/F2/+26330, S/M/R2/-27078 précitées et des amorces S/M/+26636-26655 et S/M/-26567-26548. Ils ont ensuite été clonés, comme ci-dessus, pour donner les plasmides dénommés SARS-E et SARS-M. L'ADN de ces clones a ensuite été isolé et séquencé à l'aide des amorces universelles M13 forward et M13 reverse ainsi que des amorces S/M/+26636 et S/M/-26548 précitées.

**[0173]** La séquence de l'amplicon représentant l'ADNc codant pour la protéine E (SEQ ID NO : 13) de la souche de SARS-CoV issue du prélèvement n°031589 ne comporte pas de différences par rapport aux séquences correspondantes des isolats AY274119.3-Tor2 et AY278741-Urbani. La séquence de la protéine E de la souche de SARS-CoV 031589 correspond à la séquence SEQ ID NO : 14 dans la liste de séquences jointe en annexe.

**[0174]** Le plasmide, dénommé SARS-E a été déposé sous le n° I-3046, le 28 mai 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient la séquence d'ADNc codant pour la protéine E de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle séquence correspondant aux nucléotides des positions 26082 à 26413 (SEQ ID NO : 15), en référence à la séquence Genbank n° d'accès AY274119.3.

**[0175]** La séquence de l'amplicon représentant l'ADNc codant pour la M (SEQ ID NO :16) de la souche de SARS-CoV issue du prélèvement n°031589 ne comporte pas de différences par rapport à la séquence correspondante de l'isolat AY274119.3-Tor2. En revanche, en position 26857, l'isolat AY278741-Urbani comporte un c et la séquence de la souche de SARS-CoV issue du prélèvement répertorié sous le n°031589 un t. Cette mutation aboutit à une modification de la séquence en acides aminés de la protéine correspondante: en position 154, une proline (AY278741-Urbani) est changée en sérine dans la souche de SARS-CoV issue du prélèvement répertorié sous le n°031589. La séquence de la protéine M de la souche de SARS-CoV issue du prélèvement répertorié sous le n°031589 correspond à la séquence SEQ ID NO :17 dans la liste de séquences jointe en annexe.

**[0176]** Le plasmide, dénommé SARS-M a été déposé sous le n° I-3047, le 28 mai 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient la séquence d'ADNc codant pour la protéine M de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus ; laquelle séquence correspondant aux nucléotides des positions 26330 à 27098 (SEQ ID NO : 18), en référence à la séquence Genbank n° d'accès AY274119.3.

### 2.3) ADNc correspondant aux ORF3, ORF4, ORF7 à ORF11

**[0177]** La même stratégie d'amplification, de clonage et de séquençage a été utilisée pour obtenir les fragments d'ADNc correspondant respectivement aux ORF suivantes: ORF 3, ORF4, ORF7, ORF8, ORF9, ORF10 et ORF11. Les couples d'amorces utilisées pour la première amplification sont :

- ORF3 et ORF4 : S/SE/F1/+25069-25088 et S/SE/R1/-26300-26281
- ORF7 à ORF11 : S/MN/F1/+26898-26917 et S/MN/R1/-28287-28266

**[0178]** Les couples d'amorces utilisées pour la deuxième amplification sont :

- ORF3 et ORF4 : S/SE/F2/+25110-25129 et S/SE/R2/-26244-26225
- ORF7 à ORF11 : S/MN/F2/+26977-26996 et S/MN/R2/-28218-28199

**[0179]** Les conditions de la première amplification (RT-PCR) sont les suivantes : 45 min à 42°C, 10 min à 55°C, 2

min à 94°C suivi de 40 cycles comprenant une étape de dénaturation à 94°C pendant 15 sec, une étape d'hybridation à 58°C pendant 30 sec et une étape d'élongation à 68°C pendant 1 min, avec 5 sec d'incrément par cycle et enfin une étape d'élongation terminale à 68°C pendant 7 min.

Les conditions de la PCR nichée sont les suivantes : une étape de dénaturation à 94°C pendant 2 min a été suivie de 40 cycles comprenant une étape de dénaturation à 94°C pendant 20 sec, une étape d'hybridation à 58°C pendant 30 sec et une étape d'élongation à 72°C pendant 50 sec, avec 4 sec d'incrément par cycle et enfin une étape d'élongation terminale à 72°C pendant 7 min.

**[0180]** Les produits d'amplification obtenus correspondant aux ADNc contenant respectivement les ORF3 et 4 et les ORF7 à 11 ont été séquencés à l'aide des amorces : S/SE+/25363, S/SE+/25835, S/SE-/25494, S/SE-/25875, S/MN+/27839, S/MN+/27409, S/MN-/27836 S/MN-/27799 et clonés comme ci-dessus pour les autres ORF, pour donner les plasmides dénommés SARS-SE et SARS-MN. L'ADN de ces clones a été isolé et séquencé à l'aide de ces mêmes amorces et des amorces universelles M13 sens et M13 anti-sens.

**[0181]** La séquence de l'amplicon représentant l'ADNc de la région contenant les ORF 3 et 4 (SEQ ID NO :7) de la souche de SARS-CoV issue du prélèvement n°031589 comporte une différence nucléotidique par rapport à la séquence correspondante de l'isolat AY274119-Tor2. Cette mutation en position 25298 aboutit à une modification de la séquence en acides aminés de la protéine correspondante (ORF 3): en position 11, une arginine (AY274119-Tor2) est changée en glycine dans la souche de SARS-CoV issue du prélèvement n°031589. En revanche, aucune mutation n'a été identifiée par rapport à la séquence correspondante de l'isolat AY278741-Urbani. Les séquences des ORF 3 et 4 la souche de SARS-CoV issue du prélèvement n°031589 correspondent respectivement aux séquences SEQ ID NO :10 et 12 dans la liste de séquences jointe en annexe.

**[0182]** Le plasmide, dénommé SARS-SE a été déposé sous le n° I-3126, le 13 novembre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient l'ADNc correspondant à la région située entre l'ORF-S et l'ORF-E et chevauchant l'ORF-E de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle région correspondant aux nucléotides des positions 25110 à 26244 (SEQ ID NO :8), en référence à la séquence Genbank n° d'accès AY274119.3,

**[0183]** La séquence de l'amplicon représentant l'ADNc correspondant à la région contenant les ORF7 à ORF11 (SEQ ID NO :19) de la souche de SARS-CoV issue du prélèvement n°031589 ne comporte pas de différences par rapport aux séquences correspondantes des isolats AY274119-Tor2 et AY278741-Urbani. Les séquences des ORF7 à 11 de la souche de SARS-CoV issue du prélèvement n°031589 correspondent respectivement aux séquences SEQ ID NO : 22, 24, 26, 28 et 30 dans la liste de séquences jointe en annexe.

**[0184]** Le plasmide dénommé SARS-MN a été déposé sous le n° I-3125, le 13 novembre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient la séquence d'ADNc correspondant à la région située entre l'ORF-M et l'ORF-N de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589 et prélevée à Hanoi, telle que définie ci-dessus, laquelle séquence correspondant aux nucléotides des positions 26977 à 28218 (SEQ ID NO :20 ), en référence à la séquence Genbank n° d'accès AY274119.3.

**[0185]** La séquence de l'amplicon représentant l'ADNc correspondant à la région contenant les ORF7 à ORF11 (SEQ ID NO :19) de la souche de SARS-CoV issue du prélèvement n°031589 ne comporte pas de différences par rapport aux séquences correspondantes des isolats AY274119-Tor2 et AY278741-Urbani. Les séquences des ORF7 à 11 de la souche de SARS-CoV issue du prélèvement n°031589 correspondent respectivement aux séquences SEQ ID NO : 22, 24, 26, 28 et 30 dans la liste de séquences jointe en annexe.

#### 2.4) ADNc codant pour la protéine N et incluant les ORF13 et ORF14

**[0186]** L'ADNc a été synthétisé et amplifié comme décrit ci-dessus pour les fragments Sa et Sb. De manière plus précise, le mélange réactionnel contenant : 5 µl d'ARN, 5 µl d'H<sub>2</sub>O ppi 4 µl de tampon de reverse transcriptase 5X, 2 µl de dNTP (5 mM), 2 µl d'oligo 20T (5 µM), 0,5 µl de RNasin (40 UI/ul) et 1, 5 µl de AMV-RT (10 UI/ul Promega) a été incubé dans un thermocycleur dans les conditions suivantes : 45 min à 42°C, 15 min à 55°C, 5 min à 95°C, puis il a été maintenu à +4°C.

**[0187]** Une première amplification PCR a été réalisée avec la paire d'amorces S/N/F3+/28023 et S/N/R3-/29480.

**[0188]** Le mélange réactionnel comme ci-dessus pour l'amplification des fragments S1 et S2 a été incubé dans un thermocycleur, dans les conditions suivantes : une étape initiale de dénaturation à 94°C pendant 2 min a été suivie de 40 cycles comprenant une étape de dénaturation à 94°C pendant 20 sec, une étape d'hybridation à 55°C pendant 30 sec puis une étape d'élongation à 72°C pendant 1 min 30 sec avec 10 sec d'élongation supplémentaire à chaque cycle, puis d'une étape finale d'élongation à 72°C pendant 5 min.

**[0189]** L'amplicon obtenu à la première amplification PCR a subi une seconde étape d'amplification PCR (PCR nichée) avec la paires d'amorce S/N/F4+/28054 et S/N/R4-/29430 dans des conditions identiques à celles de la première amplification.

**[0190]** Le produit d'amplification obtenu correspondant à l'ADNc codant pour la protéine N de la souche de SARS-



CoV issue du prélèvement n°031589a été séquencé à l'aide des amorces: S/N/F4/+28054, S/N/R4/-29430, S/N/+28468, S/N/+28918 et S/N/-28607 et cloné comme ci-dessus pour les autres ORF, pour donner le plasmide dénommé SARS-N. L'ADN de ces clones a été isolé et séquencé à l'aide des amorces universelles M13 sens et M13 anti-sens, ainsi que des amorces S/N/+28468, S/N/+28918 et S/N/-28607.

**[0191]** La séquence de l'amplicon représentant l'ADNc correspondant à l'ORF-N et incluant les ORF13 et ORF14 (SEQ ID NO :36) de la souche de SARS-CoV issue du prélèvement n°031589 ne comporte pas de différences par rapport aux séquences correspondantes des isolats AY274119.3-Tor2 et AY278741-Urbani. La séquence de la protéine N de la souche de SARS-CoV issue du prélèvement n°031589 correspond à la séquence SEQ ID NO : 37 dans la liste de séquences jointe en annexe.

**[0192]** Les séquences des ORF13 et 14 de la souche de SARS-CoV issue du prélèvement n°031589 correspondent respectivement aux séquences SEQ ID NO : 32 et 34 dans la liste de séquences jointe en annexe.

**[0193]** Le plasmide dénommé SARS-N a été déposé sous le n° I-3048, le 5 juin 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient l'ADNc codant pour la protéine N de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle séquence correspondant aux nucléotides des positions 28054 à 29430 (SEQ ID NO :38), en référence à la séquence Genbank n° d'accès AY274119.3.

## 2.5) extrémités 5' et 3' non-codantes

### a) extrémité 5' non-codante (5'NC)

#### a<sub>1</sub>) synthèse de l'ADNc

**[0194]** Les ARN issus du prélèvement 031589, extraits comme ci-dessus, ont été soumis à une transcription inverse dans les conditions suivantes :

**[0195]** L'ARN (15 µl) et l'amorce S/L/-443 (3 µl à la concentration de 5µm, ont été incubés 10 min à 75°C.

**[0196]** Ensuite, du Tampon de transcriptase inverse 5X (6 µl, INVITROGEN), des dNTP 10 mM (1 µl), du DTT 0,1M (3 µl) ont été ajoutés et le mélange a été incubé à 50°C pendant 3 min.

**[0197]** Enfin la transcriptase inverse (3 µl de Superscript®, INVITROGEN) a été ajoutée au mélange précédent qui a été incubé à 50°C pendant 1h30 puis à 90 °C pendant 2 min.

**[0198]** L'ADNc ainsi obtenu a été purifié à l'aide du kit *QIAquick PCR purification* (QIAGEN), selon les recommandations du fabricant.

#### b<sub>1</sub>) Réaction à la Terminal Transferase (TdT)

**[0199]** L'ADNc (10 µl) est incubé 2 min à 100°C, conservé dans la glace, puis sont ajoutés : H<sub>2</sub>O (2,5 µl), tampon TdT 5X (4 µl, AMERSHAM), dATP 5mM (2 µl) et TdT (1,5 µl, AMERSHAM). Le mélange ainsi obtenu est incubé 45 min à 37°C puis 2 min à 65°C.

**[0200]** Le produit obtenu est amplifié par une première réaction PCR à l'aide des amorces : S/L/-225-206 et ancre 14T : 5'-AGATGAATTCGGTACCTTTTTTTTTTTTTTTT-3' (SEQ ID NO :68). Les conditions de l'amplification sont les suivantes : une étape initiale de dénaturation à 94°C pendant 2 min est suivie de 10 cycles comprenant une étape de dénaturation à 94°C pendant 10 sec, une étape d'hybridation à 45°C pendant 30 sec puis une étape d'élongation à 72°C pendant 30 sec puis de 30 cycles comprenant une étape de dénaturation à 94°C pendant 10 sec, une étape d'hybridation à 50°C pendant 30 sec puis une étape d'élongation à 72°C pendant 30 sec, puis d'une étape finale d'élongation à 72°C pendant 5 min.

**[0201]** Le produit de la première amplification PCR a subi une seconde étape d'amplification à l'aide des amorces : S/L/-204-185 et ancre 14T précitée dans des conditions identiques à celles de la première amplification. L'amplicon ainsi obtenu a été purifié, séquencé à l'aide de l'amorce S/L/-182-163 puis il a été cloné comme ci-dessus pour les différentes ORF, pour donner le plasmide dénommé SARS-5'NC. L'ADN de ce clone a été isolé et séquencé à l'aide des amorces universelles M13 sens et M13 anti-sens et de l'amorce S/L/-182-163 précitée.

**[0202]** L'amplicon représentant l'ADNc correspondant à l'extrémité 5'NC de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589 correspond à la séquence SEQ ID NO : 72 dans la liste de séquences jointe en annexe ; cette séquence ne comporte pas de différences par rapport aux séquences correspondantes des isolats AY274119.3-Tor2 et AY278741-Urbani.

**[0203]** Le plasmide dénommé SARS-5'NC a été déposé sous le n° I- 3124, le 7 novembre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15 ; il contient l'ADNc correspondant à l'extrémité 5' non codante du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle séquence correspondant aux nucléotides des positions 1 à 204 (SEQ

ID NO :39 ), en référence à la séquence Genbank n° d'accès AY274119.3.

b) extrémité 3'non-codante (3'NC)

a<sub>1</sub>) synthèse de l'ADNc

**[0204]** Les ARN issus du prélèvement 031589, extraits comme ci-dessus, ont été soumis à une transcription inverse, selon le protocole suivant : le mélange réactionnel contenant : ARN (5 µl), H<sub>2</sub>O (5 µl), tampon de transcriptase inverse 5X (4 µl), dNTP 5 mM (2 µl), Oligo 20T 5µM (2 µl), RNasin 40 U/ µl (0,5 µl) et RT-AMV 10 U/ µl (1,5 µl, PROMEGA) a été incubé dans un thermocycleur, dans les conditions suivantes : 45 min à 42°C, 15 min à 55°C, 5 min à 95°C, puis il a été maintenu à +4°C.

**[0205]** L'ADNc obtenu a été amplifié par une première réaction PCR à l'aide des amorces S/N/+28468-28487 et ancre 14T précitée. Les conditions de l'amplification sont les suivantes : une étape initiale de dénaturation à 94°C pendant 2 min est suivie de 10 cycles comprenant une étape de dénaturation à 94°C pendant 20 sec, une étape d'hybridation à 45°C pendant 30 sec puis une étape d'élongation à 72°C pendant 50 sec puis de 30 cycles comprenant une étape de dénaturation à 94°C pendant 20 sec, une étape d'hybridation à 50°C pendant 30 sec puis une étape d'élongation à 72°C pendant 50 sec, puis d'une étape finale d'élongation à 72°C pendant 5 min.

**[0206]** Le produit de la première amplification PCR a subi une seconde étape d'amplification à l'aide des amorces S/N/+28933-28952 et ancre 14T précitée, dans des conditions identiques à celles de la première amplification. L'amplicon ainsi obtenu a été purifié, séquencé à l'aide de l'amorce S/N/+29257-29278 et cloné comme ci-dessus pour les différentes ORF, pour donner le plasmide dénommé SARS-3'NC. L'ADN de ce clone a été isolé et séquencé à l'aide des amorces universelles M13 sens et M13 anti-sens et de l'amorce S/N/+29257-29278 précitée.

**[0207]** L'amplicon représentant l'ADNc correspondant à l'extrémité 3'NC de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589 correspond à la séquence SEQ ID NO :73 dans la liste de séquences jointe en annexe ; cette séquence ne comporte pas de différences par rapport aux séquences correspondantes des isolats AY274119.3-Tor2 et AY278741-Urbani.

**[0208]** Le plasmide dénommé SARS-3'NC a été déposé sous le n° I-3123 le 7 novembre 2003, auprès de la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15. ; il contient la séquence d'ADNc correspondant à l'extrémité 3'non codante du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, telle que définie ci-dessus, laquelle séquence correspondant à celle située entre le nucléotide en position 28933 à 29727 (SEQ ID NO :40), en référence à la séquence Genbank n° d'accès AY274119.3, se termine par une série de nucléotides a.

## 2.6) ORF1a et ORF1b

**[0209]** L'amplification de la région 5' contenant les ORF1a et ORF1b du génome du SARS-CoV issu du prélèvement 031589 a été réalisée en pratiquant des réactions de RT-PCR suivies de PCR nichées selon les mêmes principes que ceux précédemment décrits pour les autres ORF. Les fragments amplifiés sont chevauchants sur plusieurs dizaines de bases, permettant ainsi la reconstruction informatique de la séquence complète de cette partie du génome. En moyenne, les fragments amplifiés sont de deux kilobases.

**[0210]** 14 fragments chevauchants dénommés L0 à L12 ont ainsi été amplifiés à l'aide des amorces suivantes :

**Tableau II: Amorces utilisées pour l'amplification de la région 5'(ORF1a et ORF1b)**

REGION AMPLIFIEE ET SEQUENCEE (ne tient pas compte des amorces)	Amorce sens RT- PCR	Amorce antisens RT-PCR	Amorce sens PCR nichée	Amorce antisens PCR nichée
<b>L0 50-480</b>	S/L0/F1/+30	S/L0/R1/-481		
<b>L1 231-2240</b>	S/L1/F1/+147	S/L1/R1/-2336	S/L1/F2/+211	S/L1/R2/-2241
<b>L2 2156-4167</b>	S/L2/F1/+2033	S/L2/R1/-4192	S/L2/F2/+2136	S/L2/R2/-4168
<b>L3 3913-5324</b>	S/L3bis/F1/+3850	S/L3bis/R1/-5365	S/L3bis/F2/+3892	S/L3bis/R2/-5325
<b>L4b 4952-6023</b>	S/L4b/F1/+4878	S/L4b/R1/-6061	S/L4b/F2/+4932	S/L4b/R2/-6024
<b>L4 5325-7318</b>	S/L4/F1/+5272	S/L4/R1/-7392	S/L4/F2/+5305	S/L4/R2/-7323

# EP 1 694 829 B1

(suite)

REGION AMPLIFIEE ET SEQUENCEE (ne tient pas compte des amorces)	Amorce sens RT- PCR	Amorce antisens RT-PCR	Amorce sens PCR nichée	Amorce antisens PCR nichée
L5 7296-9156	S/L5/F1/+7111	S/L5/R1/-9253	S/L5/F2/+7275	S/L5/R2/-9157
L6 9053-11066	S/L6/F1/+8975	S/L6/R1/-11151	S/L6/F2/+9032	S/L6/R2/-11067
L7 10928-12962	S/L7/F1/+10883	S/L7/R1/-13050	S/L7/F2/+10928	S/L7/R2/-12963
L8 12835-14834	S/L8/F1/+12690	S/L8/R1/-14857	S/L8/F2/+12815	S/L8/R2/-14835
L9 14765-16624	S/L9/F1/+14688	S/L9/R1/-16678	S/L9/F2/+14745	S/L9/R2/-16625
L10 16534-18570	S/L10/P1/+16451	S/L10/R1/-18594	S/L10/F2/+16514	S/L10/R2/-18571
L11 18521-20582	S/L11/F1/+18441	S/L11/R1/-20612	S/L11/F2/+18500	S/L11/R2/-20583
L12 20338-22205	S/L12/F1/+20279	S/L12/R1/-22229	S/L12/F2/+20319	S/L12/R2/-22206

**[0211]** Tous les fragments ont été amplifiés dans les conditions suivantes, excepté le fragment L0 qui a été amplifié comme décrit ci-dessus pour l'ORF-M :

- RT-PCR: 30 min à 42°C, 15 min à 55°C, 2 min à 94°C, puis l'ADNc obtenu est amplifié dans les conditions suivantes : 40 cycles comprenant : une étape de dénaturation à 94°C pendant 15 sec, une étape d'hybridation à 58°C pendant 30 sec puis une étape d'élongation à 68°C pendant 1 min 30 sec, avec 5 sec d'élongation supplémentaire à chaque cycle, puis une étape finale d'élongation à 68°C pendant 7 min.
- PCR nichée : une étape initiale de dénaturation à 94°C pendant 2 min est suivie de 3 cycles comprenant : une étape de dénaturation à 94°C pendant 15 sec, une étape d'hybridation à 60°C pendant 30 sec puis une étape d'élongation à 72°C pendant 1 min 30 sec, avec 5 sec d'élongation supplémentaire à chaque cycle, puis une étape finale d'élongation à 72°C pendant 7 min.

**[0212]** Les produits d'amplifications ont été séquencés à l'aide des amorces définies dans le Tableau III ci-après :

**Tableau III : Amorces utilisées pour le séquençage de la région 5' (ORF1a et ORF1b)**

Noms	Séquences (SEQ ID NO : 76 à 139)
S/L3/+4932	5'-CCACACACAGCTTGTGGATA-3'
S/L4/+6401	5'-CCGAAGTTGTAGGCAATGTC-3'
S/L4/+6964	5'-TTTGCTGCTCCTTCTTATTG-3'
S/L4/-6817	5'-CCGGCATCCAAACATAATTT-3'
S/L5/-7633	5'-TGGTCAGTAGGGTTGATTGG-3'
S/L5/-8127	5'-CATCCTTTGTGTCAACATCG-3'
S/L5/-8633	5'-GTCACGAGTGACACCATCCT-3'
S/L5/+7839	5'-ATGCGACGAGTCTGCTTCTA-3'
S/L5/+8785	5'-TTCATAGTGCTGGCTTACC-3'
S/L5/+8255	5'-ATCTTGGCGCATGTATTGAC-3'
S/L6/-9422	5'-TGCATTAGCAGCAACAACAT-3'
S/L6/-9966	5'-TCTGCAGAACAGCAGAAGTG-3'
S/L6/-10542	5'-CCTGTGCAGTTTGTCTGTCA-3'
S/L6/+10677	5'-CCTGTGGCAATGAAGTACA-3'
S/L6/+10106	5'-ATGTCATTTGCACAGCAGAA-3'
S/L6/+9571	5'-CTTCAATGGTTTGCCATGTT-3'
S/L7/-11271	5'-TGCGAGCTGTCATGAGAATA-3'
S/L7/-11801	5'-AACCGAGAGCAGTACCACAG-3'

# EP 1 694 829 B1

(suite)

Noms	Séquences (SEQ ID NO : 76 à 139)
S/L7/-12383	5'-TTTGGCTGCTGTAGTCAATG-3'
S/L7/+12640	5'-CTACGACAGATGTCCTGTGC-3'
S/L7/+12088	5'-GAGCAGGCTGTAGCTAATGG-3'
S/L7/+11551	5'-TTAGGCTATTGTTGCTGCTG-3'
S/L8/-13160	5'-CAGACAACATGAAGCACCAC-3'
S/L8/-13704	5'-CGCTGACGTGATATATGTGG-3'
S/L8/-14284	5'-TGCACAATGAAGGATACACC-3'
S/L8/+14453	5'-ACATAGCTCGCGTCTCAGTT-3'
S/L8/+13968	5'-GGCATTGTAGGCGTACTGAC-3'
S/L8/+13401	5'-GTTTGCAGGTGTAAGTGCAG-3'
S/L9/-15098	5'-TAGTGGCGGCTATTGACTTC-3'
S/L9/-15677	5'-CTAAACCTTGAGCCGCATAG-3'
S/L9/-16247	5'-CATGGTCATAGCAGCACTTG-3'
S/L9/+16323	5'-CCAGGTTGTGATGTCACTGAT-3'
S/L9/+15858	5'-CCTTACCCAGATCCATCAAG-3'
S/L9/+15288	5'-CGCAAACATAAACACTTGCTG-3'
S/L10/-16914	5'-AGTGTGGGTACAAGCCAGT-3'
S/L10/-17466	5'-GTTCCAAGGAACATGTCTGG-3'
S/L10/-18022	5'-AGGTGCCTGTGTAGGATGAA-3'
S/L10/+18245	5'-GGGCTGTCATGCAACTAGAG-3'
S/L10/+17663	5'-TCTTACACGCAATCCTGCTT-3'
S/L10/+17061	5'-TACCCATCTGCTCGCATAGT-3'
S/L11/-18877	5'-GCAAGCAGAATTAACCCTCA-3'
S/L11/-19396	5'-AGCACCACCTAAATTGCATC-3'
S/L11 /-20002	5'-TGGTCCCTTTGAAGGTGTTA-3'
S/L11 /+20245	5'-TCGAACACATCGTTTATGGA-3'
S/L11/+19611	5'-GAAGCACCTGTTTCCATCAT-3'
S/L11/+19021	5'-ACGATGCTCAGCCATGTAGT-3'
SARS/L1/F3/+800	5'-GAGGTGCAGTCACTCGCTAT-3'
SARS/L1/F4/+1391	5'-CAGAGATTGGACCTGAGCAT-3'
SARS/L1/F5/+1925	5'-CAGCAAACCACTCAATTCCT-3'
SARS/L1/R3/-1674	5'-AAATGATGGCAACCTCTTCA-3'
SARS/L1/R4/-1107	5'-CACGTGGTTGAATGACTTTG-3'
SARS/L1/R5/-520	5'-ATTTCTGCAACCAGCTCAAC-3'
SARS/L2/F3/+2664	5'-CGCATTGTCTCCTGGTTTAC-3'
SARS/L2/F4/+3232	5'-GAGATTGAGCCAGAACCAGA-3'
SARS/L2/F5/+3746	5'-ATGAGCAGGTTGTCATGGAT-3'
SARS/L2/R3/-3579	5'-CTGCCTTAAGAAGCTGGATG-3'
SARS/L2/R4/-2991	5'-TTTCTTCACCAGCATCATCA-3'
SARS/L2/R5/-2529	5'-CACCGTTCTTGAGAACAACC-3'
SARS/L3/F3/+4708	5'-TCTTTGGCTGGCTCTTACAG-3'
SARS/L3/F4/+5305	5'-GCTGGTGATGCTGCTAACTT-3'
SARS/L3/F5/+5822	5'-CCATCAAGCCTGTGTCGTAT-3'
SARS/L3/R3/-5610	5'-CAGGTGGTGCAGACATCATA-3'
SARS/L3/R4/-4988	5'-AACATCAGCACCATCCAAGT-3'
SARS/L3/R5/-4437	5'-ATCGGACACCATAGTCAACG-3'

**[0213]** Les séquences des fragments L0 à L12 de la souche de SARS-CoV issue du prélèvement répertorié sous le n° 031589, correspondent respectivement aux séquences SEQ ID NO :41 à SEQ ID NO :54 dans la liste de séquences jointe en annexe. Parmi ces séquences, seule celle correspondant aux fragments L5 comporte une différence nucléo-

tidique par rapport à la séquence correspondante de l'isolat AY278741-Urbani. Cette mutation t/c en position 7919 aboutit à une modification de la séquence en acides aminés de la protéine correspondante, codée par l'ORF la: en position 2552, une valine (codon gtt ; AY278741) est changée en alanine (codon gct) dans la souche de SARS-CoV 031589. En revanche, aucune mutation n'a été identifiée par rapport à la séquence correspondante de l'isolat AY274119.3-Tor2. Les autres fragments ne présentent pas de différences par rapport aux séquences correspondantes des isolats Tor2 et Urbani.

## **Exemple 2 : Production et purification de protéines N et S recombinantes de la souche de SARS-CoV issue du prélèvement répertorié sous le numéro 031589**

**[0214]** La protéine entière N et deux fragments polypeptidiques de la protéine S de la souche de SARS-CoV issue du prélèvement répertorié sous le numéro 031589 ont été produites chez *E. coli*, sous forme de protéines de fusion comprenant une étiquette polyhistidine N-ou C-terminale. Dans les deux polypeptides S, les séquences hydrophobes N et C-terminales de la protéine S (peptide signal : positions 1 à 13 et hélice transmembranaire : positions 1196 à 1218) ont été déléetées alors que l'hélice  $\beta$  (positions 565 à 687) et les deux motifs de type coiled-coils (positions 895 à 980 et 1155 à 1186) de la protéine S ont été préservés. Ces deux polypeptides sont constitués par : un fragment long ( $S_L$ ) correspondant aux positions 14 à 1193 de la séquence en acides aminés de la protéine S et un fragment court ( $S_C$ ) correspondant aux positions 475 à 1193 de la séquence en acides aminés de la protéine S. **1) Clonage des ADNc N,  $S_L$  et  $S_C$  dans les vecteurs d'expression pIVEX2.3 et**

### **pIVEX2.4**

**[0215]** Les ADNc correspondant à la protéine N et aux fragments  $S_L$  et  $S_C$  ont été amplifiés par PCR dans des conditions standard, à l'aide de l'ADN polymérase Platinum Pfx® (INVITROGEN). Les plasmides SRAS-N et SRAS-S ont été utilisés comme matrice et les oligonucléotides suivants comme amorces :

5'-CCCATATGCTCTGATAATGGACCCCAATCAAAC-3' (N sens, SEQ ID NO :55)  
 5'-CCCCCGGGTGCCTGAGTTGAATCAGCAGAAGC-3' (N antisens, SEQ ID NO :56)  
 5'-CCCATATGAGTGACCTTGACCGGTGCACCAC-3' ( $S_C$  sens, SEQ ID NO :57)  
 5'-CCCATATGAAACCTTGACCCACCTGCTC-3' ( $S_L$  sens, SEQ ID NO :58)  
 5'-CCCCCGGGTTTAATATATTGCTCATATTTTCCC-3' ( $S_C$  et  $S_L$  antisens, SEQ ID NO :59).

**[0216]** Les amorces sens introduisent un site *NdeI* (souligné) alors que les amorces antisens introduisent un site *XmaI* ou *SmaI* (souligné). Les 3 produits d'amplification ont été purifiés sur colonne (kit *QIAquick PCR Purification*, QIAGEN) et clonés dans un vecteur approprié. L'ADN plasmidique purifié des 3 constructions (kit *QIAfilter Midi Plasmid*, QIAGEN) a été vérifié par séquençage et digéré par les enzymes *NdeI* et *XmaI*. Les 3 fragments correspondants aux ADNc N,  $S_L$  et  $S_C$  ont été purifiés sur gel d'agarose puis insérés dans les plasmides pIVEX2.3MCS (étiquette polyhistidine C-terminale) et pIVEX2.4d (étiquette polyhistidine N-terminale) préalablement digérés par les mêmes enzymes. Après vérification des constructions, les 6 vecteurs d'expressions ainsi obtenus (pIV2.3N, pIV2.3 $S_C$ , pIV2.3 $S_L$ , pIV2.4N, pIV2.4 $S_C$  également dénommé pIV2.4 $S_L$ , pIV2.4 $S_L$ ) ont été ensuite utilisés, d'une part pour tester l'expression des protéines *in-vitro*, et d'autre part pour transformer la souche bactérienne BL21(DE3)pDIA17 (NOVAGEN). Ces constructions codent pour des protéines dont la masse moléculaire attendue est la suivante : pIV2.3N (47174 Da), pIV2.3 $S_C$  (82897 Da), pIV2.3 $S_L$  (132056 Da), pIV2.4N (48996 Da), pIV2.4 $S_L$  (81076 Da) et pIV2.4 $S_L$  (133877 Da). Des bactéries transformées par pIV2.3N ont été déposées à la CNCM le 23 octobre 2003, sous le numéro I-3117, et des bactéries transformées par pIV2.4 $S_L$  ont été déposées à la CNCM le 23 octobre 2003, sous le numéro I-3118.

## **2) Analyse de l'expression des protéines recombinantes *in-vitro* et *in vivo***

**[0217]** L'expression de protéines recombinantes à partir des 6 vecteurs recombinants a été testée, dans un premier temps, dans un système *in-vitro* (RTS100, Roche). Les protéines produites *in vitro*, après une incubation des vecteurs recombinants pIVEX, 4h à 30°C, dans le système RTS100, ont été analysées par western-blot à l'aide d'un anticorps anti-(his)<sub>6</sub> couplé à la peroxydase. Le résultat d'expression *in-vitro* (Figure 1) montre que seule la protéine N est exprimée en quantités importantes, cela quelle que soit la position, N- ou C-terminale, de l'étiquette polyhistidine. Dans une seconde étape, l'expression des protéines N et S a été testée *in-vivo* à 30°C dans du milieu LB, en présence ou en l'absence d'inducteur (IPTG 1mM). La protéine N est très bien produite dans ce système bactérien (Figure 2) et se retrouve principalement dans une fraction soluble après lyse des bactéries. En revanche, la version longue de S ( $S_L$ ) est très peu produite et complètement insoluble (Figure 3). La version courte ( $S_C$ ) présente également une très faible solubilité, mais un taux d'expression beaucoup plus élevé que celui de la version longue. Par ailleurs, la construction

S<sub>C</sub> fusionnée à une étiquette polyhistidine en position C-terminale présente une taille plus faible que celle attendue. Une expérience d'immunodétection avec un anticorps anti-polyhistidine a montré que cette construction était incomplète. En conclusion, les deux constructions, pIV2.3N et pIV2.4S<sub>I</sub>, exprimant respectivement la protéine N entière fusionnée à l'étiquette polyhistidine en C-terminal et la protéine S courte fusionnée à l'étiquette polyhistidine en N-terminal, ont été retenues pour produire les deux protéines en grande quantité afin de les purifier. Les plasmides pIV2.3N et pIV2.4S<sub>I</sub> ont été déposés respectivement sous le n° I-3117 et I-3118 auprès de la CNCM, 25 rue du Docteur Roux, 75724 PARIS 15, le 23 octobre 2003.

### 3) Analyse de l'activité antigénique des protéines recombinantes

[0218] L'activité antigénique des protéines N, S<sub>L</sub> et S<sub>C</sub> a été testée par western-blot, à l'aide de deux échantillons de sérum, provenant d'un même patient infecté par le SARS-CoV, prélevés 8 jours (M12) et 29 jours-(M13) après le début des symptômes du SRAS. Le protocole expérimental est comme décrit à l'exemple 3. Les résultats illustrés par la figure 4 montrent (i) la séroconversion du patient, et (ii) que la protéine N possède une plus forte réactivité antigénique que la protéine S courte.

### 4) Purification de la protéine N à partir de pIV2.3N

[0219] Plusieurs expériences de purification de la protéine N, produite à partir du vecteur pIV2.3N, ont été réalisées selon le protocole suivant. Les bactéries BL21(DE3)pDIA17, transformées par le vecteur d'expression pIV2.3N, ont été cultivées à 30°C dans 1 litre de milieu de culture contenant 0,1 mg/ml d'ampicilline, et induites par 1 mM IPTG quand la densité cellulaire, équivalente à A<sub>600</sub> = 0,8, est atteinte (environ 3 heures). Après 2 heures de culture en présence d'inducteur, les cellules ont été récupérées par centrifugation (10 min à 5000 rpm), remises en suspension dans le tampon de lyse (50 mM NaH<sub>2</sub>PO<sub>4</sub>, NaCl 0,3 M, 20 mM imidazole, pH 8 contenant le mélange d'inhibiteurs de protéases Complete®, Roche), et lysées par la presse de French (12000 psi). Après centrifugation du lysat bactérien (15 min à 12000 rpm), le surnageant (50 ml) a été déposé à un débit de 1 ml/min sur une colonne (15 ml) de chélation métallique (Ni-NTA superflow, Qiagen), équilibrée par le tampon de lyse. Après lavage de la colonne par 200 ml de tampon de lyse, la protéine N a été éluée par un gradient d'imidazole (20 → 250 mM) en 10 volumes de colonne. Les fractions contenant la protéine N ont été rassemblées et analysées par électrophorèse en gel de polyacrylamide en conditions dénaturantes puis coloration au bleu de Coomassie. Les résultats illustrés par la figure 5 montrent que le protocole employé permet de purifier la protéine N avec une homogénéité très satisfaisante (95%) et un rendement moyen de 15 mg de protéine par litre de culture.

### 5) Purification de la protéine S<sub>C</sub> à partir de pIV2.4S<sub>C</sub> (pIV2.4S<sub>I</sub>)

[0220] Le protocole suivi pour purifier la protéine S courte est très différent de celui décrit ci-dessus car la protéine est fortement agrégée dans le système bactérien (corps d'inclusion). Les bactéries BL21(DE3)pDIA17, transformées par le vecteur d'expression pIV2.4S<sub>I</sub> ont été cultivées à 30°C dans 1 litre de milieu de culture contenant 0,1 mg/ml d'ampicilline, et induites par 1 mM IPTG quand la densité cellulaire, équivalente à A<sub>600</sub> = 0,8, est atteinte (environ 3 heures). Après 2 heures de culture en présence d'inducteur, les cellules ont été récupérées par centrifugation (10 min à 5000 rpm), remises en suspension dans le tampon de lyse (0,1 M Tris-HCl, EDTA 1 mM, pH 7,5), et lysées par la presse de French (1200 psi). Après centrifugation du lysat bactérien (15 min à 12000 rpm), le culot a été remis en suspension dans 25 ml de tampon de lyse contenant 2% Triton X100 et 10 mM β-mercaptoéthanol, puis centrifugé pendant 20 min à 12000 rpm. Le culot a été remis en suspension dans un tampon Tris-HCl 10 mM contenant 7 M urée, et mis en agitation douce pendant 30 min à température ambiante. Ce dernier lavage des corps d'inclusion avec 7 M urée est nécessaire pour éliminer la plupart des protéines membranaires d'*E. coli* qui co-sédimentent avec la protéine S<sub>C</sub> agrégée. Après une dernière centrifugation pendant 20 min à 12000 rpm, le culot final est remis en suspension dans le tampon Tris-HCl 10 mM. L'analyse électrophorétique de cette préparation (Figure 6) montre que la protéine S courte peut être purifiée avec une homogénéité satisfaisante (environ 90%) à partir des corps d'inclusion (extrait insoluble).

### Exemple 3 : Immunodominance de la protéine N

[0221] La réactivité des anticorps présents dans le sérum des patients atteints de pneumopathie atypique causée par le coronavirus associé au SRAS (SARS-CoV), vis-à-vis des différentes protéines de ce virus, a été analysée par western-blot dans les conditions décrites ci-après.

## 1) Matériel

### a) lysats de cellules infectées par le SARS-CoV

**[0222]** Des cellules Vero E6 ( $2 \times 10^6$ ) ont été infectées par le SARS-CoV (isolat répertorié sous le numéro FFM/MA104) à une multiplicité d'infection (M.O.I.) de  $10^{-1}$  ou  $10^{-2}$  puis incubées dans du milieu DMEM contenant 2% de SVF, à 35°C dans une atmosphère contenant 5% de  $\text{CO}_2$ . 48 heures plus tard, le tapis cellulaire a été lavé avec du PBS puis lysé avec 500  $\mu\text{l}$  de tampon de dépôt préparé selon Laemmli et contenant du  $\beta$ -mercaptoéthanol. Les échantillons ont ensuite été bouillis 10 minutes puis soniqués 3 fois 20 secondes.

### b) anticorps

#### b<sub>1</sub>) sérum de patient atteint de pneumopathie atypique

**[0223]** Le sérum référencé au Centre National de Référence des virus influenzae (Région-Nord) sous le N° 20033168 est celui d'un patient français atteint d'une pneumopathie atypique causée par le SARS-CoV prélevé au jour 38 après le début des symptômes ; le diagnostic d'infection par le SARS-CoV a été réalisé par RT-PCR nichée et PCR quantitative.

#### b<sub>2</sub>) sérums polyclonaux de lapin monospécifiques dirigés contre la protéine N ou la protéine S

**[0224]** Les sérums sont ceux produits à partir des protéines recombinantes N et S<sub>C</sub> (exemple 2), selon le protocole d'immunisation décrit à l'exemple 4 ; il s'agit du sérum du lapin P13097 (sérum anti-N) et du sérum du lapin P11135 (sérum anti-S).

## 2) Méthode

**[0225]** 20  $\mu\text{l}$  de lysat de cellules infectées par le SARS-CoV à des M.O.I. de  $10^{-1}$  et  $10^{-2}$  et, à titre de contrôle, 20  $\mu\text{l}$  d'un lysat de cellules non infectées (mock) ont été séparés sur un gel SDS à 10% de polyacrylamide puis transférés sur une membrane de nitrocellulose. Après blocage dans une solution de PBS/lait 5%/Tween 0,1% et lavage en PBS/Tween 0,1%, cette membrane a été hybridée pendant une nuit à 4°C avec : (i) l'immun-sérum N° 20033168 dilué au 1/300, 1/1000 et 1/3000 dans le tampon PBS/BSA 1%/Tween 0,1%, (ii) le sérum du lapin P13097 (sérum anti-N) dilué au 1/50000 dans le même tampon et (iii) le sérum du lapin P11135 (sérum anti-S) dilué au 1/10000 dans le même tampon. Après lavage en PBS/Tween, une hybridation secondaire a été réalisée à l'aide, soit d'anticorps polyclonaux de mouton dirigés contre les chaînes lourdes et légères des immunoglobulines G humaines et couplés à la peroxidase (NA933V, Amersham), soit d'anticorps polyclonaux d'âne dirigés contre les chaînes lourdes et légères des immunoglobulines G de lapin et couplés à la peroxidase (NA934V, Amersham). Les anticorps fixés ont été révélés à l'aide du kit ECL+ (Amersham) et de films d'autoradiographie Hyperfilm MP (Amersham). Une échelle de masse moléculaire (kDa) est portée sur la figure.

## 3) Résultats

**[0226]** La figure 7 montre que trois polypeptides de masse moléculaire apparente 35, 55 et 200 kDa sont détectés spécifiquement dans les extraits de cellules infectées par le SARS-CoV.

**[0227]** Afin d'identifier ces polypeptides, deux autres immunoempreintes (figure 8) ont été réalisées sur les mêmes échantillons et dans les mêmes conditions avec des anticorps polyclonaux de lapins spécifique de la nucléoprotéine N (lapin P13097, figure 8A) et de la protéine de spicule S (lapin P11135, figure 8B) Cette expérience montre que le polypeptide de 200 kDa correspond à la glycoprotéine de spicule S du SARS-CoV, que le polypeptide de 55 kDa correspond à la nucléoprotéine N tandis que le polypeptide de 35 kDa représente vraisemblablement une forme tronquée ou dégradée de la N.

**[0228]** Les données présentées dans la figure 7 montrent donc que le sérum 20033168 réagit fortement avec la N et beaucoup plus faiblement avec la S du SARS-CoV, puisque les polypeptides de 35 et 55 kDa sont révélés sous la forme de bandes intenses pour des dilutions de 1/300, 1/1000 et 1/3000 de l'immunsérum alors que le polypeptide de 200 kDa n'est que faiblement révélé pour une dilution de 1/300. On peut noter également qu'aucun autre polypeptide du SARS-CoV n'est détecté pour des dilutions supérieures au 1/300 du sérum 20033168.

**[0229]** Cette expérience indique que la réponse en anticorps spécifique de la N du SARS-CoV domine les réponses en anticorps spécifiques des autres polypeptides du SARS-CoV et en particulier la réponse en anticorps dirigée contre la glycoprotéine S. Elle indique une immunodominance de la nucléoprotéine N lors des infections humaines par le SARS-CoV.

#### **Exemple 4 : Réparation d'anticorps polyclonaux monospécifiques dirigés contre les protéines N et S du coronavirus associé au SRAS (SARS-CoV)**

##### **1) Matériel et méthode**

**[0230]** Trois lapins (P13097, P13081, P13031) ont été immunisés avec le polypeptide recombinant purifié correspondant à l'intégralité de la nucléoprotéine (N), préparé selon le protocole décrit à l'exemple 2. Après une première injection de 0,35 mg par lapin de protéine émulsionnée en adjuvant complet de Freund (voie intradermique), les animaux ont reçus 3 injections de rappel à 3 puis 4 semaines d'intervalle, de 0,35 mg de protéine recombinante émulsionnée en adjuvant incomplet de Freund.

**[0231]** Trois lapins (P11135, P13042, P14001) ont été immunisés avec le polypeptide recombinant correspondant au fragment court de la protéine S ( $S_C$ ), produit comme décrit à l'exemple 2. Comme ce polypeptide est retrouvé principalement sous la forme de corps d'inclusion dans le cytoplasme bactérien, les animaux ont reçus 4 injections intra-dermiques à 3-4 semaines d'intervalle d'une préparation de corps d'inclusion correspondant à 0,5 mg de protéine recombinante émulsionnée en adjuvant incomplet de Freund. Les 3 premières injections ont été réalisées avec une préparation de corps d'inclusion préparés selon le protocole décrit à l'exemple 2, tandis que la quatrième injection a été réalisée avec une préparation de corps d'inclusion qui ont été préparés selon le protocole décrit à l'exemple 2 puis purifiés sur gradient de saccharose et lavés en 2 % Triton X100.

**[0232]** Pour chaque lapin, un sérum pré-immun (p.i.) a été préparé avant la première immunisation et un immun-sérum (I.S.) 5 semaines après la quatrième immunisation.

**[0233]** Dans un premier temps, la réactivité des sérums a été analysée par test ELISA vis à vis de préparations de protéines recombinantes semblables à celles utilisées pour les immunisations ; les tests ELISA ont été réalisés selon le protocole et avec les réactifs tels que décrits à l'exemple 6.

**[0234]** Dans un deuxième temps, la réactivité des sérums a été analysée en réalisant une immunoempreinte (western blot) d'un lysat de cellules infectées par le SARS-CoV, en suivant le protocole tel que décrit à l'exemple 3.

##### **2) Résultats**

**[0235]** Les tests ELISA (figure 9) démontrent que les préparations de protéine N recombinante et de corps d'inclusion du fragment court de la protéine S ( $S_C$ ) sont immunogènes chez l'animal et que le titre des sérums immuns est élevé (plus de 1/25000).

**[0236]** L'immunoempreinte (figure 8) montre que le sérum immun du lapin P13097 reconnaît deux polypeptides présents dans les lysats de cellules infectées par le SARS-CoV : un polypeptide dont la masse moléculaire apparente (50-55 kDa selon les expériences) est compatible avec celle de la nucléoprotéine N (422 résidus, masse moléculaire prédite de 46 kDa) et un polypeptide de 35 kDa, qui représente vraisemblablement une forme tronquée ou dégradée de la N.

**[0237]** Cette expérience montre également que le sérum du lapin P11135 reconnaît principalement un polypeptide dont la masse moléculaire apparente (180-220 kDa selon les expériences) est compatible avec une forme glycosylée de la S (1255 résidus, chaîne polypeptidique non glycosylée de 139 kDa), ainsi que des polypeptides plus légers, qui représentent vraisemblablement des formes tronquées et/ou non glycosylées de la S.

**[0238]** En conclusion, l'ensemble de ces expériences démontrent que des polypeptides recombinants exprimés chez *E. coli* et correspondant aux protéines N et S du SARS-CoV permettent d'induire chez l'animal des anticorps polyclonaux capables de reconnaître les formes natives de ces protéines.

#### **Exemple 5 : Préparation d'anticorps polyclonaux monospécifiques dirigés contre les protéines M et E du coronavirus associé au SRAS (SARS-CoV)**

##### **1) Analyse de la structure des protéines M et E**

###### **a) Protéine E**

**[0239]** La structure de la protéine E du SARS-CoV (76 acides aminés) a été analysée *in silico*, à l'aide de différents logiciels comme signalP v1.1, NetNGlyc 1.0, THMM 1.0 et 2.0 (Krogh et al., 2001, J. Mol. Biol., 305(3):567-580) ou encore TOPPRED (von Heijne, 1992, J. Mol. Biol. 225, 487-494). L'analyse montre que ce polypeptide non glycosylé est une protéine membranaire de type 1, contenant une seule hélice transmembranaire (aa 12-34 d'après THMM), et dont la plus grande partie du domaine hydrophile (42 résidus) est localisée à l'extrémité C-terminale et vraisemblablement à l'intérieur de la particule virale (endodomaine). On peut noter une inversion dans la topologie prédite par les versions 1.0 (N-ter est externe) et 2.0 (N-ter est interne) du logiciel THMM, mais que d'autres algorithmes, notamment TOPPRED



et THUMBUP (Zhou et Zhou, 2003, Protein Science 12 :1547-1555) confirment une localisation externe de l'extrémité N-terminale de E.

## b) Protéine M

**[0240]** Une analyse similaire réalisée sur la protéine M du SARS-CoV (221 acides aminés) montre que ce polypeptide ne possède pas de peptide signal (d'après le logiciel signalP v1.1) mais trois domaines transmembranaires (résidus 15-37, 50-72, 77-99 d'après THMM2.0) et un grand domaine hydrophile (aa 100-221) localisé à l'intérieur de la particule virale (endodomaine). Elle est vraisemblablement glycosylée sur l'asparagine en position 4 (d'après NetNGlyc 1.0).

**[0241]** Ainsi, en accord avec les données expérimentales connues pour les autres coronavirus, il est remarquable que les deux protéines M et E présentent des endodomains correspondant à la majeure partie des polypeptides et des ectodomains de très petite taille.

- l'ectodomaine de E correspond vraisemblablement aux résidus 1 à 11 ou 1 à 12 de la protéine : MYSFVSEETGT (L), SEQ ID NO : 70. En effet, la probabilité associée à la localisation transmembranaire du résidu 12 est intermédiaire (0,56 d'après THMM 2.0).
- l'ectodomaine de M correspond vraisemblablement aux résidus 2 à 14 de la protéine : ADNGTITVEELKQ, SEQ ID NO : 69. En effet, la méthionine N-terminale de M est très probablement clivée du polypeptide mature car le résidu en position 2 est une Alanine (Varshavsky, 1996, 93:12142-12149).

**[0242]** Par ailleurs, l'analyse de l'hydrophobicité (Kyte & Doolittle, Hopp & Woods) de la protéine E met en évidence que l'extrémité C-terminale de l'endodomaine de E est hydrophile et donc vraisemblablement exposée à la surface de ce domaine. Ainsi, un peptide synthétique correspondant à cette extrémité est un bon candidat immunogène pour induire chez l'animal des anticorps dirigés contre l'endodomaine de E. En conséquence, un peptide correspondant aux 24 résidus C-terminaux de E a été synthétisé.

## 2) Préparation d'anticorps dirigés contre l'ectodomaine des protéines M et E et l'endodomaine de la protéine E

**[0243]** Les peptides M2-14 (ADNGTITVEELKQ, SEQ ID NO : 69), E1-12 (MYSFVSEETGTL, SEQ ID NO: 70) et E53-76 (KPTVYVYSRV KNLNSSEGVP DLLV, SEQ ID NO : 71) ont été synthétisés par Neosystem. Ils ont été couplés à la KLH (*Keyhole Limpet Heinocyanin*) à l'aide du MBS (m-maleimidobenzoyl-N-hydroxysuccinimide ester) via une cystéine ajoutée au cours de la synthèse soit en N-terminal du peptide (cas de E53-76) soit en C-terminal (cas de M2-14 et E1-12).

**[0244]** Deux lapins ont été immunisés avec chacun des conjugués, en suivant le protocole d'immunisation suivant : après une première injection de 0,5 mg de peptide couplé à la KLH et émulsionné en adjuvant complet de Freund (voie intradermique), les animaux reçoivent 2 à 4 injections de rappel à 3 ou 4 semaines d'intervalle de 0,25 mg de peptide couplé à la KLH et émulsionné en adjuvant incomplet de Freund.

**[0245]** Pour chaque lapin, un sérum pré-immun (p.i.) a été préparé avant la première immunisation et un immun-sérum (I.S.) est préparé 3 à 5 semaines après les injections de rappel.

**[0246]** La réactivité des sérums a été analysée par western blot à l'aide d'extraits de cellules infectées par le SRAS-CoV (figure 43B) ou à l'aide d'extraits de cellules infectées par un virus recombinant de la vaccine exprimant la protéine E (VV-TG-E, figure 43A) ou M (VV-TN-M, figure 43C) de l'isolat 031589 du SRAS-CoV.

**[0247]** Les immunosérums des lapins 22234 et 22240, immunisés par le conjugué KLH-E53-76, reconnaissent un polypeptide d'environ 9 à 10kD, qui est présent dans les extraits de cellules infectées par le SRAS-CoV mais absent dans les extraits de cellules non infectées (figure 43B). La masse apparente de ce polypeptide est compatible avec la masse prédite de la protéine E, qui est de 8,4 kD. De façon similaire, l'immunosérum du lapin 20047, immunisé par le conjugué KLH-E1-12, reconnaît un polypeptide présent dans les extraits de cellules infectées par le virus VV-TG-E, dont la masse molaire apparente est compatible avec celle de la protéine E (figure 43A).

**[0248]** L'immunosérum des lapin 20013 et 20080, immunisés par le conjugué KLH-M2-14, reconnaît un polypeptide présent dans les extraits de cellules infectées par le virus VV-TN-M (figure 43C), dont la masse molaire apparente (18 kD environ) est compatible avec celle de la glycoprotéine M, qui est de 25,1 kD et présente un point isoélectrique élevé (9.1 pour le polypeptide nu).

**[0249]** Ces résultats démontrent que les peptides E1-12 et E53-76 d'une part et le peptide M2-14 d'autre part permettent d'induire chez l'animal des anticorps polyclonaux capables de reconnaître les formes natives des protéines E et M respectivement du SRAS-CoV.

**Exemple 6 : Analyse de la réactivité en ELISA de la protéine N recombinante, vis-à-vis de sérums de patients atteints de SRAS****1) Matériel**

[0250] L'antigène utilisé pour préparer les phases solides est la nucléoprotéine N recombinante purifiée préparée selon le protocole décrit à l'exemple 2.

[0251] Les sérums à tester (Tableau IV) ont été choisis sur la base des résultats d'analyse de leur réactivité par immunofluorescence (titre IF-SRAS), vis-à-vis de cellules infectées par le SARS-CoV.

**Tableau IV: Sérums testés en ELISA**

Référence	N° sérum	Type de sérum	Date du Sérum***	Titre IF-SRAS
3050	A	Témoin	na*	nt**
3048	B	Témoin	na	nt
033168	D	Patient 1- SRAS	27/04/03 (J38)	320
033397	E	Patient-1 SRAS	11/05/03 (J52)	320
032632	F	Patient-2 SRAS	21/03/03 (J17)	2500
032791	G	Patient-3 SRAS	04/04/03 (J3)	<40
033258	H	Patient-3 SRAS	28/04/03 (J27)	160
*na : non-applicable. ** nt : non-testé. *** les dates indiquées correspondent au nombre de jours après le début des symptômes de SRAS.				

**2) Méthode**

[0252] La protéine N (100 µl) diluée à différentes concentrations dans du tampon carbonate 0,1 M, pH 9,6 (1, 2 ou 4 µg/ml) est distribuée dans les puits de plaques ELISA, puis les plaques sont incubées une nuit à température du laboratoire. Les plaques sont lavées avec du tampon PBS-Tween, saturées avec du tampon PBS-lait écrémé-saccharose (5 %). Les sérums à tester (100 µl) préalablement dilués (1/50, 1/100, 1/200, 1/400, 1/800, 1/1600 et 1/3200) sont ajoutés, puis les plaques sont incubées 1 h à 37° C. Après 3 lavages, le conjugué anti-IgG humaines marqué à la peroxydase (référence 209-035-098, JACKSON) dilué au 1/18000 est ajouté puis les plaques sont incubées 1h à 37 °C. Après 4 lavages, le chromogène (TMB) et le substrat (H<sub>2</sub>O<sub>2</sub>) sont ajoutés et les plaques sont incubées 30min à température ambiante, à l'abri de la lumière. La réaction est ensuite arrêtée puis l'absorbance à 450 nm est mesurée à l'aide d'un lecteur automatique.

**3) Résultats**

[0253] Les tests ELISA (figure 10) démontrent que la préparation de protéine N recombinante est reconnue spécifiquement par les anticorps de sérums de patients atteints de SRAS prélevés en phase tardive de l'infection (≥ 17 jours après le début des symptômes) alors qu'elle n'est pas reconnue de façon significative par les anticorps d'un sérum de patient prélevé en phase précoce de l'infection (3 jours après le début des symptômes) ni par des sérums témoins de sujets non atteints de SRAS.

**Exemple 7 : Tests ELISA élaborés pour une détection très spécifique et sensible d'une infection par le coronavirus associé au SRAS, à partir de sérums de patients****1) Test ELISA IgG indirect****a Réactifs****Préparation des plaques**

[0254] Les plaques sont sensibilisées par une solution de protéine N à 2µg/mL dans un tampon PBS 10mM pH 7,2, rouge de phénol à 0,25mL /L. 100 µL de solution sont déposés dans les puits et laissés incubés à température ambiante

## EP 1 694 829 B1

pendant une nuit. La saturation se fait par un prélavage en tampon PBS 10mM / tween 0,1%, suivi d'un lavage avec une solution de saturation PBS, 25% lait /saccharose.

### *Diluant serums*

**[0255]** Tampon TRIS 0,48g/L, PBS 10mM, EDTA 3,7g/L, lait 15% v/v , pH 6,7

### *Diluant conjugué*

**[0256]** Tampon citrate (15g/L), tween 0,5% , sérum bovin 25%, NaCl 12%, lait écrémé 6% v/v PH 6,5

### *Conjugué*

**[0257]** Conjugué anti-IgG humaines 50X, commercialisés par Bio-Rad : kit Platelia H. pylori ref 72778

### *Autres solutions :*

**[0258]** Solution de lavage R2, Solutions de révélation au TMB R8 diluant, R9 chromogène, R10 solution d'arrêt : réactifs commercialisés commercialisées par Bio-Rad (ex : kit *Platelia pylori*, ref 72778)

### b) Mode opératoire

#### **[0259]**

Diluer les serums au **1/200** dans le diluant échantillons

Distribuer 100 $\mu$ L/puits

Incubation 1h à 37°C

3 lavages en solution de LAVAGE R2 10x préalablement diluée 10 fois en eau déminéralisée (*i.e.*, solution de lavage 1X)

Distribuer 100 $\mu$ L de conjugué (conjugué 50x à diluer extemporanément dans le diluant conjugué fourni)

Incubation 1h à 37°C

4 lavages en solution de lavage 1X

distribuer 200 $\mu$ L/puits de solution de révélation (à diluer extemporanément ex: mL de R9 dans 10mL de R8)

Incubation 30 min à température ambiante à l'obscurité

arrêter la réaction avec 100 $\mu$ L/puits de R10

LECTURE à 450/620nm

**[0260]** Les résultats peuvent être interprétés en prenant un sérum SEUIL donnant une réponse au delà de laquelle les serums testés seront considérés comme positifs. Ce sérum est choisi et dilué de façon à donner un signal significativement supérieur au bruit de fond.

## **2) Test ELISA DOUBLE EPITOPE**

### a Réactifs

#### *Préparation des plaques*

**[0261]** Les plaques sont sensibilisées par une solution de protéine N à 1 $\mu$ g/mL dans un tampon PBS 10mM pH 7,2, rouge de phénol à 0,25mL /L. 100  $\mu$ L de solution sont déposés dans les puits et laissés incubés à température ambiante pendant une nuit. La saturation se fait par un prélavage en tampon PBS 10mM / O,1% tween suivi d'un lavage avec une solution de saturation PBS 10mM, lait 25% (V/V)

#### *Diluant serums et conjugué*

**[0262]** Tampon TRIS saline 50mMpH8, lait 2%

## Conjugué

**[0263]** Il s'agit de la protéine N recombinante purifiée, couplée à la peroxidase selon le protocole de Nakane (Nakane P.K. and Kawaoi A; (1974) : Peroxydase-labeled antibody, a new method of conjugation. The Journal of Histochemistry and Cytochemistry Vol22, N)23, pp 1084-1091.), dans des ratios molaires respectifs 1/2. Ce conjugué ProtN POD est utilisé à une concentration de 2µg/mL dans du diluant serum/conjugué.

## Autres solutions :

**[0264]** Solution de lavage R2, Solutions de révélation au TMB R8, diluant, R9 chromogène, R10 solution d'arrêt : réactifs commercialisés par Bio-Rad (ex kit platelia pylori ref 72778).

## b) Mode opératoire

### 1ere étape en plaque de "prédilution"

#### [0265]

- Diluer chaque sérum au 1/5 dans la plaque de prédilution (48 µL de diluant +12 µL de sérum).
- Après avoir dilués l'ensemble des sérums, distribuer 60µL de conjugué
- Le cas échéant, le mélange sérum + conjugué est laissé à incuber.

### 2eme étape en plaque de "réaction"

#### [0266]

- Transférer 100µL de mélange/puits dans la plaque de réaction
- Incubation 1h 37°C
- 5 lavages en solution de LAVAGE R2 10x préalablement diluée 10 fois en eau déminéralisée (--> solution de lavage 1x)
- distribuer 200µL/puits de solution de révélation (à diluer extemporanément ex: 1 mL de R9 dans 10mL de R8)
- incubation 30min à température ambiante à l'abri de la lumière
- arrêter la réaction avec 100µL/puits de R10
- LECTURE à 450/620nm

**[0267]** De même que pour le test ELISA indirect, les résultats peuvent être interprétés en utilisant un serum "valeur seuil". Tout serum ayant une réponse supérieure au sérum valeur seuil sera considéré comme positif.

## 2) Résultats

**[0268]** Les sérums de patients classés comme cas probables de SRAS de l'hôpital français de Hanoi, Vietnam ou en relation avec l'hôpital français de Hanoi (JYK) ont été analysés en utilisant le test IgG-N indirect et le test N double épitope.

**[0269]** Les résultats du test IgG-N indirect (figures 14 et 15) et N double épitope (figures 16 et 17) montrent une excellente corrélation entre eux ainsi qu'avec un test ELISA indirect comparant la réactivité des sérums vis-à-vis d'un lysat de cellules VeroE6 non infectées ou infectées par le SRAS-CoV (ELISA-lysate SRAS-CoV ; voir le Tableau V ci-après). Tous les sérums prélevés 12 jours ou plus après le début des symptômes ont été trouvés positifs, y compris chez des patients pour lesquels l'infection par le virus du SRAS-CoV n'avait pas pu être documentée par analyse de prélèvements respiratoires par RT-PCR, vraisemblablement en raison d'un prélèvement trop tardif au cours de l'infection ( $\geq$  J12). Dans le cas du patient TTH pour lequel un prélèvement nasal réalisé à J7 a été trouvé négatif par RT-PCR, la qualité du prélèvement pourrait être en cause.

**[0270]** Certains sérums ont été trouvés négatifs alors que la présence de SRAS-CoV a été détectée par RT-PCR. Il s'agit dans tous les cas de sérums précoces prélevés moins de 10 jours après le début des symptômes (ex : sérum # 032637). Dans le cas d'un patient PTTH (sérum # 032673), seule une suspicion de SRAS était évoquée au moment où les prélèvements ont été réalisés.

**[0271]** En conclusion, les tests sérologiques IgG-N indirect et N-double épitope permettent de documenter l'infection par le SRAS-CoV chez tous les patients pour des sérums prélevés 12 jours ou plus après l'infection.

Tableau V : Résultats des tests ELISA

Num Pvt	Patient	Jour	PCR-SARS (1)	ELISA lysat SRAS-CoV (2)	IgG-N (2 <sup>ème</sup> série)	2Xepitope (2 <sup>ème</sup> série)
033168	JYK	38	POS	+++	>5000	NT
033597	JYK	74	POS	NT	≈ 5000	NT
032552	VTT	8	NEG-J3&J8&J12	NEG	<200	<5
032544	CTP	16	NEG J16&J20	++	>5000	>>20
032546	CJF	15	NEG J15&J19	++	>5000	>>20
032548	PTL	17	NEG J17&J21	++	>5000	>>20
032550	NTH	17	NEG-J17&J21	++	>5000	>>20
032553	VTT	8	NEG-J3&J8&J12	NEG	<200	<5
032554	NTBV	4	POS	NEG	<200	<5
032555	NTBV	4	POS	NEG	<200	
032564	NTP	15	POS	++	>5000	>>20
032629	NVH	4	POS	NEG	<200	<5
032631	BTTX	9	POS	NEG	<200	<5
032635	NHH	4	POS	NEG	<200	<5
032637	NHB	10	POS	NEG	<200	<5
032642	BTTX	9	POS	NEG	<200	<5
032643	LTDH	1	POS	NEG	<200	<5
032644	NTBV	4	POS	NEG	<200	<5
032646	TTH	12	NEG J7&J12&J16	++	>5000	>>20
032647	DTH	17	NEG J17&J21	++	>5000	>>20
032648	NNT	15	NEG J15&J19	++	>5000	>>20
032649	PTH	17	NEG J17&J21	++	>5000	>>20
032672	LVV	16	NEG J16&J20	+	>5000	>>20
032673	PTTH	NA	NEG	NEG	<200	<5
032674	PNB	17	NEG J17&J21	++	>5000	>>20
032682	VTH	12	NEG J12&J16	++	>5000	>>20
032683	DTV	17	NEG J17&J21	+	>1000	>>20
Notes: (1) : Les analyses par RT-PCR ont été réalisées par RT-PCR nichée BNI, LC Artus et LC-N sur des écouvillonnages nasaux ou pharyngés; POS signifie qu'au moins un prélèvement a été trouvé positif chez ce patient. (2) : La réactivité des sérums dans le test ELISA utilisant un lysat de cellules infectées par le SRAS-CoV a été classée en très fortement réactif (+++), fortement réactif (++), réactif (+) et négatif en fonction de la valeur DO obtenue aux dilutions testées						

**Exemple 8 : Détection du coronavirus associé au SRAS (SARS-CoV) par RT-PCR****1) Mise au point de conditions de RT-PCR en temps réel à l'aide d'amorces spécifiques du gène de la protéine de nucléocapside - test "Light Cycler N"**a) conception des amorces et des sondes

**[0272]** La conception des amorces et sondes a été réalisée à partir de la séquence du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le numéro 031589, à l'aide du programme "Light Cycler Probe Design (Roche)". Ainsi les deux séries d'amorces et de sondes suivantes ont été sélectionnées :

**- série 1 (SEQ ID NO : 60, 61, 64, 65):****[0273]**

- amorce sens : N/+28507 : 5'-GGC ATC GTA TGG GTT G-3' [28507-28522]
- amorce antisens : N/-28774 : 5'-CAG TTT CAC CAC CTC C-3' [28774-28759]
- sonde 1 : 5'-GGC ACC CGC AAT CCT AAT AAC AAT GC-fluorescéine 3' [28561-28586]
- sonde 2 : 5' Red705 -GCC ACC GTG CTA CAA CTT CCT-phosphate [28588-28608]

**- série 2 (SEQ ID NO : 62, 63, 66, 67)****[0274]**

- amorce sens : N/+28375 : 5'-GGC TAC TAC CGA AGA G-3' [28375-28390]
- amorce antisens : N/-28702 : 5'-AAT TAC CGC GAC TAC G-3' [28702-28687]
- sonde 1 : SRAS/N/FL : 5'-ATA CAC CCA AAG ACC ACA TTG GC - fluorescéine 3' [28541-28563]
- sonde 2 : SRAS/N/LC705 : 5' Red705 -CCC GCA ATC CTA ATA ACA ATG CTG C-phosphate 3' [28565-28589]

b) analyse de l'efficacité des deux couples amorces

**[0275]** Afin de tester l'efficacité respective des deux couples d'amorces, une amplification par RT-PCR a été réalisée sur un ARN synthétique correspondant aux nucléotides 28054-29430 du génome de la souche de SARS-CoV issue du prélèvement répertorié sous le numéro 031589 et contenant la séquence du gène N.

**[0276]** De manière plus précise :

Cet ARN synthétique a été préparé par transcription *in vitro* à l'aide de l'ARN polymérase du phage T7, d'une matrice d'ADN obtenu par linéarisation du plasmide SRAS-N avec l'enzyme *Bam* H1. Après élimination de la matrice d'ADN par digestion à l'aide de DNase 1, les ARN synthétiques sont purifiés par une extraction au phénol-chloroforme suivie de deux précipitations successives en acétate d'ammonium et isopropanol. Ils sont alors quantifiés par mesure de l'absorbance à 260 nm et leur qualité est contrôlée par le rapport des absorbances à 260 et 280 nm ainsi que par une électrophorèse en gel d'agarose. Ainsi, la concentration de la préparation d'ARN synthétique utilisée pour ces études est de 1,6 mg/ml, ce qui correspond à  $2,1 \cdot 10^{15}$  copies/ml d'ARN.

**[0277]** Des quantités décroissantes d'ARN synthétique ont été amplifiées par RT-PCR à l'aide du kit "Superscript™ One-Step RT-PCR with Platinum® Taq" et les couples d'amorces n° 1 (N/+28507, N/-28774) (figure 1A) et n° 2 (N/+28375, N/-28702) (figure 1B), en suivant les indications du fournisseur. Les conditions d'amplification utilisées sont les suivantes : l'ADNc a été synthétisé par incubation 30 min à 45 °C, 15 min à 55°C puis 2 min à 94 °C puis il a été amplifié par 5 cycles comprenant : une étape de dénaturation à 94°C pendant 15 sec, une étape d'hybridation à 45°C pendant 30 sec puis une étape d'élongation à 72°C pendant 30 sec, suivis de 35 cycles comprenant : une étape de dénaturation à 94°C pendant 15 sec, une étape d'hybridation à 55°C pendant 30 sec puis une étape d'élongation à 72°C pendant 30 sec, avec 2 sec d'élongation supplémentaire à chaque cycle, et d'une étape finale d'élongation à 72°C pendant 5 min. Les produits d'amplification obtenus ont ensuite été maintenus à 10°C.

**[0278]** Les résultats présentés à la figure 11 montrent que le couple d'amorces n° 2 (N/+28375, N/-28702) permet de détecter jusqu'à 10 copies d'ARN (bande de faible intensité) ou  $10^2$  copies (bande de bonne intensité) contre  $10^4$  copies pour le couple d'amorces n° 1 (N/+28507, N/-28774). Les amplicons sont respectivement de 268 pb (couple 1) et de 328 pb (couple 2).

c) mise au point de la RT-PCR en temps réel

**[0279]** Une RT-PCR en temps réel a été mise au point à l'aide du couple d'amorces n°2 et du couple de sonde constitué par SRAS/N/FL et SRAS/N/LC705 (figure 2).

**[0280]** L'amplification a été réalisée sur un LightCycler™ (Roche) à l'aide du kit "Light Cycler RNA Amplification Kit Hybridization Probes" (référence 2 015 145, Roche) dans les conditions optimisées suivantes. Un Mélange réactionnel contenant : H<sub>2</sub>O (6,8 µl), MgCl<sub>2</sub> 25 mM (0,8 µl, 4 µM final de Mg<sup>2+</sup>), mélange réactionnel 5X (4 µl), sonde SRAS/N/FL 3 µM (0,5 µl, 0,075 µM final), sonde SRAS/N/LC705 3 µM (0,5 µl, 0,075 µM final), amorce N/+28375 10 µM (1 µl, 0,5 µM final), amorce N/-28702 10 µM (1 µl, 0,5 µM final), mélange d'enzyme (0,4 µl) et échantillon (ARN viral, 5 µl) a été amplifié en suivant le programme suivant :

- Transcription inverse : **50°C** 10:00min analysis mode: none
- Dénaturation : 95°C 30sec x1 analysis mode: none
- Amplification : 95°C 2sec }
- 50°C** . 15sec analysis mode: quantification\* } x45
- 72°C 13sec rampe thermique 2,0°C/sec }
- refroidissement : 40°C 30sec x1 analysis mode: none

\*La mesure de fluorescence se fait à la fin de l'hybridation et à chaque cycle (en mode SINGLE).

**[0281]** Les résultats présentés à la figure 12 montrent que cette RT-PCR en temps réel est très sensible puisqu'elle permet de détecter 10<sup>2</sup> copies d'ARN synthétique dans 100% des 5 échantillons analysés (29/29 échantillons dans 8 expériences) et jusqu'à 10 copies d'ARN dans 100% des 5 échantillons analysés (40/45 échantillons dans 8 expériences). Elle montre également que cette RT-PCR permet de détecter la présence du génome du SARS-CoV dans un échantillon et de quantifier le nombre de génomes présents. A titre d'exemple, l'ARN viral d'un stock de SARS-CoV cultivé sur cellules Vero E6 a été extrait à l'aide du kit "Qiaamp viral RNA extraction" (Qiagen), dilué à 0,05.10<sup>-4</sup> et analysé par RT-PCR en temps réel selon le protocole décrit ci-dessus; l'analyse présentée à la figure 12 montre que ce stock de virus contient 6,5.10<sup>9</sup> génomes -équivalents/ml (geq/ml), ce qui est tout à fait similaire à la valeur de 1,0.10<sup>10</sup> geq/ml mesurée à l'aide du kit "RealArt™ HPA-Coronavirus LC RT PCR Reagents" commercialisé par Artus.

## 2) Mise au point de conditions de RT-PCR nichée ciblant le gène de l'ARN polymérase - test "RT-PCR nichée CDC (Centers for Disease Control and Prevention) /IP"

### a) Extraction de l'ARN viral

**[0282]** Echantillon clinique : QIAamp viral RNA Mini Kit (QIAGEN) selon les indications du fabricant, ou une technique équivalente. L'ARN est élué dans un volume de 60 µl.

### b) RT-PCR nichée "SNE/SAR"

*Première étape : RT-PCR couplée « SNE »*

**[0283]** Le kit Invitrogen "Superscript™ One-Step RT-PCR with Platinum® Taq" a été utilisé, mais le kit "Titan" de Roche Boehringer peut lui être substitué avec des résultats similaires.

Oligonucléotides :

### **[0284]**

- SNE-S1 5' GGT TGG GAT TAT CCA AAA TGT GA 3'
- SNE-AS1 5'GCA TCA TCA GAA AGA ATC ATC ATG 3'

-> Taille attendue : **440 pb**

1. Préparer un mix :

H2O	6,5 µl
Reaction mix 2X	12,5 µl
Oligo SNE-S1 50 µM	0,2 µl

## EP 1 694 829 B1

(suite)

Oligo SNE-AS1 50 µM	0,2 µl
RNAasin 40 U/µl	0,12 µl
RT/Platinum Taq mix	0,5 µl

2. A 20 µl du mix, ajouter 5 µl d'ARN et procéder à l'amplification sur thermocycleur (conditions ABI 9600) :

2.1	45°C	30 min.	
	55°C	15 min.	
	94°C	2 min.	
2.2.	94°C	15 sec.	}
	45°C	30 sec.	} x 5 cycles
	72°C	30 sec.	}
2.3.	94°C	15 sec.	}
	55°C	30 sec.	} x 35 cycles
	72°C	30 sec. + 2sec./cycle	}
2.4.	72°C	5 min.	
2.5	10°C	∞	

Conservation à +4°C.

**[0285]** La RNAasin (N2511/N2515) de Promega a été utilisée comme inhibiteurs de RNase.

**[0286]** Des ARN synthétiques ont servi de témoin positif. A titre de contrôle, 10<sup>3</sup>, 10<sup>2</sup> et 10 copies d'ARN synthétique R<sub>SNE</sub> ont été amplifiées dans chaque expérience.

*Seconde étape : PCR nichée "SAR"*

Oligonucléotides :

**[0287]**

- SAR1-S 5' CCT CTC TTG TTC TTG CTC GCA 3'
- SAR1-AS 5' TAT AGT GAG CCG CCA CAC ATG 3'

-> Taille attendue : **121 pb**

1. Préparer un mix :

H2O	35,8 µl
Tampon Taq 10X	5 µl
MgCl <sub>2</sub> 25 mM	4 µl
Mix dNTPs 5 mM	2 µl
Oligo SAR1-S 50 µM	0,5 µl
Oligo SAR1-AS 50 µM	0,5 µl
Taq ADN pol 5 U/µl	0,25 µl

L'*AmpliTaq DNA Pol* d'Applied Biosystems a été utilisée (tampon 10X sans MgCl<sub>2</sub>, réf 27216601).

2. A 48 µl du mix, ajouter 2 µl du produit de la première PCR et procéder à l'amplification (conditions ABI 9600) :

2.1.	94°C	2 min.	
2.2.	94°C	30 sec.	}
	45°C	45 sec.	} x 5 cycles
	72°C	30 sec.	}
2.3.	94°C	30 sec.	}
	55°C	30 sec.	} x 35 cycles



(suite)

	72°C	30 sec. + 1sec./cycle	}
2.4.	72°C	5 min.	
2.5	10°C	∞	

3. Analyser 10 µl du produit réactionnel sur gel "*low-melting*" (type Seakem GTG) à 3% d'agarose.

La sensibilité du test niché est en routine, dans les conditions décrites, de 10 copies d'ARN.

4. Les fragments peuvent ensuite être purifiés sur QIAquick PCR kilt (QIAGEN) et séquencés avec les oligos SAR1-S et SAR1-AS.

### 3) Détection de l'ARN du SARS-CoV par PCR à partir de prélèvements respiratoires

#### a) Première étude comparative

**[0288]** Une étude comparative a été réalisée sur une série de prélèvements respiratoires reçus par le Centre National de Référence du Virus Influenzae (région nord) et susceptibles de contenir du SARS-CoV. Pour ce faire, l'ARN a été extrait des prélèvements à l'aide du kit "Qiaamp viral RNA extraction" (Qiagen) et analysé par RT-PCR en temps réel, d'une part à l'aide des couples d'amorces et de sondes de la série n° 2 dans les conditions décrites ci-dessus d'une part, et d'autre part à l'aide du kit "LightCycler SARS-CoV quantification kit" commercialisé par Roche (référence 03 604 438). Les résultats sont résumés dans le Tableau VI ci-dessous. Ils montrent que 18 des 26 prélèvements sont négatifs et 5 des 26 prélèvements sont positifs pour les deux kits, tandis qu'un prélèvement est positif pour le seul kit Roche et deux pour les seuls réactifs N "série2". En outre, pour 3 prélèvements (20032701, 20032712, 20032714) les quantités d'ARN détectés sont nettement supérieures avec les réactifs (sondes et amorces) de la série n°2. Ces résultats indiquent que les amorces et sondes N "série2" sont plus sensibles pour la détection du génome du SARS-CoV dans des prélèvements biologiques que celles du kit actuellement disponible.

**Tableau VI: Analyse par RT-PCR en temps réel des ARN extraits d'une série de prélèvements de 5 patients à l'aide des couples d'amorces et de sondes de la série n° 2 (N "série 2") ou du kit "LightCycler SARS-CoV quantification kit" (Roche). Le type de prélèvement est indiqué ainsi que le nombre de copies de génome viral mesurées dans chacun des deux tests. NEG : RT-PCR négative.**

Prélèvements n°	Patient	Type de prélèvement	KIT ROCHE	N "série2"
20033082	K	nasal	NEG	NEG
20033083	K	pharyngé	NEG	NEG
20033086	K	nasal	NEG	NEG
20033087	K	pharyngé	NEG	NEG
20032802	M	nasal	NEG	NEG
20032803	M	expectoration	NEG	NEG
20032806	M	nasal ou pharyngé	NEG	NEG
20031746ARN2	C	pharyngé	NEG	NEG
20032711	C	nasal ou pharyngé	<b>39</b>	NEG
20032910	B	nasal	NEG	NEG
20032911	B	pharyngé	NEG	NEG
20033356	V	expectoration	NEG	NEG
20033357	V	expectoration	NEG	NEG
20031725	K	asp. endotrachéale	NEG	<b>150</b>
20032657	K	asp. endotrachéale	NEG	NEG
20032698	K	asp. endotrachéale	NEG	NEG
20032720	K	asp. endotrachéale	<b>3</b>	<b>5</b>
20033074	K	selles	<b>115</b>	<b>257</b>
20032701	M	pharyngé	<b>443</b>	<b>1676</b>
20032702	M	expectoration	NEG	<b>249</b>
20031747ARN2	C	pharyngé	NEG	NEG

(suite)

Prélèvements n°	Patient	Type de prélèvement	KIT ROCHE	N "série2"
20032712	C	inconnu	634	6914
20032714	C	pharyngé	17	223
20032800	B	nasal	NEG	NEG
20033353	V	nasal	NEG	NEG
20033384	V	nasal	NEG	NEG

## b) Deuxième étude comparative

**[0289]** Les performances de différentes méthodes de RT-PCR nichée et de RT-PCR en temps réel ont ensuite été comparées pour 121 prélèvements respiratoires de cas possibles de SRAS de l'hôpital français de Hanoi, Vietnam, réalisés entre le 4<sup>ème</sup> et le 17<sup>ème</sup> jour après le début des symptômes. Parmi ces prélèvements, 14 avaient été trouvés positifs lors d'un premier test utilisant la méthode de RT-PCR nichée ciblant l'ORF1b (codant pour la réplicase) telle que décrite initialement par le Bernhard Nocht Institute (RT-PCR nichée BNI). Des informations concernant ce test sont disponibles sur internet, à l'adresse [http://www15.bni-hamburg.de/bnibn12/neu2/getfile.acgi?area\\_engl=diagnostics&pid=4112](http://www15.bni-hamburg.de/bnibn12/neu2/getfile.acgi?area_engl=diagnostics&pid=4112).

**[0290]** Les différents tests comparés dans cette étude sont :

- la méthode de RT-PCR quantitative selon l'invention, avec les amorces et sonde N "série 2" décrite ci-dessus (colonne Light Cycler N),
- le test de RT-PCR nichée ciblant le gène de l'ARN polymérase décrit ci-dessus, développé par le CDC, le BNI et l'Institut Pasteur (RT-PCR nichée CDC/IP),
- le kit ARTUS de référence "HPA Corona LC RT-PCR Kit # 5601-02", qui est un test de RT-PCR en temps réel ciblant le gène ORF1b,
- le test de RT-PCR nichée du BNI, ciblant également le gène de l'ARN polymérase, mentionné ci-dessus.

**[0291]** Les inventeurs ont constaté :

- 1) une variabilité inter-test pour une même technique, liée à la dégradation de la préparation d'ARN lors de décongelations répétées, notamment pour les échantillons contenant les quantités d'ARN les plus faibles,
- 2) une sensibilité réduite de la RT-PCR nichée CDC/IP par rapport à la RT-PCR nichée BNI, et
- 3) une sensibilité comparable du test de RT-PCR quantitative selon l'invention (Light Cycler N) par rapport au test Light Cycler (LC) Artus.

**[0292]** Ces résultats, présentés dans le Tableau VII ci-dessous, montrent que le test par RT-PCR quantitative selon l'invention constitue un excellent complément - ou une alternative - aux tests actuellement disponibles. En effet, le coronavirus lié au SRAS est un virus émergent, susceptible d'évoluer rapidement. En particulier, le gène de la RNA polymérase du coronavirus lié au SRAS, qui est ciblé dans la plupart des tests actuellement disponibles, peut recombiner avec celui d'autres coronavirus non liés au SRAS. L'utilisation d'un test ciblant exclusivement ce gène pourrait alors conduire à l'obtention de faux négatifs.

**[0293]** Le test par RT-PCR quantitative selon l'invention ne cible pas la même région génomique que le kit ARTUS, puisqu'il cible le gène codant pour la protéine N. En réalisant un test de diagnostic ciblant deux gènes différents du coronavirus lié au SRAS, on peut donc espérer s'affranchir de résultats de type faux négatifs, qui pourraient être dus à l'évolution génétique du virus.

**[0294]** De plus, il apparaît particulièrement avantageux de cibler le gène de la protéine de nucléocapside, car il est très stable, du fait de la forte pression de sélection liée aux contraintes structurales élevées concernant cette protéine.

**Tableau VII: Comparaison de différentes méthodes d'analyse par amplification génique, à partir de 121 Prélèvements de cas probables de SRAS de l'hôpital français de Hanoi, Vietnam (épidémie 2003)**

N° CNR	Type prélèvement (1)	Jour prélèvement	Patient	RT-PCR nichée CDC/IP	RT-PCR nichée BNI	kit Light Cyclar Artus	Light Cyclar N (IP)
107 prélèvements	N et P			Négatif	Négatif	Négatif	Négatif
032529	P	10	NHB	Négatif	Positif	Négatif	Négatif
032530	N	10	NHB	Positif	Positif	3,10E+01	4,20E+01
032531	P	7	LP	Positif	Positif	7,70E+00	3,10E+00
032534	N	15	BND	Positif	Positif	1,60E+00	Négatif
032600	P	4	NHH	Négatif	Positif	Négatif	1,30E+02
032612	P	17	NTS	Négatif	Positif	Négatif	Négatif
032688	P	9	BTX	Positif	Positif	Négatif	Négatif
032689	N	4	NVH	Positif	Positif	1,20E+01	2,30E+02
032690	P	4	NVH	Négatif	Positif	1,60E+00	Négatif
032727	P	8	NVH	Positif	Positif	2,30E+02	4,00E+02
032728	N	8	NVH	Positif	Positif	1,10E+03	1,60E+04
032729	P	14	NHB	Positif	Positif	5,90E+00	3,40E+01
032730	N	14	NHB	Positif	Positif	1,30E+02	4,80E+02
032741	P	8	NHH	Positif	Positif	2,10E+02	1,30E+02
positifs				10	14	10	9
fraction détectée des 14 positifs				71,4%	100,0%	71,4%	64,3%
(1) P= écouvillonnage pharyngé N= écouvillonnage nasal							

#### Exemple 9 : Obtention et caractérisation d'anticorps monoclonaux dirigés contre la protéine N

**[0295]** Des souris Balb C ont été immunisées à l'aide de la protéine N recombinante purifiée et leurs cellules spléniques fusionnées avec un myélome murin approprié selon les techniques de Köhler et Milstein.

**[0296]** Dix-neuf hybridomes sécrétant d'anticorps anti N ont été présélectionnés et leurs immuno réactivités précisées. Ces anticorps reconnaissent bien la protéine N recombinante (en ELISA) avec des intensités variables, ainsi que la protéine virale, naturelle N en ELISA et/ou en Western Blot. Les figures 18 à 20 montrent les résultats de ces tests pour 15 de ces 19 anticorps monoclonaux.

**[0297]** Les clones 12, 17, 28, 57, 72, 76, 86, 87, 98, 103, 146, 156, 166, 170, 199, 212, 218, 219 et 222, fortement réactifs, ont été sous clonés. Les études de spécificité ont été poursuivies avec les outils appropriés afin de préciser les épitopes reconnus et vérifier l'absence de réactivité vis-à-vis des autres Coronavirus humains et de certains virus respiratoires.

**[0298]** Les études de cartographie (*mapping*) épitopique (réalisées sur membrane spot, à l'aide de peptides chevauchants de 15 aa) et les études supplémentaires réalisées sur la protéine N naturelle en Western Blot ont révélé l'existence de 4 groupes d'anticorps monoclonaux :

1° Anticorps monoclonaux spécifiques d'un épitope linéaire majeur en position N-ter (75-81, séquence : INTNSVP)  
Le représentant de ce groupe est l'anticorps 156. L'hybridome produisant cet anticorps a été déposé à la Collection Nationale de Cultures de Microorganismes (CNCM) de l'Institut Pasteur (Paris, France) le 1<sup>er</sup> décembre 2004, sous le numéro I-3331. Ce même épitope est également reconnu par un sérum de lapin (polyclonal anti N) obtenu par immunisation classique à l'aide de cette même protéine N.

2° Anticorps monoclonaux spécifiques d'un épitope linéaire majeur situé en position centrale (position 217-224, séquence : ETALALL); les représentants de ce groupe sont les anticorps monoclonaux 87 et 166. L'hybridome produisant l'anticorps 87 a été déposé à la CNCM le 1<sup>er</sup> décembre 2004, sous le numéro I-3328.

3° Anticorps monoclonaux spécifiques d'un épitope linéaire majeur situé en position C-terminale (position 403-408, séquence : DFFRQL), les représentants de ce groupe sont les anticorps 28, 57 et 143. L'hybridome produisant l'anticorps 57 a été déposé à la CNCM le 1<sup>er</sup> décembre 2004, sous le numéro I-3330.

4° Anticorps monoclonaux spécifiques d'un épitope discontinu, conformationnel. Ce groupe d'anticorps ne reconnaît aucun des peptides recouvrant la séquence de la protéine N, mais réagissent fortement sur la protéine naturelle non dénaturée. Le représentant de ce dernier groupe est l'anticorps 86. L'hybridome produisant cet anticorps a été déposé à la CNCM le 1<sup>er</sup> décembre 2004, sous le numéro I-3329.

**[0299]** Le Tableau VIII ci-après résume les résultats de cartographie épitopique obtenus :

**Tableau VIII : Mapping épitopique des anticorps monoclonaux**

Anticorps	Epitope	Position	Région
28	DFFRQL Q	403 ... 408	C-Ter.
143	DFFRQL Q		
76	DFFRQL Q		
57	DFFRQL Q		
	FFGMS RI	315 .... 319	
146	LPQRQ	383...387	
166	ETALALLL	217 ...224	centrale
87	ETALALL	217... 224	
156	INTNSGP	75... 81	N - Ter.
86	Conformationnel		
212	Conformationnel		
170	Conformationnel		

**[0300]** En outre, comme illustré notamment aux figures 18 et 19, ces anticorps ne présentent pas de réactivité en ELISA et/ou en WB, vis à vis de la protéine N du coronavirus humain 229 E.

**Exemple 10 : Combinaisons des anticorps monoclonaux pour le développement d'un test d'immunocapture sensible et spécifique de l'antigène viral N dans le sérum ou les fluides biologiques des patients contaminés par le virus SARS Co V**

**[0301]** Les anticorps listés ci-dessous ont été retenus en raison de leurs propriétés bien particulières pour une étude supplémentaire de capture et détection de la protéine virale N, dans le sérum des sujets ou patients.

**[0302]** Ces anticorps ont été produits en ascite sur souris, purifiés par chromatographie d'affinité et utilisés seuls ou en combinaison comme anticorps de capture, et comme anticorps signal.

**[0303]** Liste des anticorps sélectionnés pour le propos :

- Acn anti région C-ter (n° 28, 57, 143)
- Acn anti région centrale (n° 87, 166)
- Acn anti région N-ter (n° 156)
- Acn anti épitope discontinu, conformationnel (86)

#### 1) Préparation des réactifs :

##### a) Plaques ELISA immunocapture

**[0304]** Les plaques sont sensibilisées avec les solutions d'anticorps à 5 µg/ml en tampon carbonate 0,1 M, pH.9,6.

Les solutions (monovalentes ou plurivalentes) sont déposées sous volume de 100 µl dans les puits et incubées pendant une nuit à température ambiante. Ces plaques sont ensuite lavées en tampon PBS (10 mM pH 7,4 additionné de 0,1% de Tween 20) puis saturées avec une solution de PBS additionnée de 0,3% de BSA et de 5% de Saccharose). Les plaques sont ensuite séchées puis conditionnées dans un sachet en présence d'un dessiccant. Elles sont prêtes à l'emploi.

## b) Conjugués

**[0305]** Les anticorps purifiés ont été couplés à la peroxydase selon le protocole de Nakane (Nakane et al. - 1974, J. of Histo and cytochemistry, vol.22, p1084-1091) dans un rapport de une molécule d'IgG pour 3 molécules de peroxydase. Ces conjugués ont été purifiés par chromatographie d'exclusion et conservés concentrés (concentration comprise entre 1 à 2 mg/ml) en présence de 50% de glycerol et à -20°C. Ils sont dilués pour leur mise en oeuvre dans les essais à la concentration finale de 1 ou 2 µg/ml en tampon PBS (pH 7,4) additionné de 1% de BSA.

## c) Autres réactifs

### [0306]

- Sérums humains négatifs pour tous les marqueurs sériques des virus HIV, HBV, HCV et THLV
- Pool de sérums humains négatifs additionné de 0,5% de Triton X 100
- Ag viral inactivé : surnageant de culture virale inactivé par irradiation et inactivation vérifiée après mise en culture sur cellules sensibles - titre de la suspension avant inactivation environ  $10^7$  particules infectieuses par ml ou encore environ  $5 \times 10^9$  particules virales physiques par ml d'antigène
- Les échantillons d'Ag dilués en sérum humain négatif : ces échantillons ont été préparés par dilution au 1:100<sup>e</sup> puis par dilution en série de raison 5.
- Ces échantillons non infectieux miment des échantillons humains supposés contenir des concentrations faibles à très faibles de nucléoprotéine virale N. De tels échantillons ne sont pas accessibles pour les travaux de routine
- Solution de lavage R2, solution de révélation TMB R8, chromogène R9 et solution d'arrêt R10, sont les réactifs génériques commercialisés par Bio-Rad dans ses trousse ELISA (ex : trousse *Platelia Pylori* ref.72778).

## 2) Mode opératoire

**[0307]** Les échantillons de sérums humains surchargés en Ag viral inactivé sont distribués à raison de 100 µl par cupule, directement dans les plaques sensibilisées, prêtes à l'emploi puis incubés pendant 1 heure à 37°C (incubation Bio-Rad IPS).

**[0308]** Le matériel non retenu par la phase solide est éliminé par 3 lavages (lavage avec solution R2 diluée, laveur automatique LP 35).

**[0309]** Les conjugués appropriés, dilués à la concentration finale de 1 ou 2 µg/ml sont distribués à raison de 100 µl par cupule et les plaques incubées à nouveau pendant une heure à 37° C (incubation IPS).

**[0310]** L'excès de conjugué est éliminé par 4 lavages successifs (solution R2 diluée - laveur LP 35).

**[0311]** La présence de conjugué fixé sur les plaques est révélée après addition de 100 µl de solution de révélation préparée avant emploi (1 ml de R9 et 10 ml de R8) et après incubation pendant 30 minutes, à température ambiante et à l'abri de la lumière.

**[0312]** La réaction enzymatique est finalement bloquée par addition de 100 µl de réactif R10 (H2SO4 IN) dans toutes les cupules.

**[0313]** La lecture est effectuée à l'aide d'un lecteur de microplaque approprié à double longueur d'onde (450/620 nm).

**[0314]** Les résultats peuvent être interprétés en utilisant comme valeur seuil provisoire la moyenne d'au moins deux contrôles négatifs multiplié par un facteur 2 ou encore la moyenne de 100 sérums négatifs additionnée d'un incrément correspondant à 6 SD (*Standart Deviation* calculée sur les 100 mesures individuelles).

## 3) Résultats

**[0315]** Différentes combinaisons anticorps de capture et anticorps signal ont été testées en se fondant sur les propriétés des anticorps sélectionnés, et en évitant les combinaisons d'anticorps spécifiques des mêmes épitopes en phase solide et en conjugués.

**[0316]** Les meilleurs résultats ont été obtenus avec les 4 combinaisons listées ci-dessous. Ces résultats sont reproduits dans le tableau IX ci-après.

### 1. Combinaison F/28

## EP 1 694 829 B1

Phase solide (Acm 166 + 87 région centrale) : conjugué anticorps 28 (C-ter)

2. Combinaison G/28

Phase solide (Acm 86 - épitope conformationnel) : conjugué anticorps 28 (C-ter)

3. Combinaison H/28

Phase solide (Acms 86, 166 et 87 région centrale et épitope conformationnel) : conjugué anticorps 28 (C-ter)

4. Combinaison H/28 + 87

Phase solide (Acms 86, 166 et 87 région centrale et épitope conformationnel) : conjugué mixte anticorps 28 (C-ter) et 87 (central)

5. Combinaison G/87

Phase solide (Acm 86 - épitope conformationnel) : conjugué anticorps 87 (région centrale)

Les 4 premières combinaisons présentent des performances équivalentes et reproduites, supérieures aux autres combinaisons mises en oeuvre (comme par exemple la combinaison G/87). Bien entendu, dans ces combinaisons, un anticorps monoclonal peut être remplacé par un autre anticorps reconnaissant le même épitope. Ainsi, on peut citer les variantes suivantes :

6. Variante de la combinaison F/28

Phase solide (Acm 87 uniquement) : conjugué anticorps 57 (C-ter)

7. Variante de la combinaison G/28

Phase solide (Acm 86 - épitope conformationnel) : conjugué anticorps 57 (C-ter)

8. Variante de la combinaison H/28

Phase solide (Acms 86 et 87 région centrale et épitope conformationnel) : conjugué anticorps 57 (C-ter)

9. Variante de la combinaison H/28 + 87

Phase solide (Acms 86 et 87 région centrale et épitope conformationnel) : conjugué mixte anticorps 57 (C-ter) et 87 (central)

Tableau IX : Contrôle d'immunoréactivité des Acm anti nucléoprotéines SARS CoV : densités optiques mesurées avec chaque combinaison d'anticorps, en fonction des dilutions de l'antigène viral inactivé.

N°	Dilution	F/28	G/28	G/87	H/28	H/28+87
0	1/100	5	5	3,495	3,900	5
1	1/500	3,795	3,814	1,379	3,702	3,804
2	1/2 500	2,815	2,950	0,275	3,268	2,680
3	1/12500	0,987	1,038	0,135	1,374	0,865
4	1/62500	0,404	0,348	0,125	0,480	0,328
5	1/312500	0,285	0,211	0,123	0,240	0,215
6	Témoin	0,210	0,200	0,098	0,186	0,156
7	Témoin	0,269	0,153	0,104	0,193	0,202

[0317] La limite de détection de ces 4 essais expérimentaux correspond à la dilution d'antigène en sérum négatif 1: 62500. Une extrapolation rapide laisse supposer la détection de moins de  $10^3$  particules infectieuses par ml de sérums.

[0318] De cette étude, il ressort que les anticorps les mieux appropriés pour la capture de la nucléoprotéine virale native sont les anticorps spécifiques de la région centrale et/ou d'un épitope conformationnel, l'un et l'autre étant des anticorps sélectionnés également pour leur forte affinité pour l'antigène natif.

[0319] Ayant déterminé les meilleurs anticorps pour la composition de la phase solide, les anticorps à retenir en priorité pour la détection des antigènes fixés sur la phase solide, sont les anticorps complémentaires spécifiques d'un épitope dominant en région C-ter. L'emploi de tout autre anticorps complémentaire, mais spécifique des épitopes localisés en région N-ter de la protéine conduit à des résultats moyens ou médiocres.

### Exemple 11 : Systèmes d'expression eucaryotes de la protéine de spicule (S) du coronavirus associé au SRAS (SRAS-CoV)

#### 1) Optimisation des conditions d'expression de la S du SRAS-CoV en cellules de mammifères.

[0320] Les conditions d'expression transitoire de la protéine de spicule (S) du SRAS-CoV ont été optimisées en cellules de mammifères (293T, VeroE6).

[0321] Pour cela, un fragment d'ADN contenant le cDNA de la S du SRAS-CoV a été amplifié par PCR à l'aide des

oligonucléotides 5'-ATAGGATCCA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' et 5'-ATACTCGAGTT AT-GTGTAAATG TAATTTGACA CCCTTG-3' à partir du plasmide pSARS-S (C.N.C.M. n° I-3059) puis inséré entre les sites BamH1 et Xho1 du plasmide pTRIPΔU3-CMV contenant un vecteur lentiviral TRIP (Sirven, 2001, Mol. Ther., 3, 438-448) pour obtenir le plasmide pTRIP-S. Le fragment BamH1 et Xho1 contenant le cDNA de la S a ensuite été sous-cloné entre les BamH1 et Xho1 du plasmide d'expression eucaryote pcDNA3.1(+) (Clontech) pour obtenir le plasmide pcDNA-S. Le fragment Nhe1 et Xho1 contenant le cADN de la S a ensuite été sous-cloné entre les sites correspondants du plasmide d'expression pCI (Promega) pour obtenir le plasmide pCI-S. Les séquences WPRE du virus de l'hépatite de la marmotte ("*Woodchuck Hepatitis Virus posttranscriptional regulatory element*") et les séquences CTE ("*constitutive transport element*") du rétrovirus simien de Mason-Pfizer ont été insérées dans chacun des deux plasmides pcDNA-S et pCI-S entre les sites Xho1 et Xba1 pour obtenir respectivement les plasmides pcDNA-S-CTE, pcDNA-S-WPRE, pCI-S-CTE et pCI-S-WPRE (figure 21). Le plasmide pCI-S-WPRE a été déposé à la CNCM, le 22 novembre 2004, sous le numéro I-3323. Tous les inserts ont été séquencés à l'aide d'un kit BigDye Terminator v1.1 (Applied Biosystems) et d'un séquenceur automatique ABI377.

**[0322]** La capacité des plasmides construits à diriger l'expression de la S du SRAS-CoV dans des cellules de mammifères a été recherchée après transfection de cellules VeroE6 (figure 22). Dans cette expérience, des monocouches de  $5 \times 10^5$  cellules VeroE6 en boîtes de Pétri de 35mm ont été transfectées avec 2  $\mu$ g des plasmides pcDNA (à titre de contrôle), pcDNA-S, pCI et pCI-S et 6  $\mu$ l de réactif Fugene6 selon les indications du fabricant (Roche). Après 48 heures d'incubation à 37°C et sous 5% de CO<sub>2</sub>, des extraits cellulaires ont été préparé en tampon de dépôt selon Laemmli, séparés sur un gel SDS à 8% de polyacrylamide, puis transférés sur une membrane de PVDF (BioRad). La détection de cette immunoempreinte (« western blot ») a été réalisée à l'aide d'un sérum polyclonal de lapin anti-S (immun-sérum du lapin P11135 : cf exemple 4 ci-dessus) et d'anticorps polyclonaux d'âne dirigés contre les IgG de lapin et couplés à la peroxidase (NA934V, Amersham). Les anticorps fixés ont été révélés par luminescence à l'aide du kit ECL+ (Amersham) et de films d'autoradiographie Hyperfilm MP (Amersham).

**[0323]** Cette expérience (figure 22) montre que le plasmide pcDNA-S ne permet pas de diriger l'expression de la S du SRAS-CoV à des niveaux détectables tandis que le plasmide pCI-S permet une expression faible, proche de la limite de détection, qui peut être mise en évidence lorsque le film est surexposé. Des résultats similaires ont été obtenus lorsque l'expression de la S a été recherchée par immunofluorescence (données non montrées). Cette impossibilité de détecter une expression efficace de la S ne peut pas être imputée aux techniques de détection utilisées puisque la protéine S peut être mise en évidence à la taille attendue (180 kDa) dans un extrait de cellules infectées par le SRAS-CoV ou dans un extrait de cellules VeroE6 infectées par le virus recombinant de la vaccine VV-TF7.3 et transfectées par le plasmide pcDNA-S. Dans cette dernière expérience, le virus VV-TF7.3 exprime l'ARN polymérase du phage T7 et permet la transcription cytoplasmique d'un ARN non coiffé susceptible d'être traduit efficacement. Cette expérience suggère que les défauts d'expression décrits ci-dessus sont dus à une incapacité intrinsèque du cADN de la S à être exprimé efficacement lorsque l'étape de transcription en ARN messager est réalisée au niveau nucléaire.

**[0324]** Dans une seconde expérience, l'effet des signaux CTE et WPRE sur l'expression de la S a été recherchée après transfection de cellules VeroE6 (figure 23A) et 293T (figure 23B) et selon un protocole similaire à celui décrit ci-dessus. Alors que l'expression de la S ne peut pas être mise en évidence après transfection des plasmides pcDNA-S-CTE et pcDNA-S-WPRE dérivés de pcDNA-S, l'insertion des signaux WPRE et CTE améliore fortement l'expression de la S dans le contexte du plasmide d'expression pCI-S.

**[0325]** Pour préciser ce résultat, une deuxième série d'expériences a été réalisée, où l'immunoempreinte est révélée de façon quantitative par luminescence et acquisition sur un dispositif d'imagerie numérique (FluorS, BioRad). L'analyse des résultats obtenus avec le logiciel QuantityOne v4.2.3 (BioRad) montre que les séquences WPRE et CTE augmentent respectivement l'expression de la S d'un facteur 20 à 42 et 10 à 26 en cellules Vero E6 (Tableau X). En cellules 293T (Tableau X), l'effet de la séquence CTE est plus modéré (4 à 5 fois) tandis que celui de la séquence WPRE reste important (13 à 28 fois).

**Tableau X analyse quantitative de l'effet des signaux CTE et WPRE sur l'expression de la S du SRAS-CoV :**

Des extraits cellulaires ont été préparés 48 heures après transfection de cellules VeroE6 ou 293T par les plasmides pCI, pCI-S, pCI-S-CTE et pCI-S-WPRE et analysés par western blot comme décrit dans la légende de la figure 22. Le western blot est révélé par luminescence (ECL+, Amersham) et acquisition sur un dispositif d'imagerie numérique (FluorS, BioRad). Les niveaux d'expression sont indiqués en fonction d'une échelle arbitraire où la valeur de 1 représente le niveau mesuré après transfection du plasmide pCI-S. Deux expériences indépendantes ont été réalisées pour chacun des deux types cellulaires. Dans l'expérience 1 sur cellules VeroE6, les transfections ont été réalisées en dupliqués et les résultats sont indiqués sous la forme de moyennes et écart-types des niveaux d'expression mesurés.

Plasmide	Cellule	exp. 1	exp. 2
PCI	VeroE6	0,0	0,0
pCI-S	VeroE6	1,0 ± 0,1	1,0
pCI-S-CTE	VeroE6	9,8 ± 0,9	26,4
pCI-S-WPRE	VeroE6	20,1 ± 2,0	42,3
PCI	293T	0,0	0,0
PCI-S	293T	1,0	1,0
PCI-S-CTE	293T	4,6	4,0
PCI-S-WPRE	293T	27,6	12,8

**[0326]** En résumé, l'ensemble de ces résultats montre que l'expression dans des cellules de mammifères du cADN de la S du SRAS-CoV sous la dépendance de séquences promotrices de l'ARN polymérase II, requiert, pour être efficace, la présence d'un signal d'épissage ainsi que de l'une ou l'autre des séquences WPRE et CTE.

**2) Obtention de lignées stables permettant l'expression de la S du SRAS-CoV**

**[0327]** Le cADN de la protéine S du SRAS-CoV a été cloné sous la forme d'un fragment BamH1-Xho1 dans le plasmide pTRIPΔU3-CMV contenant un vecteur lentiviral défectif TRIP à DNA flap central (Sirven et al, 2001, Mol. Ther., 3 : 438-448) pour obtenir le plasmide pTRIP-S (figure 24). La co-transfection transitoire selon Zennou et al (2000, Cell, 101 : 173-185) de ce plasmide, d'un plasmide d'encapsulation (p8.2) et d'un plasmide d'expression de la glycoprotéine d'enveloppe G du VSV (pHCMV-G) dans des cellules 293T a permis la préparation de pseudoparticules rétrovirales contenant le vecteur TRIP -S et pseudotypées par la protéine d'enveloppe G. Ces vecteurs TRIP-S pseudotypés ont été utilisés pour transduire des cellules 293T et FRhK-4 : aucune expression de la protéine S n'a pu être mise en évidence par western blot et immunofluorescence dans les cellules transduites (données non présentées).

**[0328]** Les cassettes d'expression optimales constituées du promoteur immédiat/précoce du virus CMV, d'un signal d'épissage, du cDNA de la S et de l'un ou l'autre des signaux post-transcriptionnels WPRE ou CTE décrites ci-dessus ont alors été substituées à la cassette EF1α-EGFP du vecteur d'expression lentiviral défectif à DNA FLAP central TRIPΔU3-EF1α (Sirven et al, 2001, Mol. Ther., 3 : 438-448) (figure 25). Ces substitutions ont été réalisées par une série de sous-clonages successifs des cassettes d'expression de S qui ont été excisées des plasmides pCT-S-CTE (BgIII-Apa1) ou respectivement pCI-S-WPRE (BgIII-Sall) puis insérées entre les sites Mlu1 et Kpn1 ou respectivement Mlu1 et Xho1 du plasmide TRIPΔU3-EF1α pour obtenir les plasmides pTRIP-SD/SA-S-CTE et pTRIP-SD/SA-S-WPRE, déposés à la CNCM, le 1<sup>er</sup> décembre 2004, sous les numéros I-3336 et I-3334, respectivement. Des vecteurs pseudotypés ont été produits selon Zennou et al (2000, Cell, 101 : 173-185) et utilisés pour transduire des cellules 293T (10000 cellules) et FRhK-4 (15000 cellules) selon une série de 5 cycles successifs de transduction avec une quantité de vecteur correspondant à 25 ng (TRIP-SD/SA-S-CTE) ou 22 ng TRIP-SD/SA-S-WPRE) de p24 par cycle.

**[0329]** Les cellules transduites ont été clonées par dilution limite et une série de clones ont été analysés pour l'expression de la S du SRAS-CoV qualitativement par immunofluorescence (données non montrées), puis quantitativement par western blot (figure 25) à l'aide d'un sérum polyclonal de lapin anti-S. Les résultats présentés dans la figure 25 montrent que les clones 2 et 15 de cellules FRhK4-s-CTE transduites par TRIP-SD/SA-S-CTE et les clones 4, 9 et 12 de cellules FRhK4-S-WPRE transduites par TRIP-SD/SA-S-WPRE permettent l'expression de la S du SRAS-CoV à des niveaux respectivement faibles et modérés si on les compare à ceux que l'on peut observer au cours de l'infection par le SRAS-CoV.

**[0330]** En résumé, les vecteurs TRIP-SD/SA-S-CTE et TRIP-SD/SA-S-WPRE permettent l'obtention de clones stables de cellules FRhK-4 et de façon similaire 293T exprimant la S du SRAS-CoV, alors que les essais réalisés avec le vecteur



"de base" TRIP-S sont restés infructueux, ce qui démontre la nécessité d'un signal d'épissage ainsi que de l'une, ou l'autre, des séquences CTE et WPRE pour l'obtention de clones cellulaires stables exprimant la protéine S.

**[0331]** En outre, ces modifications du vecteur TRIP (insertion d'un signal d'épissage et d'un signal post-transcriptionnel comme CTE et WPRE) pourraient s'avérer intéressantes pour améliorer l'expression d'autres cDNA que celui de la S.

3) Obtention de lignées stables permettant l'expression d'une forme soluble de la S du SRAS-CoV. Purification de cet antigène recombinant

**[0332]** Un cDNA codant pour une forme soluble de la protéine S (Ssol) a été obtenu en fusionnant les séquences codant pour l'ectodomaine de la protéine (acides aminés 1 à 1193) à celles d'une étiquette (FLAG : DYKDDDDK) via un linker BspEI codant pour le dipeptide SG. Pratiquement, pour obtenir le plasmide pcDNA-Ssol, un fragment d'ADN codant pour l'ectodomaine de la S du SRAS-CoV a été amplifié par PCR à l'aide des oligonucléotides 5'-ATAGGATCCA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' et 5'-ACCTCCGGAT TTAATATATT GTCATATTT TCCCAA-3' à partir du plasmide pcDNA-S, puis inséré entre les sites uniques BamH1 et BspEI uniques d'un plasmide d'expression eucaryote pcDNA3.1(+) (Clontech) modifié contenant entre ses sites BamH1 et Xho1 la séquence de l'étiquette FLAG :

```
// GGATCC ...nnn... TCC GGA GAT TAT AAA GAT GAC GAC GAT AAA TAA
      BamH1           S  G  D  Y  K  D  D  D  D  K  ter
CTCGAG //
Xho1
```

**[0333]** Les fragments Nhe1-Xho1 et BamH1-Xho1, contenant le cADN de la S, ont ensuite été excisés du plasmide pcDNA-Ssol1, et sous-clonés entre les sites correspondants du plasmide pTRIP-SD/SA-S-CTE et du plasmide pTRIP-SD/SA-S-WPRE, respectivement, pour obtenir les plasmides pTRIP-SD/SA-Ssol-CTE et pTRIP-SD/SA-Ssol-WPRE, déposés à la CNCM, le 1<sup>er</sup> décembre 2004, sous les numéros I-3337 et I-3335, respectivement.

**[0334]** Des vecteurs pseudotypés ont été produits selon Zennou et coll. (2000, Cell, 101 : 173-185) et utilisés pour transduire des cellules FRhK-4 (15000 cellules) selon une série de 5 cycles successifs de transduction (15000 cellules) avec une quantité de vecteur correspondant à 24 ng (TRIP-SD/SA-Ssol-CTE) ou 40 ng TRIP-SD/SA-Ssol-WPRE) de p24 par cycle. Les cellules transduites ont été clonées par dilution limite et une série de 16 clones transduits par TRIP-SD/SA-Ssol-CTE et de 15 clones par TRIP-SD/SA-Ssol-WPRE ont été analysés pour l'expression du polypeptide Ssol par western blot révélé par un anticorps monoclonal anti-FLAG (figure 26 et données non montrées), ainsi que par un ELISA-capture spécifique du polypeptide Ssol qui a été mis au point dans ce but (Tableau XI et données non montrées). Une partie du processus de sélection des meilleurs clones sécréteurs est montré dans la figure 26. L'ELISA-capture repose sur l'utilisation de phases solides recouvertes d'anticorps polyclonaux de lapins immunisés par du SRAS-CoV purifié et inactivé. Ces phases solides permettent la capture du polypeptide Ssol sécrété dans les surnageants cellulaires, dont la présence est ensuite révélée par une série d'étapes impliquant successivement la fixation d'un anticorps monoclonal anti-FLAG (M2, SIGMA), d'anticorps biotinylés polyclonaux de lapin anti-IgG(H+L) de souris (Jackson) et d'un conjugué streptavidine-peroxidase (Amersham) puis l'addition de chromogène et de substrat (TMB + H<sub>2</sub>O<sub>2</sub>, KPL).

**Tableau XI : analyse de l'expression de polypeptide Ssol par des lignées cellulaires transduites par les vecteurs lentiviraux TRIP-SD/SA-Ssol-WPRE et TRIP-SD/SA-Ssol-CTE.** La sécrétion du polypeptide Ssol a été recherchée dans le surnageant d'une série de clones cellulaires isolés après transduction de cellules FRhK-4 par les vecteurs lentiviraux TRIP-SD/SA-Ssol-WPRE et TRIP-SD/SA-Ssol-CTE. Les surnageants dilués au 1/50 ont été analysés par un test ELISA-capture spécifique de la S du SRAS-CoV.

Vecteur	Clone	DO (450 nm)
Témoin	-	0,031
TRIP-SD/SA-Ssol-CTE	CTE2	0,547
	CTE3	0,668
	CTE9	0,171
	CTE12	0,208
	CTE13	0,133

(suite)

Vecteur	Clone	DO (450 nm)
TRIP-SD/SA-Ssol-WPRE	WPRE1	0,061
	WPRE10	0,134

**[0335]** La lignée cellulaire sécrétant les quantités les plus élevées de polypeptide Ssol dans le surnageant de culture est la lignée FRhK4-Ssol-CTE3. Elle a été soumise à une deuxième série de 5 cycles de transduction par le vecteur TRIP-SD/SA-Ssol-CTE dans des conditions similaires à celles décrites ci-dessus puis clonée. Le sous-clone sécrétant les quantités les plus élevées de Ssol a été sélectionné par une combinaison d'analyse par western blot et ELISA-capture : il s'agit du sous-clone FRhK4-Ssol-30, qui a été déposé à la CNM, le 22 novembre 2004, sous le numéro I-3325.

**[0336]** La lignée FRhK4-Ssol-30 permet la production et la purification en quantité du polypeptide recombinant Ssol. Dans une expérience type où les conditions expérimentales de croissance, de production et de purification ont été optimisées, les cellules de la lignée FRhK4-Ssol-30 sontensemencées en milieu de culture standard (DMEM sans pyruvate contenant 4.5g/l de glucose et supplémenté par 5% de SVF, 100 U/ml de pénicilline et 100 µg/ml de streptomycine) sous la forme d'une monocouche sous-confluente (1 million de cellules pour chaque 100 cm<sup>2</sup> dans 20 ml de milieu). A confluence, le milieu standard est remplacé par le milieu de sécrétion où la quantité de SVF est abaissée à 0.5% et la quantité de milieu réduite à 16 ml pour chaque 100 cm<sup>2</sup>. Le surnageant de culture est prélevé après 4 à 5 jours d'incubation à 35°C et sous 5% de CO<sub>2</sub>. Le polypeptide recombinant Ssol est purifié à partir du surnageant par l'enchaînement d'étapes de filtration sur membrane de polyethersulfone (PES) de 0.1 µm, de concentration par ultrafiltration sur une membrane de PES de point de coupure 50kD, de chromatographie d'affinité sur matrice anti-FLAG avec élution par une solution de peptide FLAG (DYKDDDDK) à 100 µg/ml en TBS (Tris 50mM pH 7.4, 150 mMNaCl) puis de chromatographie de gel filtration en TBS sur billes de sephadex G-75 (Pharmacia). La concentration du polypeptide recombinant Ssol purifié a été déterminée par test micro-BCA (Pierce) puis ses caractéristiques biochimiques analysées.

**[0337]** L'analyse par gel SDS à 8 % d'acrylamide coloré au nitrate d'argent met en évidence un polypeptide majoritaire dont la masse moléculaire est d'environ 180kD et dont le degré de pureté peut être évalué à 98% (figure 27A). Par spectrométrie de masse SELDI-TOF (Cypherger), deux pics principaux sont mis en évidence : ils correspondent à des formes simplement, et doublement chargées d'un polypeptide majoritaire dont la masse moléculaire est ainsi déterminée à 182,6 ± 3,7 kD (figures 27B et C). Après transfert sur membrane Prosorb et rincage en TFA 0,1%, l'extrémité N-terminale du polypeptide Ssol a été séquencée en phase liquide par dégradation d'Edman sur 5 résidus (ABI494, Applied Biosystems) et déterminée comme étant SDLDR (figure 27D). Ceci démontre que le peptide signal localisé à l'extrémité N-terminale de la protéine S du SRAS-CoV, composé des aa 1 à 13 (MFIFLLFLTLTSG) d'après une analyse réalisée avec le logiciel signalP v2.0 (Nielsen et al., 1997, Protein Engineering, 10 : 1-6), est clivé du polypeptide Ssol mature. Le polypeptide recombinant Ssol est donc constitué des acides aminés 14 à 1193 de la protéine S du SRAS-CoV fusionnés en C-terminal à une séquence SGDYKDDDDK contenant la séquence de l'étiquette FLAG (soulignée). L'écart entre la masse molaire théorique du polypeptide Ssol nu (132,0 kD) et la masse molaire réelle du polypeptide mature (182,6 kD) suggère que le polypeptide Ssol est glycosylé.

**[0338]** Une préparation de polypeptide Ssol purifié et dont la concentration protéique a été déterminée par test micro-BCA, permet de réaliser une gamme étalon pour mesurer à l'aide du test ELISA-capture décrit plus haut les concentrations de Ssol présent dans les surnageants de culture et de revisiter les caractéristiques des lignées sécrétrices. Selon ce test, la lignée FRhK4-Ssol-CT3 sécrète 4 à 6 µg/ml de polypeptide Ssol tandis que la lignée FRhK4-Ssol-30 sécrète 9 à 13 µg/ml de Ssol après 4 à 5 jours de culture à confluence. En outre, le schéma de purification présenté plus haut permet en routine de purifier de 1 à 2 mg de polypeptide Ssol par litre de surnageant de culture.

#### **Exemple 12 : Immunisation génique visant la protéine de spicule (S) du coronavirus associé au SRAS (SRAS-CoV)**

**[0339]** L'effet d'un signal d'épissage et des signaux post-transcriptionnels WPRE et CTE a été analysé après immunisation génique de souris BALB/c (figure 28).

**[0340]** Pour cela, des souris BALB/c ont été immunisées à des intervalles de 4 semaines par injection dans le *tibialis anterior* d'une solution saline de 50 µg d'ADN plasmidique de pcDNA-S et pCI-S ainsi que, à titre de contrôle, par 50 µg d'ADN plasmidique de pcDNA-N (dirigeant l'expression de la N du SRAS-CoV) ou de pCI-HA (dirigeant l'expression de la HA du virus grippal A/PR/8/34) et les sérums immuns collectés 3 semaines après la 2ème injection. La présence d'anticorps dirigés contre la S du SRAS-CoV a été recherchée par ELISA indirect en utilisant comme antigène un lysat de cellules VeroE6 infectées par le SRAS-CoV et à titre de contrôle, un lysat de cellules VeroE6 non infectées. Les titres (TI) en anticorps anti-SRAS-CoV sont calculés comme l'inverse de la dilution produisant une DO spécifique de 0,5

(différence entre DO mesurée sur lysat de cellules infectées et DO mesurée sur lysat de cellules non infectées) après révélation par un anticorps polyclonal anti-IgG de souris couplé à la peroxidase (NA931V, Amersham) et du TMB additionné de H<sub>2</sub>O<sub>2</sub> (KPL) (figure 28A).

**[0341]** Dans ces conditions, le plasmide d'expression pcDNA-S ne permet l'induction que de faibles titres d'anticorps dirigées contre la S du SRAS-CoV chez 3 souris sur 6 ( $\text{LOG}_{10}(\text{TI})=1,9\pm0,6$ ) alors que le plasmide pcDNA-N permet l'induction d'anticorps anti-N à des titres élevés ( $\text{LOG}_{10}(\text{TI})=3,9\pm0,3$ ) chez tous les animaux, et les plasmides contrôles (pCI, pCI-HA) n'entraînent aucun anticorps détectable ( $\text{LOG}_{10}(\text{TI})<1,7$ ). Le plasmide pCI-S muni d'un signal d'épissage permet l'induction d'anticorps à des titres élevés ( $\text{LOG}_{10}(\text{TI})=3,7\pm0,2$ ), qui sont environ 60 fois supérieurs à ceux observés après injection du plasmide pcDNA-S ( $p<10^{-5}$ ).

**[0342]** L'efficacité des signaux post-transcriptionnels a été étudiée en réalisant une étude dose-réponse des titres en anticorps anti-S induits chez la souris BALB/c en fonction de la quantité d'ADN plasmidique utilisé comme immunogène (2 µg, 10 µg et 50 µg). Cette étude (figure 28B) démontre que le signal post-transcriptionnel WPRE améliore fortement l'efficacité de l'immunisation génique lorsque de faibles doses d'ADN sont utilisées ( $p<10^{-5}$  pour une dose de 2 µg d'ADN et  $p<10^{-2}$  pour une dose de 10 µg), alors que l'effet du signal CTE reste marginal ( $p=0,34$  pour une dose de 2 µg d'ADN).

**[0343]** Enfin, les anticorps induits chez la souris après immunisation génique neutralisent l'infectivité du SRAS-CoV *in vitro* (figures 29A et 29B) à des titres qui sont en rapport avec les titres mesurés par ELISA.

**[0344]** En résumé, l'utilisation d'un signal d'épissage et du signal post-transcriptionnel WPRE du virus de l'hépatite de la marmotte améliore de façon considérable l'induction d'anticorps neutralisants dirigés contre le SRAS-CoV après immunisation génique à l'aide d'ADN plasmidique dirigeant l'expression du cADN de la S du SRAS-CoV.

### Exemple 13 : Applications diagnostiques de la protéine S

**[0345]** La réactivité en ELISA du polypeptide recombinant Ssol a été analysée vis-à-vis de sérums de patients atteints de SRAS.

**[0346]** Les sérums de cas probables de SRAS testés ont été choisis sur la base des résultats (positifs ou négatifs) d'analyse de leur réactivité spécifique vis-à-vis des antigènes natifs du SRAS-CoV par test d'immunofluorescence sur cellules VeroE6 infectées par le SRAS-CoV et/ou par test ELISA indirect en utilisant comme antigène un lysat de cellules VeroE6 infectées par le SRAS-CoV. Les sérums de ces patients sont identifiés par un numéro d'ordre du Centre National de Référence des Virus Influenzae ainsi que par les initiales du patient et le nombre de jours écoulés depuis le début des symptômes. Tous les sérums de cas probables (cf Tableau XII) reconnaissent les antigènes natifs du SRAS-CoV, à l'exception du sérum 032552 du patient VTT, pour lequel l'infection par le SRAS-CoV n'a pas pu être confirmée par RT-PCR réalisée sur prélèvements respiratoires des jours 3, 8 et 12. Un panel de sérums témoin a été utilisé à titre de contrôle (sérums TV): il s'agit de sérums prélevés en France avant l'épidémie de SRAS survenue en 2003.

**Tableau XII : sérums de cas probables de SRAS**

sérum	patient	jour de prélèvement
031724	JYK	7
033168	JYK	38
033597	JYK	74
032632	NTM	17
032634	THA	15
032541	PHV	10
032542	NIH	17
032552	VTT	8
032633	PTU	16
032791	JLB	3
033258	JLB	27
032703	JCM	8
033153	JCM	29

**[0347]** Des phases solides sensibilisées par le polypeptide recombinant Ssol ont été préparées par adsorption d'une solution de polypeptide Ssol purifié à 2 µg/ml en PBS dans les puits d'une plaque ELISA, puis les plaques sont incubées une nuit à 4°C et lavées avec du tampon PBS-Tween (PBS, 0,1% Tween20). Après saturation des plaques ELISA par une solution de PBS-lait écrémé à 10% (poids/volume) et lavage en PBS-tween, les sérums à tester (100 µl) sont dilués au 1/400 dans du tampon PBS-lait écrémé-Tween (PBS, 3% lait écrémé, 0,1% Tween) puis ajoutés dans les puits de la plaque ELISA sensibilisée. Les plaques sont incubées 1h à 37°C. Après 3 lavages avec du tampon PBS-Tween, le conjugué anti-IgG humaines marqué à la peroxidase (réf NA933V, Amersham) dilué au 1/4000 dans du tampon PBS-lait écrémé-Tween est ajouté, puis les plaques sont incubées une heure à 37°C. Après 6 lavages avec du tampon PBS-Tween, le chromogène (TMB) et le substrat (H<sub>2</sub>O<sub>2</sub>) sont ajoutés et les plaques sont incubées 10 minutes à l'abri de la lumière. La réaction est arrêtée par ajout d'une solution 1N de H<sub>3</sub>PO<sub>4</sub>, puis l'absorbance est mesurée à 450 nm avec un référence à 620nm.

**[0348]** Les tests ELISA (figure 30) démontrent que le polypeptide recombinant Ssol est reconnu spécifiquement par les anticorps sériques de patients atteints de SRAS prélevés en phase moyenne ou tardive de l'infection (≥ 10 jours après le début des symptômes alors qu'il n'est pas reconnu de façon significative par les anticorps sériques de deux patients (JLB et JCM) prélevés en phase précoce de l'infection (3 à 8 jours après le début des symptômes) ni par des sérums témoins de sujets non atteints de SRAS. Les anticorps sériques des patients JLB et JCM montrent une séro-conversion entre les jours 3 et 27 pour le premier et 8 et 29 pour le second après le début des symptômes, ce qui confirme la spécificité de la réactivité de ces sérums vis-à-vis du polypeptide Ssol.

**[0349]** En conclusion, ces résultats démontrent que le polypeptide recombinant Ssol peut être utilisé comme antigène pour la mise au point d'un test ELISA de diagnostic sérologique de l'infection par le SRAS-CoV.

#### **Exemple 14 : Applications vaccinales de la protéine S soluble recombinante**

**[0350]** L'immunogénicité du polypeptide recombinant Ssol a été étudiée chez la souris.

**[0351]** Pour cela, un groupe de 6 souris a été immunisé à 3 semaines d'intervalle avec 10 µg de polypeptide recombinant Ssol adjuvé par 1 mg d'hydroxyde d'aluminium (Alu-gel-S, Serva) dilué en PBS. Trois immunisations successives ont été réalisées et les sérums immuns ont été prélevés 3 semaines après chacune des immunisations (IS1, IS2, IS3). A titre de contrôle, un groupe de souris (groupe mock) a reçu de l'hydroxyde d'aluminium seul selon le même protocole.

**[0352]** Les sérums immuns ont été analysés par pool pour chacun des 2 groupes par ELISA indirect en utilisant un lysat de cellules VeroE6 infectées par le SRAS-CoV comme antigène et à titre de contrôle, un lysat de cellules VeroE6 non infectées. Les titres en anticorps anti-SRAS-CoV sont calculés comme l'inverse de la dilution produisant une DO spécifique de 0,5 après révélation par un anticorps polyclonal anti-IgG(H+L) de souris couplé à la peroxidase (NA931 V, Amersham) et du TMB additionné de H<sub>2</sub>O<sub>2</sub> (KPL). Cette analyse (figure 31) montre que l'immunisation par le polypeptide Ssol induit chez la souris dès la première immunisation des anticorps dirigés contre la forme native de la protéine de spicule du SRAS-CoV présente dans le lysat de cellules VeroE6 infectées. Après 2 puis 3 immunisations, les titres en anticorps anti-S deviennent très élevés.

**[0353]** Les sérums immuns ont été analysés par pool pour chacun des 2 groupes pour leur capacité à séroneutraliser l'infektivité du SRAS-CoV. 4 points de séroneutralisation sur cellules FRhK-4 (100 TCID<sub>50</sub> de SRAS-CoV) sont réalisés pour chacune des dilutions de raison 2 testées à partir du 1/20. Le titre séroneutralisant est calculé selon la méthode de Reed et Munsch comme l'inverse de la dilution neutralisant l'infektivité de 2 cupules sur 4. Cette analyse montre que les anticorps induits chez la souris par le polypeptide Ssol sont neutralisants : les titres observés sont très élevés après 2 puis 3 immunisations (supérieurs à 2560 et 5120 respectivement, Tableau XIII).

**Tableau XIII : induction d'anticorps dirigés contre le SRAS-CoV après immunisation avec le polypeptide recombinant Ssol.** Les sérums immuns ont été analysés par pool pour chacun des 2 groupes pour leur capacité à séroneutraliser l'infektivité de 100 TCID<sub>50</sub> du SRAS-CoV sur cellules FRhK-4. 4 points sont réalisés pour chacune des dilutions de raison 2 testées à partir du 1/20. Le titre séroneutralisant est calculé selon la méthode de Reed et Munsch comme l'inverse de la dilution neutralisant l'infektivité de 2 cupules sur 4.

groupe	sérums	Ac. Neutralisants
Mock	pi	< 20
	IS1	< 20
	IS2	< 20
	IS3	< 20

(suite)

groupe	sérums	Ac. Neutralisants
Ssol	pi	< 20
	IS1	57
	IS2	> 2560
	IS3	> 5120

**[0354]** Les titres neutralisants observés chez les souris immunisées avec le polypeptide Ssol atteignent des niveaux très supérieurs aux titres observés par Yang et coll. chez la souris (2004, Nature, 428: 561-564) et à ceux observés par Buchholz chez le hamster (2004, PNAS 101 : 9804-9809), qui protègent respectivement la souris et le hamster de l'infection par le SRAS-CoV. Il est donc vraisemblable que les anticorps neutralisants induits chez la souris après immunisation par le polypeptide Ssol protègent ces animaux contre l'infection par le SRAS-CoV.

**Exemple 15 : Gène synthétique optimisé pour l'expression en cellules de mammifères de la protéine de spicule (S) du coronavirus associé au SRAS (SRAS-CoV).**

**1) Conception du gène synthétique**

**[0355]** Un gène synthétique codant pour la protéine de spicule du SRAS-CoV a été conçu à partir du gène de l'isolat 031589 (plasmide pSARS-S, C.N.C.M. n° I-3059) de façon à permettre des niveaux d'expression élevés dans des cellules de mammifères et en particulier dans les cellules d'origine humaine.

**[0356]** Pour cela :

- l'usage des codons du gène sauvage de l'isolat 031589 a été modifié de façon à se rapprocher du biais observé chez l'homme et à améliorer l'efficacité de traduction du mRNA correspondant
- le contenu global en GC du gène a été augmenté de façon à prolonger la demi-vie du mRNA correspondant
- les motifs, éventuellement cryptiques, susceptibles d'interférer avec une expression efficace du gène ont été supprimés (sites donneurs et accepteurs d'épissage, signaux de polyadénylation, séquences très riches (>80%) ou très pauvres (<30%) en GC, séquences répétées, séquences impliquées dans la formation de structures secondaires de l'ARN, boîtes TATA)
- un deuxième codon STOP a été ajouté pour permettre une terminaison efficace de la traduction.

**[0357]** En outre, des motifs CpG ont été introduits dans le gène de façon à augmenter son immunogénicité comme vaccin à ADN. Afin de faciliter la manipulation du gène synthétique, deux sites de restriction BamH1 et Xho1 ont été placés de part et d'autre de la phase ouverte de lecture de la protéine S, et les sites de restriction BamH1, Xho1, Nhe1, Kpn1, BspEI et Sal1 ont été évités dans le gène synthétique.

**[0358]** La séquence du gène synthétique conçu (gène 040530) est donnée en SEQ ID No: 140.

**[0359]** Un alignement du gène synthétique 040530 avec la séquence du gène sauvage de l'isolat 031589 du SRAS-CoV déposé à la C.N.C.M. sous le numéro I-3059 (SEQ ID NO : 4, plasmide pSRAS-S) est présenté dans la figure 32.

**2) Constructions plasmidiques**

**[0360]** Le gène synthétique SEQ ID No: 140 a été assemblé à partir d'oligonucléotides synthétiques et cloné entre les sites Kpn1 et Sac1 du plasmide pUC-Kana pour donner le plasmide 040530pUC-kana. La séquence nucléotidique de l'insert du plasmide 040530pUC-kana a été vérifiée par séquençage automatique (Applied).

**[0361]** Un fragment Kpn1-Xho1 contenant le gène synthétique 040530 a été excisé du plasmide 040530pUC-kana et sous-cloné entre les sites Nhe1 et Xho1 du plasmide d'expression pCI (Promega) pour obtenir le plasmide pCI-SSYNTH, déposé à la CNCM le 1<sup>er</sup> décembre 2004, sous le numéro I-3333.

**[0362]** Un gène synthétique codant pour la forme soluble de la protéine S a ensuite été obtenu en fusionnant les séquences synthétiques codant pour l'ectodomaine de la protéine S (acides aminés 1 à 1193) à celles de l'étiquette (FLAG : DYKDDDDK) via un linker BspEI codant pour le dipeptide SG. Pratiquement, un fragment d'ADN codant pour l'ectodomaine de la S du SRAS-CoV a été amplifié par PCR à l'aide des oligonucléotides 5'-ACTAGCTAGC GGATC-CACCA TGTCATCTT CCTG -3' et 5'- AGTATCCGGAC TTGATGTACT GCTCGTACTT GC-3' à partir du plasmide 040530pUC-kana, digéré par Nhe1 et BspEI puis inséré entre les sites uniques Nhe1 et BspEI du plasmide pCI-Ssol, pour donner le plasmide pCI-SCUBE, déposé à la CNCM le 1<sup>er</sup> décembre 2004, sous le numéro 1-3332. (Les plasmides

pCI-Ssol, pCI-Ssol-CTE et pCI-Ssol-WPRE (déposé à la CNCM, le 22 novembre 2004, sous le numéro I-3324) avaient été précédemment obtenus par sous-clonage du fragment Kpn1-Xho1 excisé du plasmide pcDNA-Ssol (voir note technique de la DI 2004-106) entre les sites Nhe1 et Xho1 des plasmides pCI, pCI-S-CTE et pCI-S-WPRE respectivement.)

**[0363]** Les plasmides pCI-Scube et pCI-Ssol codent pour le même polypeptide recombinant Sso1.

### 3) Résultats

**[0364]** La capacité du gène synthétique codant pour la protéine S à diriger efficacement l'expression de la S du SRAS-CoV dans des cellules de mammifères a été comparée à celle du gène sauvage après transfection transitoire de cellules de primates (VeroE6) et de cellules humaines (293T).

**[0365]** Dans l'expérience présentée par la figure 33 et au Tableau XIV, des monocouches de  $5 \times 10^5$  cellules VeroE6 ou  $7 \times 10^5$  cellules 293T en boîtes de Pétri de 35mm ont été transfectées avec 2  $\mu$ g des plasmides pCI (à titre de contrôle), pCI-S, pCI-S-CTE, pCI-S-WPRE et pCI-Ssynth et 6  $\mu$ l de réactif Eugene6 selon les indications du fabricant (Roche). Après 48 heures d'incubation à 37°C et sous 5% de CO<sub>2</sub>, des extraits cellulaires ont été préparés en tampon de dépôt selon Laemmli, séparés sur un gel SDS à 8% de polyacrylamide puis transférés sur une membrane de PVDF (BioRad). La détection de cette immunoempreinte (« western blot ») a été réalisée à l'aide d'un sérum polyclonal de lapin anti-S (immun-sérum du lapin P11135 : cf exemple 4 ci-dessus) et d'anticorps polyclonaux d'âne dirigés contre les IgG de lapin et couplés à la peroxidase (NA934V, Amersham). L'immunoempreinte est révélée de façon quantitative par luminescence à l'aide du kit ECL+ (Amersham) et acquisition sur un dispositif d'imagerie numérique (FluorS, BioRad).

**[0366]** L'analyse des résultats obtenus avec le logiciel QuantityOne v4.2.3 (BioRad) montre que dans cette expérience, le plasmide pCI-Synth permet l'expression transitoire de la protéine S à des niveaux élevés dans les cellules VeroE6 et 293T, alors que le plasmide pCI-S ne permet pas d'induire une expression à des niveaux suffisants pour être détectée. Les niveaux d'expression observés sont de l'ordre de 2 fois supérieurs à ceux observés avec le plasmide pCI-S-WPRE.

**Tableau XIV : utilisation d'un gène synthétique pour l'expression de la S du SRAS-CoV.** Des extraits cellulaires préparés 48 heures après transfection de cellules VeroE6 ou 293T par les plasmides pCI, pCI-S, pCI-S-CTE, pCI-S-WPRE et pCI-Ssynth ont été séparés sur un gel SDS à 8% d'acrylamide et analysés par western blot à l'aide d'un anticorps polyclonal de lapin anti-S et d'un anticorps polyclonal anti-IgG(H+L) de lapin couplé à la peroxidase (NA934V, Amersham). Le western blot est révélé par luminescence (ECL+, Amersham) et acquisition sur un dispositif d'imagerie numérique (FluorS, BioRad). Les niveaux d'expression de la protéine S ont été mesurés en quantifiant les 2 bandes majoritaires repérées sur l'image (voir figure 33) et sont indiqués en fonction d'une échelle arbitraire où la valeur de 1 représente le niveau mesuré après transfection du plasmide pCI-S-WPRE.

plasmide	VeroE6	293T
pCI	0,0	0,0
pCI-S	$\leq 0,1$	$\leq 0,1$
pCI-S-CTE	0,5	$\leq 0,1$
pCI-S-WPRE	1,0	1,0
pCI-Ssynth	1,8	1,9

**[0367]** Dans un second temps, la capacité du gène synthétique Scube à diriger efficacement la synthèse et la sécrétion du polypeptide Ssol par des cellules de mammifères a été comparée à celle du gène sauvage après transfection transitoire de cellules de hamster (BHK-21) et de cellules humaines (293T).

**[0368]** Dans l'expérience présentée par le Tableau XV, des monocouches de  $6 \times 10^5$  cellules BHK-21 et de  $7 \times 10^5$  cellules 293T en boîtes de Pétri de 35mm ont été transfectées avec 2  $\mu$ g des plasmides pCI (à titre de contrôle), pCI-Ssol, pCI-Ssol-CTE, pCI-Ssol-WPRE et pCI-Scube et 6  $\mu$ l de réactif Eugene6 selon les indications du fabricant (Roche). Après 48 heures d'incubation à 37°C et sous 5% de CO<sub>2</sub>, les surnageants cellulaires ont été prélevés et analysés de façon quantitative pour la sécrétion du polypeptide Ssol par un test ELISA-capture spécifique du polypeptide Ssol.

**[0369]** L'analyse des résultats montre que, dans cette expérience, le plasmide pCI-Scube permet l'expression du polypeptide Ssol à des niveaux 8 fois (cellules BHK-21) à 20 fois (cellules 293T) plus élevés que le plasmide pCI-Ssol. Les niveaux d'expression observés sont de l'ordre de 2 fois (cellules 293T) à 5 fois (cellules BHK-21) supérieurs à ceux observés avec le plasmide pCI-Ssol-WPRE.

**Tableau XV : utilisation d'un gène synthétique pour l'expression du polypeptide Ssol.** Les surnageants ont été récoltés 48 heures après transfection de cellules BHK ou 293T par les plasmides pCI, pCI-Ssol, pCI-Ssol-CTE, pCI-Ssol-WPRE et pCI-Scube et analysés de façon quantitative pour la sécrétion du polypeptide Ssol par un test ELISA-capture spécifique du polypeptide Ssol. Les transfections ont été réalisées en dupliqués et les résultats sont indiqués sous la forme de moyennes et écart-types des concentrations de polypeptide Ssol (ng/ml) mesurées dans les surnageants.

Plasmide	BHK	293T
Pci	< 20	< 20
pCI-Ssol	< 20	56 ± 10
pCI-Ssol-CTE	< 20	63 ± 8
pCI-Ssol-WPRE	28 ± 1	531 ± 15
PCI-Scube	152 ± 6	1140 ± 20

**[0370]** En résumé, ces résultats montrent que l'expression dans des cellules de mammifères du gène synthétique 040530 codant pour la S du SRAS-CoV sous la dépendance de séquences promotrices de l'ARN polymérase II est bien plus efficace que celle du gène sauvage de l'isolat 031589. Cette expression est même plus efficace que celle dirigée par le gène sauvage en présence des séquences WPRE du virus de l'hépatite de la marmotte.

#### 4) Applications

**[0371]** L'utilisation du gène synthétique 040530 codant pour la S du SRAS-CoV ou de sa variante Scube codant pour le polypeptide Ssol est susceptible de remplacer avantageusement le gène sauvage dans de nombreuses applications où l'expression de la S est nécessaire à des niveaux élevés. En particulier pour :

- améliorer l'efficacité de l'immunisation génique par des plasmides du type pCI-Ssynth voire pCI-Ssynth-CTE ou pCI-Ssynth-WPRE
- établir de nouvelles lignées cellulaires exprimant des quantités plus élevées de la protéine S ou du polypeptide Ssol à l'aide de vecteurs lentiviraux recombinants porteurs du gène Ssynth ou du gène Scube respectivement
- améliorer l'immunogénicité des vecteurs lentiviraux recombinants permettant l'expression de la protéine S ou du polypeptide Ssol
- améliorer l'immunogénicité de vecteurs vivants permettant l'expression de la protéine S ou du polypeptide Ssol comme des virus recombinants de la vaccine ou des virus rougeole recombinants (voir exemples 16 et 17 ci-après)

**Exemple 16 : Expression de la protéine de spicule (S) du coronavirus associé au SRAS (SRAS-CoV) à l'aide de virus recombinants de la Vaccine.**

**Application vaccinale.**

**Application à la production d'une forme soluble de la protéine de spicule (S) et conception d'un test de sérologie du SRAS.**

#### 1) Introduction

**[0372]** Le but de cet exemple est d'évaluer la capacité de virus recombinants de la vaccine (VV) exprimant différents antigènes du coronavirus associé au SRAS (SRAS-CoV) à constituer de nouveaux candidats vaccins contre le SRAS et un moyen de produire des antigènes recombinants en cellules de mammifères.

**[0373]** Pour cela, les inventeurs se sont intéressés à la protéine de spicule (S) du SRAS-CoV, qui permet d'induire après immunisation génique chez l'animal des anticorps neutralisants l'infektivité du SRAS-CoV, ainsi qu'à une forme soluble et sécrétée de cette protéine, le polypeptide Ssol, qui est composé de l'ectodomaine (aa 1-1193) de la S fusionné en son extrémité C-ter à une étiquette FLAG (DYKDDDDK) via un linker BspEI codant pour le dipeptide SG. Ce polypeptide Ssol présente une antigénicité similaire à celle de la protéine S et permet, après injection à la souris sous la forme d'une protéine purifiée adjuvée en hydroxide d'aluminium, l'induction de titres élevés d'anticorps neutralisants contre le SRAS-CoV.

**[0374]** Les différentes formes du gène S ont été placées sous la dépendance du promoteur du gène 7.5K puis introduites au sein du locus de la thymidine kinase (TK) de la souche Copenhague du virus de la vaccine par double recombinaison homologue *in vivo*. Afin d'améliorer l'immunogénicité des virus vaccine recombinants, un promoteur tardif synthétique a été choisi à la place du promoteur 7.5K, pour augmenter la production de S et Ssol au cours des phases tardives du cycle viral.

**[0375]** Après avoir isolé les virus vaccine recombinants et vérifié leur capacité à exprimer l'antigène S du SRAS-CoV, leur capacité à induire chez la souris une réponse immunitaire contre le SRAS a été testée. Après avoir purifié l'antigène Ssol du surnageant de cellules infectées, un test ELISA de sérodiagnostic du SRAS a été conçu, et son efficacité a été évaluée à l'aide de sérums de cas probables de SRAS.

## 2) Construction des virus recombinants

**[0376]** Des virus recombinants de la vaccine dirigeant l'expression de la glycoprotéine S de l'isolat 031589 du SRAS-CoV et d'une forme soluble et sécrétée de cette protéine, le polypeptide Ssol, sous la dépendance du promoteur 7.5K ont été obtenus. Dans le but d'augmenter les niveaux d'expression de S et Ssol, des virus recombinants dans lesquels les cDNA de S et de Ssol sont placés sous la dépendance d'un promoteur synthétique tardif ont également été obtenus.

**[0377]** Le plasmide pTG186poly est un plasmide de transfert pour la construction de virus recombinants de la vaccine (Kieny, 1986, Biotechnology, 4 :790-795). A ce titre, il contient le gène de la thymidine kinase du VV dans lequel a été inséré le promoteur du gène 7.5K suivi d'un site multiple de clonage permettant l'insertion de gènes hétérologues (figure 34A). Le promoteur du gène 7.5K contient en fait un tandem de deux séquences promotrices actives respectivement durant les phases précoces ( $P_E$ ) et tardives ( $P_L$ ) du cycle de réplication du virus de la vaccine. Les fragments BamH1-Xho1 ont été excisés des plasmides pTRIP-S et pcDNA-Ssol respectivement et insérés entre les sites BamH1 et Sma1 du plasmide pTG186poly pour donner les plasmides pTG-S et pTG-Ssol (figure 34A). Les plasmides pTG-S et pTG-Ssol ont été déposés à la CNCM, le 2 décembre 2004, sous les numéros I-3338 et I-3339, respectivement.

**[0378]** Les plasmides pTN480, pTN-S et pTN-Ssol ont été obtenus à partir des plasmides pTG186poly, pTG-S et pTG-Ssol respectivement, en substituant le fragment Nde1-Pst1 contenant le promoteur 7.5K par un fragment d'ADN contenant le promoteur tardif synthétique 480, qui a été obtenu par hybridation des oligonucléotides 5'- TATGAGCTTT TTTTTTTTTT TTTTTTTGGC ATATAAATAG ACTCGGCGCG CCATCTGCA-3' et 5'- GATGGCGCGC CGAGTCTATT TATATGCCAA AAAAAAAAAA AAA.AAAAAGC TCA-3' (figure 34B). L'insert a été séquencé à l'aide d'un kit BigDye Terminator v1.1 (Applied Biosystems) et d'un séquenceur automatique ABI377. La séquence du promoteur synthétique tardif 480 tel que cloné dans les plasmides de transfert de la série pTN est indiquée figure 34C. Les plasmides pTN-S et pTN-Ssol ont été déposés à la CNCM, le 2 décembre 2004, sous les numéros I-3340 et I-3341, respectivement.

**[0379]** Les virus recombinants de la vaccine ont été obtenus par double recombinaison homologue *in vivo* entre la cassette TK des plasmides de transfert des séries pTG et pTN et le gène TK de la souche Copenhague du virus de la vaccine selon une procédure décrite par Kieny et coll. (1984, Nature, 312: 163-166). Brièvement, des cellules CV-1 sont transfectées à l'aide de DOTAP (Roche) par de l'ADN génomique de la souche copenhague du virus de la vaccine et chacun des plasmides de transfert des séries pTG et pTN décrits ci-dessus, puis surinfectées par le virus vaccine auxiliaire VV-ts7 pendant 24 heures à 33°C. Le virus auxiliaire est contre-sélectionné par incubation à 40°C pendant 2 jours puis les virus recombinants (phénotype TK-) sélectionnés par deux cycles de clonage sous milieu gélosé sur cellules 143Btk- en présence de BuDr (25 µg/ml). Les 6 virus VV-TG, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S et VV-TN-Ssol ont été respectivement obtenus à l'aide des plasmides de transfert pTG186poly, pTG-S, pTG-Ssol, pTN480, pTN-S, pTN-Ssol. Les virus VV-TG et VV-TN n'expriment aucun gène hétérologue et ont été utilisés comme contrôle TK- dans les expériences. Les préparations de virus recombinants ont été réalisés sur monocouches de cellules CV-1 ou BHK-21 et le titre en unités formant plaque (u.f.p.) déterminé sur cellules CV-1 selon Earl et Moss (1998, Current Protocols in Molecular Biology, 16.16.1-16.16.13).

## 3) Caractérisation des virus recombinants

**[0380]** L'expression des transgènes codant la protéine S et le polypeptide Ssol a été recherchée par western blot.

**[0381]** Des monocouches de cellules CV-1 ont été infectées à une multiplicité de 2 par les différents virus vaccine recombinants VV-TG, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S et VV-TN-Ssol. Après 18 heures d'incubation à 37°C et sous 5% de CO<sub>2</sub>, des extraits cellulaires ont été préparés en tampon de dépôt selon Laemmli, séparés sur un gel SDS à 8% de polyacrylamide puis transférés sur une membrane de PVDF (BioRad). La détection de cette immunoempreinte (« western blot ») a été réalisée à l'aide d'un sérum polyclonal de lapin anti-S (immun-sérum du lapin P11135 : cf exemple 4) et d'anticorps polyclonaux d'âne dirigés contre les IgG de lapin et couplés à la peroxidase (NA934V, Amersham). Les anticorps fixés ont été révélés par luminescence à l'aide du kit ECL+ (Amersham) et de films d'autoradiographie Hyperfilm MP (Amersham).

**[0382]** Comme le montre la figure 35A, le virus recombinant VV-TN-S dirige l'expression de la protéine S à des niveaux



qui sont comparables à ceux que l'on peut observer 8h après infection par le SRAS-CoV mais qui sont bien plus élevés que ceux que l'on peut observer après infection par W-TG-S. Dans une deuxième expérience (figure 35B), l'analyse de quantités variables d'extraits cellulaires montre que les niveaux d'expression observés après infection par les virus de la série TN (VV-TN-S et VV-TN-Ssol) sont environ 10 fois plus élevés que ceux observés avec les virus de la série TG (VV-TG-S et VV-TG-Ssol respectivement). En outre, le polypeptide Ssol est sécrété dans le surnageant de cellules CV-1 infectées par le virus VV-TN-Ssol plus efficacement que dans le surnageant de cellules infectées par VV-TG-Ssol (figure 36A). Dans cette expérience, le virus VV-TN-Sflag a été utilisé à titre de contrôle, car il exprime la forme membranaire de la protéine S fusionnée en son extrémité C-ter à l'étiquette FLAG. La protéine Sflag n'est pas détectée dans le surnageant des cellules infectées par VV-TN-Sflag, démontrant que le polypeptide Ssol est bien sécrété de façon active après infection par VV-TN-Ssol.

**[0383]** Ces résultats démontrent que les virus vaccine recombinants sont bien porteurs des transgènes et permettent l'expression de la glycoprotéine du SRAS sous sa forme membranaire (S) ou sous une forme soluble et sécrétée (Ssol). Les virus vaccine porteurs du promoteur synthétique 480 permettent l'expression de S et la sécrétion de Ssol à des niveaux bien plus élevés que les virus porteurs du promoteur du gène 7.5K.

#### 4) Application à la production d'une forme soluble de la S du SRAS-CoV. Purification de cet antigène recombinant et applications diagnostiques.

**[0384]** La lignée BHK-21 est la lignée cellulaire qui sécrète les quantités les plus élevées de polypeptide Ssol après infection par le virus VV-TN-Ssol parmi les lignées testées (BHK-21, CV1, 293T et FrhK-4, figure 36B); elle permet la production et la purification en quantité du polypeptide recombinant Ssol. Dans une expérience type où les conditions expérimentales d'infection, de production et de purification ont été optimisées, les cellules BHK-21 sont ensemencées en milieu de culture standard (DMEM sans pyruvate contenant 4.5g/l de glucose et supplémenté par 5% de TPB, 5% de SVF, 100 U/ml de pénicilline et 100 µg/ml de streptomycine) sous la forme d'une monocouche sous-confluente (10 millions de cellules pour chaque 100 cm<sup>2</sup> dans 25 ml de milieu). Après 24h d'incubation à 37°C sous 5% de CO<sub>2</sub>, les cellules sont infectées à une M.O.I. de 0.03 et le milieu standard remplacé par le milieu de sécrétion où la quantité de SVF est abaissée à 0.5% et le TPB supprimé. Le surnageant de culture est prélevé après 2,5 jours d'incubation à 35°C et sous 5% de CO<sub>2</sub> et le virus de la vaccine inactivé par addition de triton X-100 (0,1%). Après filtration sur membrane de polyethersulfone (PES) de 0.1 µm, le polypeptide recombinant Ssol est purifié par une chromatographie d'affinité sur matrice anti-FLAG avec élution par une solution de peptide FLAG (DYKDDDDK) à 100 µg/ml en TBS (Tris 50mM pH 7.4, 150 mMNaCl).

**[0385]** L'analyse par gel SDS à 8 % d'acrylamide coloré au nitrate d'argent a mis en évidence un polypeptide majoritaire dont la masse moléculaire est d'environ 180kD et dont le degré de pureté est supérieur à 90% (figure 37). La concentration du polypeptide recombinant Ssol purifié a été déterminée par comparaison avec les marqueurs de masse moléculaire et estimée à 24 ng/µl.

**[0386]** Cette préparation de polypeptide Ssol purifié permet de réaliser une gamme étalon pour mesurer à l'aide d'un test ELISA-capture les concentrations de Ssol présentes dans les surnageants de culture. Selon ce test, la lignée BHK-21 sécrète environ 1 µg/ml de polypeptide Ssol dans les conditions de production décrites plus haut. En outre, le schéma de purification présenté permet de purifier de l'ordre de 160 µg de polypeptide Ssol par litre de surnageant de culture.

**[0387]** La réactivité en ELISA du polypeptide recombinant Ssol a été analysée vis-à-vis de sérums de patients atteints de SRAS.

**[0388]** Les sérums de cas probables de SRAS testés ont été choisis sur la base des résultats (positifs ou négatifs) d'analyse de leur réactivité spécifique vis à vis des antigènes natifs du SRAS-CoV par test d'immunofluorescence sur cellules VeroE6 infectées par le SRAS-CoV et/ou par test ELISA indirect en utilisant comme antigène un lysat de cellules VeroE6 infectées par le SRAS-CoV. Les sérums de ces patients sont identifiés par un numéro d'ordre du Centre National de Référence des Virus Influenzae ainsi que par les initiales du patient et le nombre de jours écoulés depuis le début des symptômes. Tous les sérums de cas probables (cf Tableau XVI) reconnaissent les antigènes natifs du SRAS-CoV à l'exception du sérum 032552 du patient VTT, pour lequel l'infection par le SRAS-CoV n'a pas pu être confirmée par RT-PCR réalisée sur prélèvements respiratoires des jours 3, 8 et 12. Un panel de sérums témoin a été utilisé à titre de contrôle (sérums TV): il s'agit de sérums prélevés en France avant l'épidémie de SRAS survenue en 2003.

**Tableau XVI : sérums de cas probables de SRAS**

sérum	patient	jour de prélèvement
033168	JYK	38
033597	JYK	74
032632	NTM	17

(suite)

sérum	patient	jour de prélèvement
032634	THA	15
032541	PHV	10
032542	NIH	17
032552	VTT	8
032633	PTU	16

**[0389]** Des phases solides sensibilisées par le polypeptide recombinant Ssol ont été préparées par adsorption d'une solution de polypeptide Ssol purifié à 4 µg/ml en PBS dans les puits d'une plaque ELISA. Les plaques sont incubées une nuit à 4°C puis lavées avec du tampon PBS-Tween (PBS, 0,1% Tween20). Après lavage en PBS-tween, les sérums à tester (100 µl) sont dilués au 1/100 et 1/400 dans du tampon PBS-lait écrémé-Tween (PBS, 3% lait écrémé, 0,1% Tween) puis ajoutés dans les puits de la plaque ELISA sensibilisée. Les plaques sont incubées 1h à 37°C. Après 3 lavages avec du tampon PBS-Tween, le conjugué anti-IgG humaines marqué à la peroxidase (réf NA933V, Amersham) dilué au 1/4000 dans du tampon PBS-lait écrémé-Tween est ajouté puis les plaques sont incubées une heure à 37°C. Après 6 lavages avec du tampon PBS-Tween, le chromogène (TMB) et le substrat (H<sub>2</sub>O<sub>2</sub>) sont ajoutés et les plaques sont incubées 10 minutes à l'abri de la lumière. La réaction est arrêtée par ajout d'une solution 1M de H<sub>3</sub>PO<sub>4</sub> puis l'absorbance est mesurée à 450 nm avec une référence à 620nm.

**[0390]** Les tests ELISA (figure 38) démontrent que le polypeptide recombinant Ssol est reconnu spécifiquement par les anticorps sériques de patients atteints de SRAS prélevés en phase moyenne ou tardive de l'infection (≥ 10 jours après le début des symptômes), alors qu'il n'est pas reconnu de façon significative par les anticorps sériques des sérums témoins de sujets non atteints de SRAS.

**[0391]** En conclusion, ces résultats démontrent que le polypeptide recombinant Ssol peut être purifié à partir du surnageant de cellules de mammifères infectées par le virus vaccine recombinant W-TN-Ssol et être utilisé comme antigène pour la mise au point d'un test ELISA de diagnostic sérologique de l'infection par le SRAS-CoV.

## 5) Applications vaccinales

**[0392]** L'immunogénicité des virus vaccine recombinants a été étudiée chez la souris.

**[0393]** Pour cela, des groupes de 7 souris BALB/c ont été immunisés par voie i.v. à deux reprises à 4 semaines d'intervalle par 10<sup>6</sup> u.f.p. de virus vaccine recombinants VV-TG, VV-TG-S, W-TG-Ssol, W-TN, W-TN-S et VV-TN-Ssol ainsi que, à titre de contrôle, VV-TG-HA qui dirige l'expression de l'hémagglutinine de la souche A/PR/8/34 du virus de la grippe. Les sérums immuns ont été prélevés 3 semaines après chacune des immunisations (IS1, IS2).

**[0394]** Les sérums immuns ont été analysés par pool pour chacun des groupes par ELISA indirect en utilisant un lysat de cellules VeroE6 infectées par le SRAS-CoV comme antigène et à titre de contrôle, un lysat de cellules VeroE6 non infectées. Les titres (TI) en anticorps anti-SRAS-CoV sont calculés comme l'inverse de la dilution produisant une DO spécifique de 0,5 après révélation par un anticorps polyclonal anti-IgG(H+L) de souris couplé à la peroxidase (NA931V, Amersham) et du TMB additionné de H<sub>2</sub>O<sub>2</sub> (KPL). Cette analyse (figure 39A) montre que l'immunisation par le virus VV-TG-S et W-TN-S induit chez la souris dès la première immunisation des anticorps dirigés contre la forme native de la protéine de spicule du SRAS-CoV présente dans le lysat de cellules VeroE6 infectées. Les réponses induites par le virus VV-TN-S sont plus élevées que celles induites par le virus VV-TG-S après la première (TI=740 et TI=270 respectivement) et la deuxième (TI=3230 et TI=600 respectivement) immunisation. Le virus VV-TN-Ssol induit de forts titres d'anticorps anti-SRAS-CoV après deux immunisations (TI=640), alors que le virus VV-TG-Ssol induit une réponse à la limite de la détection (TI=40).

**[0395]** Les sérums immuns ont été analysés par pool pour chacun des groupes pour leur capacité à séroneutraliser l'infektivité du SRAS-CoV. 4 points de séroneutralisation sur cellules FRhK-4 (100 TCID<sub>50</sub> de SRAS-CoV) sont réalisés pour chacune des dilutions de raison 2 testées à partir du 1/20. Le titre séroneutralisant est calculé selon la méthode de Reed et Munsch comme l'inverse de la dilution neutralisant l'infektivité de 2 cupules sur 4. Cette analyse montre que les anticorps induits chez la souris par les virus vaccine exprimant la protéine S ou le polypeptide Ssol sont neutralisants et que les virus à promoteurs synthétiques sont des immunogènes plus efficaces que les virus porteurs du promoteur 7.5K : les titres les plus élevés (640) sont observés après 2 immunisations par le virus VV-TN-S (figure 39B).

**[0396]** Le pouvoir protecteur des anticorps neutralisants induits chez la souris après immunisation par les virus vaccine recombinants est évalué à l'aide d'une infection d'épreuve par le SRAS-CoV.

## 6) Autres applications

**[0397]** Des virus vaccine recombinants de troisième génération sont construits en substituant les séquences sauvages des gènes S et Ssol par des gènes synthétiques optimisés pour l'expression en cellules de mammifères, décrits ci-dessus. Ces virus vaccine recombinants sont susceptibles d'exprimer des quantités plus importantes des antigènes S et Ssol et donc de présenter une immunogénicité accrue.

**[0398]** Le virus vaccine recombinant VV-TN-Ssol peut être utilisé pour la production en quantité et la purification de l'antigène Ssol en vue d'applications diagnostiques (sérologie par ELISA) et vaccinales (vaccin sous-unitaire).

**Exemple 17 : Virus recombinant de la rougeole exprimant la protéine de spicule (S) du coronavirus associé au SRAS (SRAS-CoV). Applications vaccinales.**

### 1) Introduction

**[0399]** Le vaccin rougeole (MV) induit chez l'homme une immunité protectrice de longue durée après une seule injection (Hilleman, 2002, Vaccine, 20 : 651-665). La protection conférée est très robuste et repose sur l'induction d'une réponse en anticorps et d'une réponse cellulaire CD4 et CD8. Le génome du MV est très stable et aucune réversion vers la virulence des souches vaccinales n'a jamais été observée. Le virus de la rougeole appartient au genre des *Morbillivirus* de la famille des *Paramyxoviridae* ; c'est un virus enveloppé dont le génome est un ARN monocaténaire de polarité négative de 16kb (figure 40A) et dont le cycle de réplication exclusivement cytoplasmique exclut toute possibilité d'intégration dans le génome de l'hôte. Le vaccin rougeole est ainsi l'un des vaccins vivants les plus efficaces et les plus sûrs utilisés dans la population humaine. L'équipe de Frédéric Tangy a développé récemment un vecteur d'expression sur la base de la souche Schwarz du virus de la rougeole, qui est la souche atténuée la plus sûre et la plus utilisée chez l'homme comme vaccin contre la rougeole. Cette souche vaccinale peut être isolée à partir d'un clone moléculaire infectieux en conservant son immunogénicité chez les primates ainsi que chez la souris susceptible à l'infection. Elle constitue, après insertion d'unités de transcription supplémentaires, un vecteur pour l'expression de séquences hétérologues (Combredet, 2003, J. Virol. 77 : 11546-11554). En outre, un MV Schwarz recombinant exprimant la glycoprotéine d'enveloppe du virus West Nile (WNV) induit une réponse en anticorps efficace et de longue durée qui protège la souris d'une infection d'épreuve létale par le WNV (Despres et al, 2004, J. Infect. Dis., sous presse). Toutes ces caractéristiques font de la souche atténuée Schwarz du virus de la rougeole un candidat vecteur extrêmement prometteur pour la construction de nouveaux vaccins vivants recombinants.

**[0400]** Le but de cet exemple est d'évaluer la capacité de virus recombinants de la rougeole (MV) exprimant différents antigènes du coronavirus associé au SRAS (SRAS-CoV) à constituer de nouveaux candidats vaccins contre le SRAS.

**[0401]** Les inventeurs se sont intéressés à la protéine de spicule (S) du SRAS-CoV, qui permet d'induire après immunisation génique chez l'animal des anticorps neutralisants l'infectivité du SRAS-CoV, ainsi qu'à une forme soluble et sécrétée de cette protéine, le polypeptide Ssol, qui est composé de l'ectodomaine (aa 1-1193) de la S fusionné en son extrémité C-ter à une étiquette FLAG (DYKDDDDK) via un linker BspEI codant pour le dipeptide SG. Ce polypeptide Ssol présente une antigénicité similaire à celle de la protéine S et permet, après injection à la souris sous la forme d'une protéine purifiée adjuvée en hydroxide d'aluminium, l'induction de titres élevés d'anticorps neutralisants contre le SRAS-CoV.

**[0402]** Les différentes formes du gène S ont été introduites sous la forme d'une unité de transcription supplémentaire entre les gènes P (phosphoprotéine) et M (matrice) dans le cADN de la souche Schwarz du MV précédemment décrite (Combredet, 2003, J. Virol. 77 : 11546-11554 ; Demande EP N° 02291551.6 du 20 juin 2002, et Demande EP n° 02291550.8 du 20 juin 2002). Après avoir isolé les virus recombinants MVSchw2-SARS-S et MVSchw2-SARS-Ssol et vérifié leur capacité à exprimer l'antigène S du SRAS-CoV, leur capacité à induire chez la souris puis chez le singe une réponse immunitaire protectrice contre le SRAS est testée.

### 2) Construction des virus recombinants

**[0403]** Le plasmide pTM-MVSchw-ATU2 (figure 40B) contient un cADN infectieux correspondant à l'antigénome de la souche vaccinale Schwarz du virus de la rougeole (MV) dans lequel une unité de transcription supplémentaire (ATU) a été introduite entre les gènes P (phosphoprotéine) et M (matrice) (Combredet, 2003, Journal of Virology, 77 : 11546-11554). Des génomes recombinants MVSchw2-SARS-S et MVSchw2-SARS-Ssol du virus de la rougeole ont été construits par l'insertion des ORFs de la protéine S et du polypeptide Ssol au sein de l'unité de transcription supplémentaire du vecteur MVSchw-ATU2.

**[0404]** Pour cela, un fragment d'ADN contenant le cADN de la S du SRAS-CoV a été amplifié par PCR à l'aide des oligonucléotides 5'-ATACGTACGA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' et 5'-ATAGCGCGCT CAT-TATGTGT AATGTAATTT GACACCCTTG-3' en utilisant le plasmide pcDNA-S comme matrice puis inséré dans le

plasmide pCR®2.1-TOPO (Invitrogen) pour obtenir le plasmide pTOPO-S-MV. Les deux oligonucléotides utilisés contiennent des sites de restriction BsiW1 et BssHII, de façon à permettre l'insertion ultérieure dans le vecteur rougeole, et ont été conçus de façon à générer une séquence de 3774 nt incluant les codons d'initiation et de terminaison, afin de respecter la règle des 6 qui stipule que la longueur du génome d'un virus rougeole doit être divisible par 6 (Calain & Roux, 1993, J. Virol., 67 : 4822-4830 ; Schneider et al., 1997, Virology, 227 : 314-322). L'insert a été séquencé à l'aide d'un kit BigDye Terminator v1.1 (Applied Biosystems) et d'un séquenceur automatique ABI377.

**[0405]** Afin d'exprimer une forme soluble et sécrétée de la S du SRAS-CoV, un plasmide contenant l'ADNc du polypeptide Ssol correspondant à l'ectodomaine (aa 1-1193) de la S du SRAS-CoV fusionné en son extrémité C-ter à la séquence d'une étiquette FLAG (DYKDDDDK) via un linker BspEI codant pour le dipeptide SG a ensuite été obtenu. Pour cela, un fragment d'ADN a été amplifié à l'aide des oligonucléotides 5'-CCATTCAAC AATTTGGCCG-3' et 5'-ATAGGATCCG CGCGCTCATT ATTTATCGTC GTCATCTTTA TAATC-3' à partir du plasmide pCDNA-Ssol puis inséré dans le plasmide pTOPO-S-MV entre les sites Sall et BamH1 pour obtenir le plasmide pTOPO-S-MV-SF. La séquence générée est longue de 3618 nt entre les sites BsiW1 et BssHII et respecte la règle des 6. L'insert a été séquencé comme indiqué ci-dessus.

**[0406]** Les fragments BsiW1-BssHII contenant les ADNc de la protéine S et du polypeptide Ssol ont ensuite été excisés par digestion des plasmides pTOPO-S-MV et pTOPO-S-MV-SF puis sous-clonés entre les sites correspondants du plasmide pTM-MV Schw-ATU2 pour donner les plasmides pTM-MV Schw2-SARS-S et pTM-MV Schw2-SARS-Ssol (figure 40B). Ces deux plasmides ont été déposés à la C.N.C.M. le 1<sup>er</sup> décembre 2004, sous les numéros I-3326 et I-3327, respectivement.

**[0407]** Les virus rougeole recombinants correspondants aux plasmides pTM-MV Schw2-SARS-S et pTM-MV Schw2-SARS-Ssol ont été obtenus par génétique inverse selon le système reposant sur l'utilisation d'une lignée cellulaire auxiliaire, décrit par Radecke et coll. (1995, Embo J., 14, : 5773-5784) et modifié par Parks et coll. (1999, J. Virol., 73 : 3560-3566). Brièvement, les cellules auxiliaires 293-3-46 sont transfectées selon la méthode au phosphate de calcium par 5 µg des plasmides pTM-MV Schw2-SARS-S ou pTM-MV Schw2-SARS-Ssol et 0,02 µg du plasmide pEMC-La dirigeant l'expression de la polymérase L du MV (don de M.A. Billeter). Après une nuit d'incubation à 37°C, un choc thermique est réalisé pendant 2 heures à 43°C et les cellules transfectées sont transférées sur une monocouche de cellules Vero. Pour chacun des deux plasmides, des syncytia sont apparus après 2 à 3 jours de co-culture et ont été transférés successivement sur des monocouches de cellules Vero à 70% de confluence en boîtes de pétri de 35 mm puis en flacons de 25 et 75 cm<sup>2</sup>. Quand les syncytia ont atteint 80-90% de confluence, les cellules sont récupérées à l'aide d'un grattoir puis congelées et décongelées une fois. Après une centrifugation à basse vitesse, le surnageant contenant le virus est conservé en aliquot à -80°C. Les titres des virus recombinants MV Schw2-SARS-S et MV Schw2-SARS-Ssol ont été déterminés par dilution limite sur cellules Vero et le titre en dose infectant 50% des cupules (TCID<sub>50</sub>) calculé selon la méthode de Kärber.

### 3) Caractérisation des virus recombinants

**[0408]** L'expression des transgènes codant la protéine S et le polypeptide Ssol a été recherchée par western blot et immunofluorescence.

**[0409]** Des monocouches de cellules Vero en flacons T-25 ont été infectées à une multiplicité de 0,05 par différents passages des deux virus MV Schw2-SARS-S et MV Schw2-SARS-Ssol et le virus sauvage MWSchw à titre de contrôle. Quand les syncytia ont atteint 80 à 90% de confluence, des extraits cytoplasmiques ont été préparés dans un tampon d'extraction (150mM NaCl, 50 mM Tris-HCl pH 7,2, 1% triton-X-100, 0,1% SDS, 1% DOC) puis dilués en tampon de dépôt selon Laemmli, séparés sur un gel SDS à 8% de polyacrylamide et transférés sur une membrane de PVDF (BioRad). La détection de cette immunoempreinte (« western blot ») a été réalisée à l'aide d'un sérum polyclonal de lapin anti-S (immun-sérum du lapin P11135 : cf exemple 4 ci-dessus) et d'anticorps polyclonaux d'âne dirigés contre les IgG de lapin et couplés à la peroxidase (NA934V, Amersham). Les anticorps fixés ont été révélés par luminescence à l'aide du kit ECL+ (Amersham) et de films d'autoradiographie Hyperfilm MP (Amersham).

**[0410]** Des cellules Vero en monocouches sur lamelles de verre ont été infectées par les deux virus MV Schw2-SARS-S et MV Schw2-SARS-Ssol et le virus sauvage MWSchw à titre de contrôle à des multiplicités d'infection de 0,05. Quand les syncytia ont atteint 90 à 100% (virus MV Schw2-SARS-Ssol) ou 30 à 40% (MV Schw2-SARS-S, MWSchw) de confluence, les cellules ont été fixées dans une solution de PBS-PFA 4%, perméabilisées par une solution de PBS contenant 0,2% de triton puis marquées par des anticorps polyclonaux de lapins hyperimmunisés par des virions purifiés et inactivés du SRAS-CoV et par un conjugué d'anticorps de chèvre anti-IgG(H+L) de lapin couplé au FITC (Jackson).

**[0411]** Comme le montrent les figures 41 et 42, les virus recombinants MV Schw2-SARS-S et MV Schw2-SARS-Ssol dirigent l'expression de la protéine S et du polypeptide Ssol respectivement à des niveaux comparables à ceux que l'on peut observer 8h après infection par le SRAS-CoV. L'expression de ces polypeptides est stable après 3 passages des virus recombinants en culture cellulaire. Ces résultats démontrent que les virus rougeole recombinants sont bien porteurs des transgènes et permettent l'expression de la glycoprotéine du SRAS sous sa forme membranaire (S) ou sous une

forme soluble (Ssol). On s'attend à ce que le polypeptide Ssol soit sécrété des cellules infectées par le virus MVSchw2-SARS-Ssol comme c'est le cas lorsque ce même polypeptide est exprimé dans des cellules de mammifères après transfection transitoire des séquences correspondantes (cf exemple 11 ci-dessus).

#### 4) Applications

**[0412]** Ayant montré que les virus MVSchw2-SARS-S et MVSchw2-SARS-Ssol permettent l'expression de la S du SRAS-CoV, leur capacité à induire une réponse immunitaire protectrice contre le SRAS-CoV chez la souris CD46<sup>+/-</sup> IFN- $\alpha$ βR<sup>-/-</sup>, qui est susceptible à l'infection par le MV, est évaluée. La réponse en anticorps des souris immunisées est évaluée par test ELISA contre les antigènes natifs du SRAS-CoV et pour leur capacité à neutraliser l'infectivité du SRAS-CoV *in vitro*, en utilisant les méthodologies décrites ci-dessus. Le pouvoir protecteur de la réponse sera évalué en mesurant la réduction de la charge virale pulmonaire 2 jours après une infection d'épreuve non létale par le SRAS-CoV.

**[0413]** Des virus rougeole recombinants de seconde génération sont construits en substituant les séquences sauvages des gènes S et Sol par des gènes synthétiques optimisés pour l'expression en cellules de mammifères, décrits à l'exemple 15 ci-dessus. Ces virus rougeole recombinants sont susceptibles d'exprimer des quantités plus importantes des antigènes S et Ssol et donc de présenter une immunogénicité accrue.

**[0414]** Alternativement, les gènes sauvages ou synthétiques codant pour la protéine S ou le polypeptide Ssol peuvent être insérés dans le vecteur rougeole MVSchw-ATU3 sous la forme d'une unité de transcription supplémentaire localisée entre les gènes H et L, puis les virus recombinants produits et caractérisés de façon similaire. Cette insertion est susceptible de générer des virus recombinants possédant des caractéristiques (multiplication du virus, niveau d'expression du transgène) différentes et possiblement une immunogénicité améliorée par rapport à ceux obtenus après insertion des transgènes entre les gènes P et N.

**[0415]** Le virus rougeole recombinant MVSchw2-SARS-Ssol peut être utilisé pour la production en quantité et la purification de l'antigène Ssol en vue d'applications diagnostiques et vaccinales.

#### Exemple 18 : Autres applications liées à la protéine S

##### **[0416]**

a) Les vecteurs lentiviraux permettant l'expression de la S ou de Ssol (voire de fragments de la S) peuvent constituer un vaccin recombinant contre le SRAS-CoV, pour être utilisé en prophylaxie humaine et vétérinaire. Afin de démontrer la faisabilité d'un tel vaccin, l'immunogénicité des vecteurs lentiviraux recombinants TRIP-SD/SA-S-WPRE et TRIP-SD/SA-Ssol-WPRE est étudiée chez la souris.

b) Des anticorps monoclonaux sont produits à l'aide du polypeptide recombinant Ssol. D'après les résultats présentés à l'exemple 14 ci-dessus, ces anticorps ou du moins la majorité d'entre eux reconnaîtront la forme native de la S du SRAS-CoV et seront susceptibles d'applications diagnostiques et/ou prophylactiques.

c) Un test de sérologie du SRAS est mis au point avec le polypeptide Ssol utilisé comme antigène et la méthodologie du double épitope.

#### SEQUENCE LISTING

##### **[0417]**

<110> INSTITUT PASTEUR  
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE PARIS 7  
VAN DER WERF, Sylvie  
ESCRIOU, Nicolas  
CRESCENZO-CHAIGNE, Bernadette  
MANUGUERRA, Jean-Claude  
KUNST, Franck  
CALLENDRET, Benoît  
BETTON, Jean-Michel  
LORIN, Valérie  
GERBAUD, Sylvie  
BURGUIERE, Ana Maria  
AZEBI, Saliha  
CHARNEAU, Pierre

TANGY, Frédéric  
COMBREDET, Chantal  
DELAGNEAU, Jean-François  
MARTIN, Monique

5

<120> NOUVELLE SOUCHE DE CORONAVIRUS ASSOCIE AU SRAS ET SES APPLICATIONS

<130> 226-108ext

10

<150> FR 0314152

<151> 2003-12-02

<150> FR 0314151

<151> 2003-12-02

15

<160> 158

<170> PatentIn version 3.1

20

<210> 1

<211> 29746

<212> DNA

<213> CORONAVIRUS

25

<400> 1

30

35

40

45

50

55

EP 1 694 829 B1

	atattagggtt	tttacctacc	caggaaaagc	caaccaacct	cgatctcttg	tagatctgtt	60
	ctctaaacga	actttaaaat	ctgtgtagct	gtcgtctcggc	tgcatgccta	gtgcacctac	120
5	gcagtataaa	caataataaa	ttttactgtc	gttgacaaga	aacgagtaac	tcgtccctct	180
	tctgcagact	gcttacgggt	tcgtccgtgt	tgcatgcat	catcagcata	cctagggtttc	240
	gtccgggtgt	gaccgaaagg	taagatggag	agccttggtc	ttggtgtcaa	cgagaaaaca	300
10	cacgtccaac	tcagtttgcc	tgctcttcag	gttagagacg	tgctagtgcg	tggtctcggg	360
	gactctgtgg	aagaggccct	atcggaggca	cgtgaacacc	tcaaaaatgg	cacttggtgt	420
	ctagtagagc	tggaataagg	cgtactgccc	cagcttgaa	agccctatgt	gttcattaaa	480
15	cgctctgatg	ccttaagcac	caatcacggc	cacaaggctg	ttgagctggt	tgagaaatg	540
	gacggcattc	agtacggctg	tagcgggtata	acactgggag	tactcgtgcc	acatgtgggc	600
	gaaaccccaa	ttgcataccg	caatgttctt	cttcgtaaga	acggtaataa	gggagccggt	660
20	ggcatagct	atggcatcga	tctaaagtct	tatgacttag	gtgacgagct	tggtactgat	720
	cccattgaag	attatgaaca	aaactggaac	actaagcatg	gcagtgggtg	actccgtgaa	780
	ctcactcgtg	agctcaatgg	aggtgcagtc	actcgtctatg	tcgacaacaa	tttctgtggc	840
25	ccagatgggt	accctcttga	ttgcatcaaa	gattttctcg	cacgcgcggg	caagtcaatg	900
	tgactctttt	ccgaacaact	tgattacatc	gagtcgaaga	gaggtgtcta	ctgctgccgt	960
	gaccatgagc	atgaaattgc	ctggttcact	gagcgtctctg	ataagagcta	cgagcaccag	1020
30	acacccttcg	aaattaagag	tgccaagaaa	tttgacactt	tcaaagggga	atgcccaaag	1080
	tttgtgtttc	ctcttaactc	aaaagtcaaa	gtcattcaac	cacgtgttga	aaagaaaaag	1140
	actgaggggt	tcatggggcg	tatacgctct	gtgtaccctg	ttgcatctcc	acaggagtgt	1200
35	aacaatatgc	acttgctctac	cttgatgaaa	tgtaatcatt	gcgatgaagt	ttcatggcag	1260
	acgtgcgact	ttctgaaagc	cacttggtgaa	cattgtggca	ctgaaaattt	agttattgaa	1320
	ggacctacta	catgtgggta	cctacctact	aatgctgtag	tgaaaatgcc	atgtcctgcc	1380
40	tgtcaagacc	cagagattgg	acctgagcat	agtgttgag	attatcacia	ccactcaaac	1440
	attgaaactc	gactccgcaa	gggaggtagg	actagatgtt	ttggaggctg	tgtgtttgcc	1500
	tatgttggct	gctataataa	gcgtgcctac	tggttccctc	gtgctagtgc	tgatattggc	1560
45	tcaggccata	ctggcattac	tggtgacaat	gtggagacct	tgaatgagga	tctccttgag	1620

50

55

## EP 1 694 829 B1

	atactgagtc	gtgaacgtgt	taacattaac	attggtggcg	attttcattt	gaatgaagag	1680
	gttgccatca	ttttggcatc	tttctctgct	tctacaagtg	cctttattga	cactataaag	1740
5	agtcttgatt	acaagtcttt	caaaaccatt	gttgagtcct	gcggtacta	taaagttacc	1800
	aagggaaaagc	ccgtaaaagg	tgcttggaac	attggacaac	agagatcagt	tttaacacca	1860
	ctgtgtggtt	ttccctcaca	ggctgctggt	gttatcagat	caatttttgc	gcgcacacct	1920
10	gatgcagcaa	accactcaat	tcctgatttg	caaagagcag	ctgtcaccat	acttgatggt	1980
	atttctgaac	agtcattacg	tcttgctgac	gccatggttt	atacttcaga	cctgctcacc	2040
	aacagtgtca	ttattatggc	atatgtaact	ggtggctctg	tacaacagac	ttctcagtgg	2100
15	ttgtctaata	ttttgggcac	tactgttgaa	aaactcaggc	ctatctttga	atggattgag	2160
	gcgaaactta	gtgcaggagt	tgaatttctc	aaggatgctt	gggagattct	caaatttctc	2220
	attacagggtg	tttttgacat	cgtcaagggt	caaatacagg	ttgcttcaga	taacatcaag	2280
20	gattgtgtaa	aatgcttcat	tgatgttggt	aacaaggcac	tcgaaatgtg	cattgatcaa	2340
	gtcactatcg	ctggcgcaaa	gttgcgatca	ctcaacttag	gtgaagtctt	catcgctcaa	2400
	agcaagggac	tttaccgtca	gtgtatacgt	ggcaaggagc	agctgcaact	actcatgcct	2460
25	cttaaggcac	caaaagaagt	aacctttctt	gaagggtgatt	cacatgacac	agtacttacc	2520
	tctgaggagg	ttgttctcaa	gaacggtgaa	ctcgaagcac	tcgagacgcc	cgttgatagc	2580
	ttcacaatg	gagctatcgt	tggcacacca	gtctgtgtaa	atggcctcat	gctcttagag	2640
30	attaaggaca	aagaacaata	ctgcgcattg	tctcctgggt	tactggctac	aaacaatgtc	2700
	tttcgcttaa	aaggggggtgc	accaattaaa	ggtgtaacct	ttggagaaga	tactgtttgg	2760
	gaagttcaag	gttacaagaa	tgtgagaatc	acatttgagc	ttgatgaacg	tgttgacaaa	2820
35	gtgcttaatg	aaaagtgtc	tgtctacact	gttgaatccg	gtaccgaagt	tactgagttt	2880
	gcatgtgttg	tagcagaggc	tgttgatgag	actttacaac	cagtttctga	tctccttacc	2940
	aacatgggta	ttgatcttga	tgagtggagt	gtagctacat	tctacttatt	tgatgatgct	3000
40	ggtgaagaaa	acttttcatc	acgtatgtat	tgttcctttt	accctccaga	tgaggaagaa	3060
	gaggacgatg	cagagtgtga	ggaagaagaa	attgatgaaa	cctgtgaaca	tgagtacggt	3120
	acagaggatg	attatcaagg	tctccctctg	gaatttggtg	cctcagctga	aacagttcga	3180
45	gttgaggaag	aagaagagga	agactggctg	gatgatacta	ctgagcaatc	agagattgag	3240
	ccagaaccag	aacctacacc	tgaagaacca	gttaatcagt	ttactgggta	tttaaaactt	3300
	actgacaatg	ttgccattaa	atgtgttgac	atcgtaaagg	aggcacaaag	tgctaatacct	3360
50	atggtgattg	taaatgtctg	taacatacac	ctgaaacatg	gtggtggtgt	agcagggtgca	3420
	ctcaacaagg	caaccaatgg	tgccatgcaa	aaggagagtg	atgattacat	taagctaaat	3480
	ggcctcttta	cagtaggagg	gtcttggttg	ctttctggac	ataatcttgc	taagaagtgt	3540
55	ctgcatgttg	ttggacctaa	cctaaatgca	ggtgaggaca	tccagcttct	taaggcagca	3600
	tatgaaaatt	tcaattcaca	ggacatctta	cttgccacct	tggtgtcagc	aggcatattt	3660



EP 1 694 829 B1

	ggtgctaaac	cacttcagtc	tttacaagtg	tgcgtgcaga	cggttcgtac	acagggtttat	3720
	attgcagtca	atgacaaagc	tctttatgag	caggttgtca	tggattatct	tgataacctg	3780
5	aagcctagag	tggaagcacc	taaacaagag	gagccaccaa	acacagaaga	ttccaaaact	3840
	gaggagaaat	ctgtcgtaca	gaagcctgtc	gatgtgaagc	caaaaattaa	ggcctgcatt	3900
	gatgaggtta	ccacaacact	ggaagaaact	aagtttctta	ccaataagtt	actcttgttt	3960
10	gctgatataca	atggtaagct	ttaccatgat	tctcagaaca	tgcttagagg	tgaagatatg	4020
	tctttccttg	agaaggatgc	accttacatg	gtaggtgatg	ttatcactag	tggtgatatc	4080
	acttggttg	taataccctc	caaaaaggct	ggtggcacta	ctgagatgct	ctcaagagct	4140
15	ttgaagaaag	tgccagttga	tgagtatata	accacgtacc	ctggacaagg	atgtgctggt	4200
	tatacacttg	aggaagctaa	gactgctctt	aagaaatgca	aatctgcatt	ttatgtacta	4260
	ccttcagaag	cacctaattgc	taaggaagag	attctaggaa	ctgtatcctg	gaatttgaga	4320
20	gaaatgcttg	ctcatgctga	agagacaaga	aaattaatgc	ctatatgcat	ggatgttaga	4380
	gccataatgg	caaccatcca	acgtaagtat	aaaggaatta	aaattcaaga	gggcatcggt	4440
	gactatggtg	tccgattctt	cttttatact	agtaaagagc	ctgtagcttc	tattattacg	4500
25	aagctgaact	ctctaaatga	gccgcttgct	acaatgccaa	ttggttatgt	gacacatggt	4560
	tttaatcttg	aagaggctgc	gcgctgtatg	cgttctctta	aagctcctgc	cgtagtgtca	4620
	gtatcatcac	cagatgctgt	tactacatat	aatggatacc	tcacttcgtc	atcaaagaca	4680
30	tctgaggagc	actttgtaga	aacagtttct	ttggctggct	cttacagaga	ttggtcctat	4740
	tcaggacagc	gtacagagtt	aggtgttgaa	tttcttaagc	gtggtgacaa	aattgtgtac	4800
	cacactctgg	agagccccgt	cgagtttcat	cttgacgggtg	aggttctttc	acttgacaaa	4860
35	ctaaagagtc	tcttatccct	gcgggagggt	aagactataa	aagtgttcac	aactgtggac	4920
	aacactaatc	tccacacaca	gcttggtgat	atgtctatga	catatggaca	gcagtttggt	4980
	ccaacatact	tggtatggtg	tgatgttaca	aaaattaaac	ctcatgtaaa	tcagtggggt	5040
40	aagactttct	ttgtactacc	tagtgatgac	acactacgta	gtgaagcttt	cgagtactac	5100
	catactcttg	atgagagttt	tcttggtagg	tacatgtctg	ctttaaacca	cacaaagaaa	5160
	tggaattttc	ctcaagttag	tggtttaact	tcaattaaat	gggctgataa	caattgttat	5220
45	ttgtctagt	ttttattagc	acttcaacag	cttgaagtca	aattcaatgc	accagcactt	5280
	caagaggctt	attatagagc	ccgtgctggt	gatgctgcta	acttttgtgc	actcactctc	5340
	gcttacagta	ataaaactgt	tggtgagctt	ggtgatgtca	gagaaactat	gacccatctt	5400
50	ctacagcatg	ctaatttgga	atctgcaaag	cgagttctta	atgtggtgtg	taaacattgt	5460
	ggtcagaaaa	ctactacctt	aacgggtgta	gaagctgtga	tgtatatggg	tactctatct	5520
	tatgataatc	ttaagacagg	tgtttccatt	ccatgtgtgt	gtggctgtga	tgctacacaa	5580
55	tatctagtac	aacaagagtc	ttcttttggt	atgatgtctg	caccacctgc	tgagtataaa	5640
	ttacagcaag	gtacattctt	atgtgcgaat	gagtacactg	gtaactatca	gtgtgggtcat	5700

	tacactcata taactgctaa ggagaccctc tatcgtattg acggagctca ctttacaag	5760
	atgtcagagt acaaaggacc agtgactgat gttttctaca aggaaacatc ttacactaca	5820
5	accatcaagc ctgtgtcgtg taaactcgat ggagttactt acacagagat tgaacaaaa	5880
	ttggatgggt attataaaaa ggataatgct tactatacag agcagcctat agaccttgta	5940
	ccaactcaac cattaccaaa tgcgagtttt gataatttca aactcacatg ttctaacaca	6000
10	aaatttgctg atgattttaa tcaaatgaca ggcttcacaa agccagcttc acgagagcta	6060
	tctgtcacat tcttcccaga ctggaatggc gatgtagtgg ctattgacta tagacactat	6120
	tcagcgagtt tcaagaaagg tgctaaatta ctgcataagc caattgtttg gcacattaac	6180
15	caggctacaa ccaagacaac gttcaaacca aacacttggg gtttacgttg tctttggagt	6240
	acaaagccag tagatacttc aaattcattt gaagttctgg cagtagaaga cacacaagga	6300
	atggacaatc ttgcttgtga aagtcaacaa cccacctctg aagaagtagt ggaaaatcct	6360
20	accatacaga aggaagtcac agagtgtgac gtgaaaacta ccgaagttgt aggcaatgtc	6420
	atacttaaac catcagatga aggtgtttaa gtaacacaag agttagggtca tgaggatctt	6480
	atggctgctt atgtggaaaa cacaagcatt accattaaga aacctaata gctttcacta	6540
25	gccttagggt taaaaacaat tgccactcat ggtattgctg caattaatag tgttccttgg	6600
	agtaaaaattt tggcttatgt caaacattc ttaggacaag cagcaattac aacatcaa	6660
	tgcgctaaga gattagcaca acgtgtgttt aacaattata tgccttatgt gtttacatta	6720
30	ttgttccaat tgtgtacttt tactaaaagt accaattcta gaattagagc ttactacct	6780
	acaactattg ctaaaaatag tgtaagagt gttgctaaat tatgtttgga tgccggcatt	6840
	aattatgtga agtcacccaa attttctaaa ttgttcacaa tcgctatgtg gctattgttg	6900
35	ttaagtattt gcttaggttc tctaactctgt gtaactgctg cttttgggtg actcttatct	6960
	aattttggtg ctcccttcta ttgtaatggc gttagagaat tgtatcttaa ttcgtctaac	7020
	gttactacta tggatttctg tgaaggttct tttccttgca gcatttgttt aagtggatta	7080
40	gactcccttg attcttatcc agctcttgaa accattcagg tgacgatttc atcgtaacag	7140
	ctagacttga caattttagg tctggccgct gagtgggttt tggcatatat gttgttcaca	7200
	aaattctttt atttattagg tctttcagct ataatgcagg tgttcttttg ctattttgct	7260
45	agtcatttca tcagcaattc ttggctcatg tggtttatca ttagtattgt acaaatggca	7320
	cccgtttctg caatgggttag gatgtacatc ttctttgctt ctttctacta catatggaag	7380
	agctatgttc atatcatgga tgggtgcacc tcttcgactt gcatgatgtg ctataagcgc	7440
50	aatcgtgcca cacgcgttga gtgtacaact attgttaatg gcatgaagag atctttctat	7500
	gtctatgcaa atggaggccg tggcttctgc aagactcaca attggaattg tctcaattgt	7560
	gacacatttt gactggtag tacattcatt agtgatgaag ttgctcgtga tttgtcactc	7620
55	cagtttaaaa gaccaatcaa ccctactgac cagtcacgt atattgttga tagtgttgct	7680
	gtgaaaaatg gcgcgcttca cctctacttt gacaaggctg gtcaaaagac ctatgagaga	7740

EP 1 694 829 B1

	catccgctct	cccattttgt	caatttagac	aatttgagag	ctaacaacac	taaaggttca	7800
	ctgcctatta	atgtcatagt	ttttgatggc	aagtccaaat	gcgacgagtc	tgcttctaag	7860
5	tctgcttctg	tgtactacag	tcagctgatg	tgccaaccta	ttctgttgct	tgaccaagct	7920
	cttgatcatg	acgttggaga	tagtactgaa	gtttccgtta	agatgtttga	tgcttatgtc	7980
	gacacctttt	cagcaacttt	tagtgttcct	atggaaaaac	ttaaggcact	tgttgctaca	8040
10	gctcacagcg	agtttagcaa	gggtgtagct	ttagatggtg	tcctttctac	attcgtgtca	8100
	gctgcccgcg	aaggtggttg	tgataccgat	gttgacacaa	aggatgttat	tgaatgtctc	8160
	aaactttcac	atcactctga	cttagaagtg	acaggtgaca	gttgtaacaa	tttcatgctc	8220
15	acctataata	aggttgaaaa	catgacgccc	agagatcttg	gcgcatgtat	tgactgtaat	8280
	gcaaggcata	tcaatgccc	agtagcaaaa	agtcacaatg	tttcactcat	ctggaatgta	8340
	aaagactaca	tgtctttatc	tgaacagctg	cgtaaacaaa	ttcgtagtgc	tgccaagaag	8400
20	aacaacatac	cttttagact	aacttgtgct	acaactagac	aggttgtcaa	tgtcataact	8460
	actaaaatct	cactcaaggg	tggttaagatt	gttagtactt	gttttaaact	tatgcttaag	8520
	gccacattat	tgtgcgttct	tgctgcattg	gtttgttata	tcgttatgcc	agtacataca	8580
25	ttgtcaatcc	atgatggtta	cacaaatgaa	atcattgggt	acaaagccat	tcaggatggg	8640
	gtcactcgtg	acatcatttc	tactgatgat	tgttttgcaa	ataaacatgc	tggttttgac	8700
	gcatggttta	gccagcgtgg	tggttcatac	aaaaatgaca	aaagctgccc	tgtagtagct	8760
30	gctatcatta	caagagagat	tggtttcata	gtgcctggct	taccgggtac	tgtgctgaga	8820
	gcaatcaatg	gtgactttct	gcattttcta	cctcgtgttt	ttagtgctgt	tggcaacatt	8880
	tgctacacac	cttccaaact	cattgagtat	agtgattttg	ctacctctgc	ttgcgttctt	8940
35	gctgctgagt	gtacaatttt	taaggatgct	atgggcaaac	ctgtgccata	ttgttatgac	9000
	actaatattg	tagaggggtt	tattttctat	agtgagcttc	gtccagacac	tcgttatgtg	9060
	cttatggatg	gttccatcat	acagtttcct	aacacttacc	tggaggggtt	tgttagagta	9120
40	gtaacaactt	ttgatgctga	gtactgtaga	catggtacat	gcgaaaggct	agaagtaggt	9180
	atgtgcctat	ctaccagtgg	tagatgggtt	cttaataatg	agcattacag	agctctatca	9240
	ggagttttct	gtggtgttga	tgcgatgaat	ctcatagcta	acatctttac	tcctcttggt	9300
45	caacctgtgg	gtgctttaga	tgtgtctgct	tcagtagtgg	ctggtgggtat	tattgccata	9360
	ttggtgactt	gtgctgccta	ctactttatg	aaattcagac	gtgtttttgg	tgagtacaac	9420
	catgttggtg	ctgctaattg	acttttggtt	ttgatgtctt	tcactatact	ctgtctggta	9480
50	ccagcttaca	gctttctgcc	gggagtctac	tcagtctttt	acttgacttt	gacattctat	9540
	ttcaccaatg	atgtttcatt	cttggctcac	cttcaatggg	ttgccatggt	ttctcttatt	9600
	gtgccttttt	ggataacagc	aatctatgta	ttctgtatgt	ctctgaagca	ctgccattgg	9660
55	ttctttaaca	actatcttag	gaaaagagtc	atgtttaatg	gagttacatt	tagtaccttc	9720
	gaggaggctg	ctttgtgtac	ctttttgctc	aacaaggaaa	tgtacctaaa	attgcgtagc	9780

## EP 1 694 829 B1

	gagacactgt	tgccacttac	acagtataac	aggtatcttg	ctctatataa	caagtacaag	9840
	tatttcagtg	gagccttaga	tactaccagc	tatcgtgaag	cagcttgctg	ccacttagca	9900
5	aaggctctaa	atgacttttag	caactcaggt	gctgatgttc	tctaccaacc	accacagaca	9960
	tcaatcactt	ctgctgttct	gcagagtggg	tttaggaaaa	tggcattccc	gtcaggcaaa	10020
	gttgaagggg	gcatggtaca	agtaacctgt	ggaactacaa	ctcttaatgg	attgtggttg	10080
10	gatgacacag	tatactgtcc	aagacatgtc	atttgcacag	cagaagacat	gcttaatcct	10140
	aactatgaag	atctgctcat	tcgcaaatac	aaccatagct	ttcttggtca	ggctggcaat	10200
	gttcaacttc	gtgtttattg	ccattctatg	caaaattgtc	tgcttaggct	taaagttgat	10260
15	acttctaacc	ctaagacacc	caagtataaa	tttgtccgta	tccaacctgg	tcaaacattt	10320
	tcagttctag	catgctacaa	tggttcacca	tctgggtgtt	atcagtgtgc	catgagacct	10380
	aatcatacca	ttaaagggtt	tttccttaat	ggatcatgtg	gtagtggttg	ttttaacatt	10440
20	gattatgatt	gcgtgtcttt	ctgctatatg	catcatatgg	agcttccaac	aggagtacac	10500
	gctggtactg	acttagaagg	taaattctat	gggccatttg	ttgacagaca	aactgcacag	10560
	gctgcaggta	cagacacaac	cataacatta	aatgttttgg	catggctgta	tgctgctgtt	10620
25	atcaatgggtg	ataggtgggt	tcttaataga	ttcaccacta	ctttgaatga	ctttaacctt	10680
	gtggcaatga	agtacaacta	tgaacctttg	acacaagatc	atgttgacat	attgggacct	10740
	ctttctgctc	aaacaggaat	tgccgtctta	gatatgtgtg	ctgctttgaa	agagctgctg	10800
30	cagaatggta	tgaatggctg	tactatcctt	ggtagcacta	ttttagaaga	tgagtttaca	10860
	ccatttgatg	ttgttagaca	atgctctggg	gttaccttcc	aaggtaagtt	caagaaaatt	10920
	gttaagggca	ctcatcattg	gatgctttta	actttcttga	catcactatt	gattcttggt	10980
35	caaagtacac	agtggctact	gtttttcttt	gtttacgaga	atgctttctt	gccatttact	11040
	cttgggtatta	tggcaattgc	tgcatgtgct	atgctgcttg	ttaagcataa	gcacgcattc	11100
	ttgtgcttgt	ttctgttacc	ttctcttgca	acagttgctt	actttaatat	gggtctacatg	11160
40	cctgctagct	gggtgatgcg	tatcatgaca	tggcttgaat	tggctgacac	tagcttgtct	11220
	ggttataggc	ttaaggattg	tgttatgtat	gcttcagctt	tagttttgct	tattctcatg	11280
	acagctcgca	ctgtttatga	tgatgctgct	agacgtgttt	ggacactgat	gaatgtcatt	11340
45	acacttgttt	acaaagtcta	ctatggtaat	gcttttagatc	aagctatttc	catgtgggccc	11400
	ttagttattt	ctgtaacctc	taactattct	gggtgcgtta	cgactatcat	gttttttagct	11460
	agagctatag	tgtttggtg	tggtgagtat	taccattgt	tatttattac	tggcaacacc	11520
50	ttacagtgta	tcatgcttgt	ttattgtttc	ttaggctatt	gttgctgctg	ctactttggc	11580
	cttttctggt	tactcaaccg	ttacttcagg	cttactcttg	gtgtttatga	ctacttggtc	11640
	tctacacaag	aatttaggta	tatgaactcc	caggggcttt	tgccctcctaa	gagtagtatt	11700
55	gatgctttca	agcttaacat	taagttgttg	ggtattggag	gtaaaccatg	tatcaagggt	11760
	gctactgtac	agtctaaaat	gtctgacgta	aagtgcacat	ctgtgggtact	gctctcggtt	11820

## EP 1 694 829 B1

:ttcaacaac ttagagtaga gtcattcttct aaattgtggg cacaatgtgt acaactccac 11880  
 aatgatattc ttcttgcaaa agacacaact gaagctttcg agaagatggt ttctcttttg 11940  
 5 tctgttttgc tatccatgca ggggtgctgta gacattaata ggttgtgcga ggaaatgctc 12000  
 gataaccgtg ctactcttca ggctattgct tcagaattta gttctttacc atcatatgcc 12060  
 gcttatgcca ctgcccagga ggcctatgag caggctgtag ctaatggtga ttctgaagtc 12120  
 10 gttctcaaaa agttaaagaa atctttgaat gtggctaaat ctgagtttga ccgtgatgct 12180  
 gccatgcaac gcaagttgga aaagatggca gatcaggcta tgacccaaat gtacaaacag 12240  
 gcaagatctg aggacaagag ggcaaaagta actagtgcga tgcaacaat gctcttcact 12300  
 15 atgcttagga agcttgataa tgatgcactt aacaacatta tcaacaatgc gcgtgatggt 12360  
 tgtgttccac tcaacatcat accattgact acagcagcca aactcatggt tgttgtccct 12420  
 gattatggta cctacaagaa cacttgtgat ggtaacacct ttacatatgc atctgcactc 12480  
 20 tgggaaatcc agcaagttgt tgatgcggat agcaagattg ttcaacttag tgaaattaac 12540  
 atggacaatt caccaaattt ggcttggcct cttattgtta cagctctaag agccaactca 12600  
 gctgttaaac tacagaataa tgaactgagt ccagtagcac tacgacagat gtcctgtgcg 12660  
 25 gctggtacca cacaaacagc ttgtactgat gacaatgcac ttgcctacta taacaattcg 12720  
 aaggagggta ggtttgtgct ggcattacta tcagaccacc aagatctcaa atgggctaga 12780  
 ttccctaaga gtgatggtac aggtacaatt tacacagaac tggaaccacc ttgtaggttt 12840  
 30 gttacagaca cacaaaagg gcctaaagt aaatacttgt acttcatcaa aggcctaaac 12900  
 aacctaata gaggtatggt gctgggcagt ttagctgcta cagtacgtct tcaggctgga 12960  
 aatgctacag aagtacctgc caattcaact gtgctttcct tctgtgcttt tgcagtagac 13020  
 35 cctgctaaag catataagga ttacctagca agtggaggac aaccaatcac caactgtgtg 13080  
 aagatgttgt gtacacacac tggtagagga caggcaatta ctgtaacacc agaagctaac 13140  
 atggaccaag agtcctttgg tggtagctca tgttgtctgt attgtagatg ccacattgac 13200  
 catccaaatc ctaaaggatt ctgtgacttg aaaggtaagt acgtccaaat acctaccact 13260  
 40 tgtgctaata acccagtggg ttttacctt agaaacacag tctgtaccgt ctgcggaatg 13320  
 tggaaagggt atggctgtag ttgtgaccaa ctccgcgaac ccttgatgca gtctgcggat 13380  
 gcatcaacgt ttttaaacgg gtttgcggtg taagtgcagc ccgtcttaca ccgtgcggca 13440  
 45 caggcactag tactgatgtc gtctacaggg cttttgatat ttacaacgaa aaagttgctg 13500  
 gttttgcaaa gttcctaaaa actaattgct gtcgcttcca ggagaaggat gaggaaggca 13560  
 atttattaga ctcttacttt gtagttaaga ggcatactat gtctaactac caacatgaag 13620  
 50 agactattta taacttgggt aaagattgtc cagcgggtgc tgtccatgac tttttcaagt 13680  
 ttagagtaga tggtagacat gtaccacata tatcacgtca gcgtctaact aaatacacia 13740  
 tggctgattt agtctatgct ctacgtcatt ttgatgaggg taattgtgat acattaaaag 13800  
 55 aaatactcgt cacatacaat tgctgtgatg atgattattt caataagaag gattgggatg 13860

acttcgtaga gaatcctgac atcttacgcg tatatgctaa cttaggtgag cgtgtacgcc 13920  
 aatcattatt aaagactgta caattctgcg atgctatgcg tgatgcaggc attgtaggcg 13980  
 5 tactgacatt agataatcag gatcttaatg ggaactggta cgatttcggt gatttcgtac 14040  
 aagtagcacc aggctgcgga gttcctattg tggattcata ttactcattg ctgatgccca 14100  
 tcctcacttt gactagggca ttggctgctg agtcccatat ggatgctgat ctcgcaaaac 14160  
 10 cacttattaa gtgggatttg ctgaaatatg attttacgga agagagactt tgtctcttcg 14220  
 accgttatatt taaatattgg gaccagacat accatcccaa ttgtattaac tgtttggatg 14280  
 ataggtgtat ccttcattgt gcaaaacttta atgtgttatt ttctactgtg tttccaccta 14340  
 15 caagttttgg accactagta agaaaaatat ttgtagatgg tgttcctttt gttgtttcaa 14400  
 ctggatacca ttttcgtgag ttaggagtcg tacataatca ggatgtaaac ttacatagct 14460  
 cgcgctctcag tttcaaggaa ctttttagtgt atgctgctga tccagctatg catgcagctt 14520  
 20 ctggcaattt attgctagat aaacgcacta catgcttttc agtagctgca ctaacaaaca 14580  
 atgttgcttt tcaaactgtc aaacccggta attttaataa agacttttat gactttgctg 14640  
 tgtctaaagg tttctttaag gaaggaagtt ctgttgaact aaaacacttc tcttttgctc 14700  
 25 aggatggcaa cgctgctatc agtgattatg actattatcg ttataatctg ccaacaatgt 14760  
 gtgatatcag acaactccta ttcgtagttg aagttgttga taaatacttt gattgttacg 14820  
 atggtggctg tattaatgcc aaccaagtaa tcgttaacaa tctggataaa tcagctgggt 14880  
 30 tcccatttaa taaatggggg aaggctagac tttattatga ctcaatgagt tatgaggatc 14940  
 aagatgcact ttcgcgctat actaagcgta atgtcatccc tactataact caaatgaatc 15000  
 ttaagtatgc cattagtgc aagaatagag ctcgcaccgt agctgggtgc tctatctgta 15060  
 35 gtactatgac aatagacag tttcatcaga aattattgaa gtcaatagcc gccactagag 15120  
 gagctactgt ggtaattgga acaagcaagt tttacgggtg ctggcataat atgttaaaaa 15180  
 ctgtttacag tgatgtagaa actccacacc ttatgggttg ggattatcca aaatgtgaca 15240  
 gagccatgcc taacatgctt aggataatgg cctctcttgt tcttgctcgc aaacataaca 15300  
 40 cttgctgtaa cttatcacac cgtttctaca ggtagctaa cgagtgtgcg caagtattaa 15360  
 gtgagatggg catgtgtggc ggctcactat atgttaaacc aggtggaaca tcatccggtg 15420  
 atgctacaac tgcttatgct aatagtgtct ttaacatttg tcaagctgtt acagccaatg 15480  
 45 taaatgcact tctttcaact gatggtaata agatagctga caagtatgtc cgcaatctac 15540  
 aacacaggct ctatgagtgt ctctatagaa atagggatgt tgatcatgaa ttcgtggatg 15600  
 agttttacgc ttacctgct aaacatttct ccatgatgat tctttctgat gatgccgttg 15660  
 50 tgtgctataa cagtaactat gcggctcaag gtttagtagc tagcattaag aactttaagg 15720  
 cagttcttta ttatcaaaat aatgtgttca tgtctgaggc aaaatgttgg actgagactg 15780  
 accttactaa aggacctcac gaattttgct cacagcatac aatgctagtt aaacaaggag 15840  
 55 atgattacgt gtacctgcct taccagatc catcaagaat attaggcgca ggctgttttg 15900

EP 1 694 829 B1

	tcgatgatat	tgtcaaaaca	gatggtacac	ttatgattga	aagggttcgtg	tcactgggcta	15960
	ttgatgctta	cccacttaca	aaacatccta	atcaggagta	tgctgatgtc	tttcacttgt	16020
5	atttacaata	cattagaaag	ttacatgatg	agcttactgg	ccacatggtg	gacatgtatt	16080
	ccgtaatgct	aactaatgat	aacacctcac	ggtactggga	acctgagttt	tatgaggcta	16140
	tgtacacacc	acatacagtc	ttgcaggctg	taggtgcttg	tgtattgtgc	aattcacaga	16200
10	cttcacttcg	ttgcggtgcc	tgtattagga	gaccattcct	atgttgcaag	tgctgctatg	16260
	accatgtcat	ttcaacatca	cacaaattag	tgttgtctgt	taatccctat	gtttgcaatg	16320
	ccccaggttg	tgatgtcact	gatgtgacac	aactgtatct	aggaggatatg	agctattatt	16380
15	gcaagtcaca	taagcctccc	attagttttc	cattatgtgc	taatggtcag	gttttttggtt	16440
	tatacaaaaa	cacatgtgta	ggcagtgaca	atgtcactga	cttcaatgcg	atagcaacat	16500
	gtgattggac	taatgctggc	gattacatac	ttgccaacac	ttgtactgag	agactcaagc	16560
20	ttttcgcagc	agaaacgctc	aaagccactg	aggaaacatt	taagctgtca	tatgggtattg	16620
	ccactgtacg	cgaagtactc	tctgacagag	aattgcatct	ttcatgggag	gttgaaaaac	16680
	ctagaccacc	attgaacaga	aactatgtct	ttactgggta	ccgtgtaact	aaaaatagta	16740
25	aagtacagat	tggagagtac	acctttgaaa	aaggtgacta	tgggtgatgct	gttgtgtaca	16800
	gagggtactac	gacatacaag	ttgaatggtg	gtgattactt	tgtgttgaca	tctcacactg	16860
	taatgccact	tagtgcacct	actctagtgc	cacaagagca	ctatgtgaga	attactggct	16920
30	tgtaccaaac	actcaacatc	tcagatgagt	tttctagcaa	tggtgcaa	tatcaaaagg	16980
	tcggcatgca	aaagtactct	acactccaag	gaccacctgg	tactggtaag	agtcattttg	17040
	ccatcggact	tgctctctat	tacccatctg	ctcgcatagt	gtatacggca	tgctctcatg	17100
35	cagctgttga	tgccctatgt	gaaaaggcat	taaaatattt	gcccatagat	aaatgtagta	17160
	gaatcatacc	tgcgcggtgcg	cgcgtagagt	gttttgataa	attcaaagtg	aattcaacac	17220
	tagaacagta	tgttttctgc	actgtaaatg	cattgccaga	aacaactgct	gacattgtag	17280
40	tctttgatga	aatctctatg	gctactaatt	atgacttgag	tggtgtcaat	gctagacttc	17340
	gtgcaaaaca	ctacgtctat	attggcgatc	ctgctcaatt	accagcccc	cgcacattgc	17400
	tgactaaagg	cacactagaa	ccagaatatt	ttaattcagt	gtgcagactt	atgaaaacaa	17460
45	taggtccaga	catgttcctt	ggaacttgtc	gccgttgctc	tgctgaaatt	gttgacactg	17520
	tgagtgcctt	agtttatgac	aataagctaa	aagcacacaa	ggataagtca	gctcaatgct	17580
	tcaaaatggt	ctacaaagg	gttattacac	atgatgttct	atctgcaatc	aacagacctc	17640
50	aaataggcgt	tgtaagagaa	tttcttacac	gcaatcctgc	ttggagaaaa	gctgttttta	17700
	tctcacctta	taattcacag	aacgctgtag	cttcaaaaat	cttaggattg	cctacgcaga	17760
	ctgttgattc	atcacaggg	tctgaatatg	actatgtcat	attcacacaa	actactgaaa	17820
	cagcacactc	ttgtaatgtc	aaccgcttca	atgtggctat	cacaagggca	aaaattggca	17880
55	ttttgtgcat	aatgtctgat	agagatcttt	atgacaaaact	gcaatttaca	agtctagaaa	17940

EP 1 694 829 B1

	taccacgtcg	caatgtggct	acattacaag	cagaaaatgt	aactggactt	tttaaggact	18000
	gtagtaagat	cattactggg	cttcaccccta	cacaggcacc	tacacacctc	agcgttgata	18060
5	taaagttcaa	gactgaagga	ttatgtgttg	acataccagg	cataccaaag	gacatgacct	18120
	accgtagact	catctctatg	atggggtttca	aaatgaatta	ccaagtcaat	ggttacccta	18180
	atatgtttat	cacccgcgaa	gaagctattc	gtcacgttcg	tgcgtggatt	ggctttgatg	18240
10	tagagggctg	tcatgcaact	agagatgctg	tgggtactaa	cctacctctc	cagctaggat	18300
	tttctacagg	tggttaactta	gtagctgtac	cgactgggta	tggtgacact	gaaaataaca	18360
	cagaattcac	cagagttaat	gcaaaacctc	caccagggtga	ccagtttaaa	catcttatac	18420
15	cactcatgta	taaaggcttg	ccctggaatg	tagtgcgat	taagatagta	caaagtctca	18480
	gtgatacact	gaaaggattg	tcagacagag	tcgtgttcgt	cctttgggcg	catggccttg	18540
	agcttacatc	aatgaagtac	tttgtcaaga	ttggacctga	aagaacgtgt	tgtctgtgtg	18600
20	acaaacgtgc	aacttgcttt	tctacttcat	cagatactta	tgctgtctgg	aatcattctg	18660
	tgggttttga	ctatgtctat	aaccatttta	tgattgatgt	tcagcagtgg	ggctttacgg	18720
	gtaaccttca	gagtaaccat	gaccaacatt	gccagggtaca	tggaaatgca	catgtggcta	18780
25	gttgtgatgc	tatcatgact	agatgtttag	cagtccatga	gtgctttgtt	aagcgcgttg	18840
	attgggtctgt	tgaataccct	attataggag	atgaactgag	ggttaattct	gcttgacaga	18900
	aagtacaaca	catgggtgtg	aagtctgcat	tgcttgctga	taagtttcca	gttcttcatg	18960
30	acattggaaa	tccaaaggct	atcaagtgtg	tgcttcaggc	tgaagtagaa	tggaagttct	19020
	acgatgtctca	gccatgtagt	gacaaagctt	acaaaataga	ggaactcttc	tattcttatg	19080
	ctacacatca	cgataaattc	actgatgggtg	tttgtttgtt	ttggaattgt	aacgttgatc	19140
35	gttaccacgc	caatgcaatt	gtgtgtaggt	ttgacacaag	agtcttgtca	aacttgaact	19200
	taccaggctg	tgatgggtgg	agtttgtatg	tgaataagca	tgcatccac	actccagctt	19260
	tcgataaaag	tgcatctact	aatttaaagc	aattgccttt	cttttactat	tctgatagtc	19320
40	cttgtgagtc	tcatggcaaa	caagtagtgt	cggatattga	ttatgttcca	ctcaaactctg	19380
	ctacgtgtat	tacacgatgc	aatttaggtg	gtgctgtttg	cagacaccat	gcaaagtagt	19440
	accgacagta	cttgatgca	tataatatga	tgatttctgc	tggaattagc	ctatggattt	19500
45	acaaacaatt	tgatacttat	aacctgtgga	atacatttac	caggttacag	agtttagaaa	19560
	atgtggctta	taatgttgtt	aataaaggac	actttgatgg	acacgccggc	gaagcacctg	19620
	tttccatcat	taataatgct	gtttacacaa	aggtagatgg	tattgatgtg	gagatctttg	19680
50	aaaataagac	aacacttcct	gttaatgttg	catttgagct	ttgggctaag	cgtaacatta	19740
	aaccagtgcc	agagattaag	atactcaata	atttgggtgt	tgatatcgct	gctaatactg	19800
	taatctggga	ctacaaaaga	gaagccccag	cacatgtatc	tacaataggt	gtctgcacaa	19860
55	tgactgacat	tgccaagaaa	cctactgaga	gtgcttggtc	ttcacttact	gtcttggttg	19920
	atggtagagt	ggaaggacag	gtagaccttt	ttagaaacgc	ccgtaatggg	gttttaataa	19980



EP 1 694 829 B1

	cagaagggttc	agtcaaaggt	ctaacacctt	caaagggacc	agcacaagct	agcgtcaatg	20040
	gagtcacatt	aattggagaa	tcagtaaaaa	cacagtttaa	ctactttaag	aaagtagacg	20100
5	gcattattca	acagttgcct	gaaacctact	ttactcagag	cagagactta	gaggatttta	20160
	agcccagatc	acaaatggaa	actgactttc	tcgagctcgc	tatggatgaa	ttcatacagc	20220
	gatataagct	cgagggttat	gccttcgaac	acatcgttta	tggagatttc	agtcattggac	20280
10	aacttggcgg	tcttcattta	atgataggct	tagccaagcg	ctcacaagat	tcaccactta	20340
	aattagagga	ttttatccct	atggacagca	cagtgaaaaa	ttacttcata	acagatgcgc	20400
	aaacagggttc	atcaaaatgt	gtgtgttctg	tgattgatct	tttacttgat	gactttgtcg	20460
15	agataataaa	gtcacaagat	ttgtcagtga	tttcaaaagt	ggtcaagggt	acaattgact	20520
	atgctgaaat	ttcattcatg	ctttgggtga	aggatggaca	tgttgaaacc	ttctacccaa	20580
	aactacaagc	aagtcaagcg	tggcaaccag	gtgttgcgat	gcctaacttg	tacaagatgc	20640
20	aaagaatgct	tcttgaaaag	tgtgaccttc	agaattatgg	tgaaaatgct	gttataccaa	20700
	aaggaataat	gatgaatgtc	gcaaagtata	ctcaactgtg	tcaatactta	aatacactta	20760
	ctttagctgt	accctacaac	atgagagtta	ttcacttttg	tgctggctct	gataaaggag	20820
25	ttgcaccagg	tacagctgtg	ctcagacaat	ggttgccaac	tggcacacta	cttgtcgtat	20880
	cagatcttaa	tgacttcgtc	tccgacgcag	attctacttt	aattggagac	tgtgcaacag	20940
	tacatacggc	taataaatgg	gaccttatta	ttagcgatat	gtatgaccct	aggaccaaac	21000
30	atgtgacaaa	agagaatgac	tctaaagaag	ggtttttcac	ttatctgtgt	ggattttataa	21060
	agcaaaaact	agccctgggt	ggttctatag	ctgtaaagat	aacagagcat	tcttggaatg	21120
	ctgaccttta	caagcttatg	ggccatttct	catggtggac	agcttttggt	acaaatgtaa	21180
35	atgcatcatc	atcggaagca	tttttaattg	gggctaacta	tcttggcaag	ccgaaggaac	21240
	aaattgatgg	ctataccatg	catgctaact	acattttctg	gaggaacaca	aatcctatcc	21300
	agttgtcttc	ctattcactc	tttgacatga	gcaaatttcc	tcttaaatta	agaggaactg	21360
40	ctgtaatgtc	tcttaaggag	aatcaaatca	atgatatgat	ttattctctt	ctggaaaaag	21420
	gtaggcttat	cattagagaa	aacaacagag	ttgtggtttc	aagtgatatt	cttgттаaca	21480
	actaaacgaa	catgtttatt	ttcttattat	ttcttactct	cactagtggg	agtgaccttg	21540
45	accggtgcac	cacttttgat	gatgttcaag	ctcctaatta	cactcaacat	acttcatcta	21600
	tgaggggggt	ttactatcct	gatgaaatth	ttagatcaga	cactctttat	ttaactcagg	21660
	atttatttct	tccattttat	tctaattgta	caggggtttca	tactattaat	catacgtttg	21720
50	gcaaccctgt	catacctttt	aaggatggta	tttattttgc	tgccacagag	aaatcaaattg	21780
	ttgtccgtgg	ttgggttttt	ggttctacca	tgaacaacaa	gtcacagtcg	gtgattatta	21840
	ttaacaattc	tactaatgtt	gttatacagag	catgtaactt	tgaattgtgt	gacaaccctt	21900
	tctttgctgt	ttctaaaccc	atgggtacac	agacacatac	tatgatattc	gataatgcat	21960
55	ttaattgcac	tttcgagtac	atatctgatg	ccttttcgct	tgatgtttca	gaaaagtcag	22020

## EP 1 694 829 B1

	gtaattttaa	acacttacga	gagtttgtgt	ttaaaaataa	agatggggtt	ctctatgttt	22080
	ataagggcta	tcaacctata	gatgtagttc	gtgatctacc	ttctgggttt	aacactttga	22140
5	aacctatfff	taagttgcct	cttggtatta	acattacaaa	ttttagagcc	attcttacag	22200
	ccttttcacc	tgctcaagac	atgtggggca	cgctcagctgc	agcctatfff	gttggtctatt	22260
	taaagccaac	tacatttatg	ctcaagtatg	atgaaaatgg	tacaatcaca	gatgctgttg	22320
10	attgtttctca	aaatccactt	gctgaactca	aatgctctgt	taagagcttt	gagattgaca	22380
	aaggaattta	ccagacctct	aatttcaggg	ttgttccctc	aggagatgtt	gtgagattcc	22440
	ctaataattac	aaacttgtgt	ccttttggag	aggtttttaa	tgctactaaa	ttcccttctg	22500
15	tctatgcatg	ggagagaaaa	aaaatttcta	attgtgttgc	tgattactct	gtgctctaca	22560
	actcaacatt	tttttcaacc	tttaagtgtc	atggcggttc	tgccactaag	ttgaatgatc	22620
	tttgcttctc	caatgtctat	gcagattctt	ttgtagtcaa	gggagatgat	gtaagacaaa	22680
20	tagcgccagg	acaaactggg	gttattgtctg	attataatta	taaattgcca	gatgatttca	22740
	tgggttgtgt	ccttgcttgg	aatactagga	acattgatgc	tacttcaact	ggtaattata	22800
	attataaata	taggtatctt	agacatggca	agcttaggcc	ctttgagaga	gacatatcta	22860
25	atgtgccttt	ctcccctgat	ggcaaacctt	gcaccccacc	tgctcttaat	tgttattggc	22920
	cattaaatga	ttatggtttt	tacaccacta	ctggcattgg	ctaccaacct	tacagagttg	22980
	tagtactttc	ttttgaactt	ttaaatgcac	cggccacggg	ttgtggacca	aaattatcca	23040
30	ctgaccttat	taagaaccag	tgtgtcaatt	ttaattttaa	tggactcact	ggtagtggtg	23100
	tgttaaactc	ttcttcaaa	agatttcaac	catttcaaca	atgtggccgt	gatgtttctg	23160
	atttactga	ttccgttcga	gattcctaaa	catctgaaat	attagacatt	tcaccttgct	23220
35	cttttggggg	tgtaagtgtg	attacacctg	gaacaaatgc	ttcatctgaa	gttgctgttc	23280
	tatatcaaga	tgttaaactg	actgatgttt	ctacagcaat	tcatgcagat	caactcacac	23340
	cagcttggcg	catatattct	actggaaaca	atgtattcca	gactcaagca	ggctgtctta	23400
40	taggagctga	gcatgtcgac	acttcttatg	agtgcgacat	tcctattgga	gctggcattt	23460
	gtgctagtta	ccatacagtt	tctttattac	gtagtactag	ccaaaaatct	attgtggctt	23520
	atactatgtc	tttaggtgct	gatagttcaa	ttgcttactc	taataacacc	attgctatac	23580
45	ctactaactt	ttcaattagc	attactacag	aagtaatgcc	tgtttctatg	gctaaaacct	23640
	ccgtagattg	taatattgtac	atctgcggag	attctactga	atgtgcta	ttgcttctcc	23700
	aatatggtag	cttttgcaca	caactaaatc	gtgcactctc	aggtattgct	gctgaacagg	23760
50	atcgcaacac	acgtgaagtg	ttcgctcaag	tcaacaaat	gtacaaaacc	ccaactttga	23820
	aatattttgg	tggtttttaa	ttttcacaaa	tattacctga	ccctctaaag	ccaactaaga	23880
	ggctttttat	tgaggacttg	ctctttaata	aggtgacact	cgctgatgct	ggcttcatga	23940
	agcaatatgg	cgaatgccta	ggatgatatta	atgctagaga	tctcatttgt	gcgcagaagt	24000
55	tcaatggact	tacagtgttg	ccacctctgc	tactgatga	tatgattgct	gcctacactg	24060

EP 1 694 829 B1

	ctgctctagt tagtgggtact gccactgctg gatggacatt tgggtgctggc gctgctcttc	24120
	aaataaccttt tgctatgcaa atggcatata ggttcaatgg cattggagtt acccaaaatg	24180
5	ttctctatga gaacccaaaa caaatcgcca accaatttaa caaggcgatt agtcaaattc	24240
	aagaatcact tacaacaaca tcaactgcat tgggcaagct gcaagacgtt gttaaccaga	24300
	atgctcaagc attaaacaca cttgttaaac aacttagctc taattttggt gcaatttcaa	24360
10	gtgtgctaaa tgatatacctt tcgcgacttg ataaagtcga ggcggaggta caaattgaca	24420
	ggtttaattac aggcagactt caaagccttc aaacctatgt aacacaacaa ctaatcaggg	24480
	ctgctgaaat cagggcttct gctaatacctt ctgctactaa aatgtctgag tgtgttcttg	24540
15	gacaatcaaa aagagttgac ttttgtggaa agggctacca ccttatgtcc ttcccacaag	24600
	cagccccgca tgggtgttgct ttcctacatg tcacgtatgt gccatcccag gagaggaact	24660
	tcaccacagc gccagcaatt tgtcatgaag gcaaagcata cttccctcgt gaagggtgttt	24720
20	ttgtgtttta tggcacttct tggtttatta cacagaggaa cttcttttct ccacaaataa	24780
	ttactacaga caatacattt gtctcaggaa attgtgatgt cgttattggc atcattaaca	24840
	acacagttta tgatcctctg caacctgagc ttgactcatt caaagaagag ctggacaagt	24900
25	acttcaaaaa tcatacatca ccagatgttg atcttggcga catttcaggc attaacgctt	24960
	ctgtcgtcaa cattcaaaaa gaaattgacc gcctcaatga ggtcgctaaa aatttaaattg	25020
	aatcactcat tgaccttcaa gaattgggaa aatatgagca atatattaaa tggccttggt	25080
30	atgtttggct cggcttcatt gctggactaa ttgccatcgt catggttaca atcttgcttt	25140
	gttgcatgac tagttgttgc agttgcctca aggggtgatg ctcttggtgt tcttgctgca	25200
	agtttgatga ggatgactct gagccagttc tcaaggggtg caaattacat tacacataaa	25260
35	cgaacttatg gatttgttta tgagattttt tactcttgga tcaattactg cacagccagt	25320
	aaaaattgac aatgcttctc ctgcaagtac tgttcatgct acagcaacga taccgctaca	25380
	agcctcactc ctttctggat ggcttggtat tggcgttgca tttcttgctg ttttctagag	25440
40	cgctaccaa ataattgcgc tcaataaaaag atggcagcta gccctttata agggcttcca	25500
	gttcatttgc aatttactgc tgctatttgt taccatctat tcacatcttt tgcttgctgc	25560
	tgcaggtatg gaggcgcaat ttttgtacct ctatgccttg atatattttc tacaatgcat	25620
45	caacgcatgt agaattatta tgagatgttg gctttgttgg aagtgcaaat ccaagaacct	25680
	attactttat gatgccaaact actttgtttg ctggcacaca cataactatg actactgtat	25740
	accatataac agtgtcacag atacaattgt cgttactgaa ggtgacggca tttcaacacc	25800
50	aaaactcaaa gaagactacc aaattgggtg ttattctgag gataggcact caggtgttaa	25860
	agactatgtc gttgtacatg gctatttcac cgaagtttac taccagcttg agtctacaca	25920
	aattactaca gacactggta ttgaaaatgc tacattcttc atctttaaca agcttggtta	25980
55	agaccaccg aatgtgcaaa tacacacaat cgacggctct tcaggagttg ctaatccagc	26040
	aatggatcca atttatgatg agccgacgac gactactagc gtgcctttgt aagcacaaga	26100

EP 1 694 829 B1

aagtgagtag gaacttatgt actcattcgt ttcggaagaa acaggtagct taatagttaa 26160  
 tagcgtagctt ctttttcttg ctttcgtggg attccttgcta gtcacactag ccatccttac 26220  
 5 tgcgcttcga ttgtgtgcgt actgctgcaa tattgttaac gtgagtttag taaaaccaac 26280  
 ggtttacgtc tactcgcggt ttaaaaatct gaactcttct gaaggagttc ctgatcttct 26340  
 ggtctaaacg aactaactat tattattatt ctgtttggaa ctttaacatt gcttatcatg 26400  
 10 gcagacaacg gtactattac cggtgaggag cttaaacaac tcctggaaca atggaaccta 26460  
 gtaatagggt tcctattcct agcctggatt atgttactac aatttgccca ttctaactcg 26520  
 aacagggtttt tgtacataat aaagcttggt ttcctctggc tcttggtggc agtaacactt 26580  
 15 gcttggtttt tgcttgctgc tgtctacaga attaatggg tgactggcgg gattgcgatt 26640  
 gcaatggctt gtattgtagg cttgatgtgg cttagctact tcgttgcttc cttcaggctg 26700  
 ttgtctcgta cccgctcaat gtggctattc aaccagaaa caaacattct tctcaatgtg 26760  
 20 cctctccggg ggacaattgt gaccagaccg ctcatggaaa gtgaacttgt cattggtgct 26820  
 gtgatcattc gtggtcactt gcgaatggcc ggacactccc tagggcgctg tgacattaag 26880  
 gacctgccaa aagagatcac tgtggctaca tcacgaacgc tttcttatta caaattagga 26940  
 25 gcgtcgcagc gtgtaggcac tgattcagggt ttgctgcat acaaccgcta ccgtattgga 27000  
 aactataaat taaatacaga ccacgccggg agcaacgaca atattgcttt gctagtagac 27060  
 taagtgacaa cagatgtttc atcttggtga cttccagggt acaatagcag agatattgat 27120  
 30 tatcattatg aggactttca ggattgctat ttggaatctt gacgttataa taagttcaat 27180  
 agtgagacaa ttattttaagc ctctaactaa gaagaattat tcggagttag atgatgaaga 27240  
 acctatggag ttagattatc cataaaacga acatgaaaat tattctcttc ctgacattga 27300  
 35 ttgtattttac atcttgcgag ctatatcact atcaggagtg tgtagagggt acgactgtac 27360  
 tactaaaaga accttgccca tcaggaacat acgagggcaa ttcaccattt caccctcttg 27420  
 ctgacaataa atttgacta acttgacta gcacacactt tgcttttgct tgtgctgacg 27480  
 gtactcgaca tacctatcag ctgctgcaa gatcagtttc accaaaactt ttcacagac 27540  
 40 aagaggaggt tcaacaagag ctctactcgc cacttttctt cattgttgct gctctagtat 27600  
 ttttaatact ttgcttcacc attagagaa agacagaatg aatgagctca ctttaattga 27660  
 cttctatttg tgctttttag ctttctgct attccttggt ttaataatgc ttattatatt 27720  
 45 ttggttttca ctcgaaatcc aggatctaga agaacttggt accaaagtct aaacgaacat 27780  
 gaaacttctc attgttttga cttgtatttc tctatgcagt tgcatatgca ctgtagtaca 27840  
 gcgctgtgca tctaataaac ctcatgtgct tgaagatcct tgtaaggtag aacactaggg 27900  
 50 gtaatactta tagcactgct tggctttgtg ctctaggaaa ggttttacct tttcatagat 27960  
 ggcacactat ggttcaaaca tgcacaccta atgttactat caactgtcaa gatccagctg 28020  
 gtggtgctgct tatagctagg tgttggtacc ttcatgaagg tcaccaaact gctgcattta 28080  
 55 gagacgtact tgttggttta aataaacgaa caaattaaaa tgtctgataa tggaccccaa 28140

EP 1 694 829 B1

tcaaaccaac gtagtgcccc ccgcattaca tttggtggac ccacagattc aactgacaat 28200  
 aaccagaatg gaggacgcaa tggggcaagg ccaaaacagc gccgaccca aggtttaccc 28260  
 5 aataatactg cgtcttggtt cacagctctc actcagcatg gcaaggagga acttagattc 28320  
 cctcgaggcc agggcggtcc aatcaacacc aatagtggtc cagatgacca aattggctac 28380  
 taccgaagag ctacccgacg agttcgtggt ggtgacggca aaatgaaaga gctcagcccc 28440  
 10 agatggtact tctattacct aggaactggc ccagaagctt cacttcccta cggcgctaac 28500  
 aaagaaggca tcgtatgggt tgcaactgag ggagccttga atacacccaa agaccacatt 28560  
 ggcacccgca atcctaataa caatgctgcc accgtgctac aacttcctca aggaacaaca 28620  
 15 ttgccaaaag gcttctacgc agaggggaagc agaggcggca gtcaagcctc ttctcgctcc 28680  
 tcatcacgta gtcgcggtaa ttcaagaaat tcaactcctg gcagcagtag gggaaattct 28740  
 cctgctcgaa tggctagcgg aggtggtgaa actgccctcg cgctattgct gctagacaga 28800  
 20 ttgaaccagc ttgagagcaa agtttctggt aaaggccaac aacaacaagg ccaaactgtc 28860  
 actaagaaat ctgctgctga ggcatctaaa aagcctcgcc aaaaacgtac tgccacaaaa 28920  
 cagtacaacg tcaactaagc atttgggaga cgtggtccag aacaaacca aggaaatttc 28980  
 25 ggggaccaag acctaatacag acaagggaact gattacaaac attggccgca aattgcacaa 29040  
 tttgctccaa gtgcctctgc attctttgga atgtcacgca ttggcatgga agtcacacct 29100  
 tcgggaacat ggctgactta tcatggagcc attaaattgg atgacaaaga tccacaattc 29160  
 30 aaagacaacg tcatactgct gaacaagcac attgacgcat acaaaacatt cccaccaaca 29220  
 gagcctaaaa aggacaaaaa gaaaaagact gatgaagctc agcctttgcc gcagagacaa 29280  
 aagaagcagc ccaactgtgac tcttcttcct gcggctgaca tggatgattt ctccagacaa 29340  
 35 cttcaaaatt ccatgagtgg agcttctgct gattcaactc aggcataaac actcatgatg 29400  
 accacacaag gcagatgggc tatgtaaacg ttttcgcaat tccgtttacg atacatagtc 29460  
 tactcttggtg cagaatgaat tctcgttaact aaacagcaca agtaggttta gttaacttta 29520  
 40 atctcacata gcaatcttta atcaatgtgt aacattaggg aggacttgaa agagccacca 29580  
 ctttttcacg gaggccacgc ggagtacgat cgagggtaca gtgaataatg ctaggagagag 29640  
 ctgcctatat ggaagagccc taatgtgtaa aattaatttt agtagtgcta tccccatgtg 29700  
 45 attttaatag cttcttagga gaatgacaaa aaaaaaaaaa aaaaaa 29746

<210> 2

<211> 3945

<212> DNA

50 <213> CORONAVIRUS

<220>

<221> CDS

<222> (89)..(3853)

55 <223>

<400> 2

	ttctcttctg	gaaaaaggta	ggcttatcat	tagagaaaac	aacagagttg	tggtttcaag	60									
	tgatattctt	gttaacaact	aaacgaac	atg ttt att ttc tta tta ttt ctt		112										
5				Met	Phe	Ile	Phe	Leu	Leu	Phe	Leu					
				1								5				
	act ctc act agt ggt agt gac ctt gac cgg tgc acc act ttt gat gat															160
	Thr Leu Thr Ser Gly Ser Asp Leu Asp Arg Cys Thr Thr Phe Asp Asp															
	10															
	gtt caa gct cct aat tac act caa cat act tca tct atg agg ggg gtt															208
	Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val															
	25															
	tac tat cct gat gaa att ttt aga tca gac act ctt tat tta act cag															256
15	Tyr Tyr Pro Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr Leu Thr Gln															
	45															
	gat tta ttt ctt cca ttt tat tct aat gtt aca ggg ttt cat act att															304
	Asp Leu Phe Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe His Thr Ile															
	60															
	aat cat acg ttt ggc aac cct gtc ata cct ttt aag gat ggt att tat															352
20	Asn His Thr Phe Gly Asn Pro Val Ile Pro Phe Lys Asp Gly Ile Tyr															
	75															
	ttt gct gcc aca gag aaa tca aat gtt gtc cgt ggt tgg gtt ttt ggt															400
25	Phe Ala Ala Thr Glu Lys Ser Asn Val Val Arg Gly Trp Val Phe Gly															
	90															
	tct acc atg aac aac aag tca cag tgc gtg att att att aac aat tct															448
	Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile Ile Asn Asn Ser															
	105															
	act aat gtt gtt ata cga gca tgt aac ttt gaa ttg tgt gac aac cct															496
30	Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu Cys Asp Asn Pro															
	125															
	ttc ttt gct gtt tct aaa ccc atg ggt aca cag aca cat act atg ata															544
35	Phe Phe Ala Val Ser Lys Pro Met Gly Thr Gln Thr His Thr Met Ile															
	140															
	ttc gat aat gca ttt aat tgc act ttc gag tac ata tct gat gcc ttt															592
	Phe Asp Asn Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser Asp Ala Phe															
	155															
	tcg ctt gat gtt tca gaa aag tca ggt aat ttt aaa cac tta cga gag															640
40	Ser Leu Asp Val Ser Glu Lys Ser Gly Asn Phe Lys His Leu Arg Glu															
	170															
	ttt gtg ttt aaa aat aaa gat ggg ttt ctc tat gtt tat aag ggc tat															688
45	Phe Val Phe Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr Lys Gly Tyr															
	185															
	caa cct ata gat gta gtt cgt gat cta cct tct ggt ttt aac act ttg															736
	Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu															
	205															
	aaa cct att ttt aag ttg cct ctt ggt att aac att aca aat ttt aga															784
50																

EP 1 694 829 B1

	Lys	Pro	Ile	Phe 220	Lys	Leu	Pro	Leu	Gly 225	Ile	Asn	Ile	Thr	Asn 230	Phe	Arg	
5	gcc Ala	att Ile	ctt Leu 235	aca Thr	gcc Ala	ttt Phe	tca Ser	cct Pro 240	gct Ala	caa Gln	gac Asp	att Ile	tgg Trp 245	ggc Gly	acg Thr	tca Ser	832
	gct Ala	gca Ala 250	gcc Ala	tat Tyr	ttt Phe	gtt Val	ggc Gly 255	tat Tyr	tta Leu	aag Lys	cca Pro	act Thr 260	aca Thr	ttt Phe	atg Met	ctc Leu	880
10	aag Lys 265	tat Tyr	gat Asp	gaa Glu	aat Asn	ggt Gly 270	aca Thr	atc Ile	aca Thr	gat Asp	gct Ala 275	gtt Val	gat Asp	tgt Cys	tct Ser	caa Gln 280	928
15	aat Asn	cca Pro	ctt Leu	gct Ala	gaa Glu 285	ctc Leu	aaa Lys	tgc Cys	tct Ser	gtt Val 290	aag Lys	agc Ser	ttt Phe	gag Glu	att Ile 295	gac Asp	976
	aaa Lys	gga Gly	att Ile	tac Tyr 300	cag Gln	acc Thr	tct Ser	aat Asn	ttc Phe 305	agg Arg	gtt Val	gtt Val	ccc Pro	tca Ser 310	gga Gly	gat Asp	1024
20	gtt Val	gtg Val	aga Arg 315	ttc Phe	cct Pro	aat Asn	att Ile	aca Thr 320	aac Asn	ttg Leu	tgt Cys	cct Pro	ttt Phe 325	gga Gly	gag Glu	gtt Val	1072
25	ttt Phe	aat Asn 330	gct Ala	act Thr	aaa Lys	ttc Phe	cct Pro 335	tct Ser	gtc Val	tat Tyr	gca Ala	tgg Trp 340	gag Glu	aga Arg	aaa Lys	aaa Lys	1120
	att Ile 345	tct Ser	aat Asn	tgt Cys	gtt Val	gct Ala 350	gat Asp	tac Tyr	tct Ser	gtg Val	ctc Leu 355	tac Tyr	aac Asn	tca Ser	aca Thr	ttt Phe 360	1168
30	ttt Phe	tca Ser	acc Thr	ttt Phe	aag Lys 365	tgc Cys	tat Tyr	ggc Gly	gtt Val	tct Ser 370	gcc Ala	act Thr	aag Lys	ttg Leu	aat Asn 375	gat Asp	1216
35	ctt Leu	tgc Cys	ttc Phe	tcc Ser 380	aat Asn	gtc Val	tat Tyr	gca Ala	gat Asp 385	tct Ser	ttt Phe	gta Val	gtc Val	aag Lys 390	gga Gly	gat Asp	1264
	gat Asp	gta Val	aga Arg 395	caa Gln	ata Ile	gcg Ala	cca Pro	gga Gly 400	caa Gln	act Thr	ggc Gly	gtt Val	att Ile 405	gct Ala	gat Asp	tat Tyr	1312
40	aat Asn	tat Tyr 410	aaa Lys	ttg Leu	cca Pro	gat Asp	gat Asp 415	ttc Phe	atg Met	ggc Gly	tgt Cys	gtc Val 420	ctt Leu	gct Ala	tgg Trp	aat Asn	1360
45	act Thr 425	agg Arg	aac Asn	att Ile	gat Asp	gct Ala 430	act Thr	tca Ser	act Thr	ggc Gly 435	aat Asn	tat Tyr	aat Asn	tat Tyr	aaa Lys	tat Tyr 440	1408
	agg Arg	tat Tyr	ctt Leu	aga Arg	cat His 445	ggc Gly	aag Lys	ctt Leu	agg Arg	ccc Pro 450	ttt Phe	gag Glu	aga Arg	gac Asp	ata Ile 455	tct Ser	1456
50	aat Asn	gtg Val	cct Pro	ttc Phe 460	tcc Ser	cct Pro	gat Asp	ggc Gly	aaa Lys 465	cct Pro	tgc Cys	acc Thr	cca Pro	cct Pro 470	gct Ala	ctt Leu	1504
55	aat Asn	tgt Cys	tat Tyr 475	tgg Trp	cca Pro	tta Leu	aat Asn	gat Asp 480	tat Tyr	ggc Gly	ttt Phe	tac Tyr	acc Thr 485	act Thr	act Thr	ggc Gly	1552
	att	ggc	tac	caa	cct	tac	aga	gtt	gta	gta	ctt	tct	ttt	gaa	ctt	tta	1600

EP 1 694 829 B1

	Ile	Gly 490	Tyr	Gln	Pro	Tyr	Arg 495	Val	Val	Val	Leu	Ser 500	Phe	Glu	Leu	Leu	
5	aat Asn 505	gca Ala	ccg Pro	gcc Ala	acg Thr	gtt Val 510	tgt Cys	gga Gly	cca Pro	aaa Lys	tta Leu 515	tcc Ser	act Thr	gac Asp	ctt Leu	att Ile 520	1648
	aag Lys	aac Asn	cag Gln	tgt Cys	gtc Val 525	aat Asn	ttt Phe	aat Asn	ttt Phe	aat Asn 530	gga Gly	ctc Leu	act Thr	ggg Gly	act Thr	ggg Gly	1696
10	gtg Val	tta Leu	act Thr	cct Pro 540	tct Ser	tca Ser	aag Lys	aga Arg	ttt Phe 545	caa Gln	cca Pro	ttt Phe	caa Gln	caa Gln 550	ttt Phe	ggc Gly	1744
15	cgt Arg	gat Asp	gtt Val 555	tct Ser	gat Asp	ttc Phe	act Thr	gat Asp 560	tcc Ser	gtt Val	cga Arg	gat Asp	cct Pro 565	aaa Lys	aca Thr	tct Ser	1792
	gaa Glu	ata Ile 570	tta Leu	gac Asp	att Ile	tca Ser	cct Pro 575	tgc Cys	tct Ser	ttt Phe	ggg Gly 580	ggg Gly	gta Val	agt Ser	gta Val	att Ile	1840
20	aca Thr 585	cct Pro	gga Gly	aca Thr	aat Asn	gct Ala 590	tca Ser	tct Ser	gaa Glu	gtt Val	gct Ala 595	gtt Val	cta Leu	tat Tyr	caa Gln	gat Asp 600	1888
25	gtt Val	aac Asn	tgc Cys	act Thr	gat Asp 605	gtt Val	tct Ser	aca Thr	gca Ala	att Ile 610	cat His	gca Ala	gat Asp	caa Gln	ctc Leu 615	aca Thr	1936
	cca Pro	gct Ala	tgg Trp	cgc Arg 620	ata Ile	tat Tyr	tct Ser	act Thr	gga Gly 625	aac Asn	aat Asn	gta Val	ttc Phe	cag Gln 630	act Thr	caa Gln	1984
30	gca Ala	ggc Gly	tgt Cys 635	ctt Leu	ata Ile	gga Gly	gct Ala	gag Glu 640	cat His	gtc Val	gac Asp	act Thr	tct Ser 645	tat Tyr	gag Glu	tgc Cys	2032
35	gac Asp	att Ile 650	cct Pro	att Ile	gga Gly	gct Ala	ggc Gly 655	att Ile	tgt Cys	gct Ala	agt Ser	tac Tyr 660	cat His	aca Thr	gtt Val	tct Ser	2080
	tta Leu 665	tta Leu	cgt Arg	agt Ser	act Thr	agc Ser 670	caa Gln	aaa Lys	tct Ser	att Ile	gtg Val 675	gct Ala	tat Tyr	act Thr	atg Met	tct Ser 680	2128
40	tta Leu	ggg Gly	gct Ala	gat Asp	agt Ser 685	tca Ser	att Ile	gct Ala	tac Tyr	tct Ser 690	aat Asn	aac Asn	acc Thr	att Ile	gct Ala 695	ata Ile	2176
45	cct Pro	act Thr	aac Asn	ttt Phe 700	tca Ser	att Ile	agc Ser	att Ile	act Thr 705	aca Thr	gaa Glu	gta Val	atg Met	cct Pro 710	gtt Val	tct Ser	2224
	atg Met	gct Ala	aaa Lys 715	acc Thr	tcc Ser	gta Val	gat Asp	tgt Cys 720	aat Asn	atg Met	tac Tyr	atc Ile	tgc Cys 725	gga Gly	gat Asp	tct Ser	2272
50	act Thr	gaa Glu 730	tgt Cys	gct Ala	aat Asn	ttg Leu	ctt Leu 735	ctc Leu	caa Gln	tat Tyr	ggg Gly 740	agc Ser 745	ttt Phe	tgc Cys	aca Thr	caa Gln	2320
55	cta Leu 745	aat Asn	cgt Arg	gca Ala	ctc Leu	tca Ser 750	ggg Gly	att Ile	gct Ala	gct Ala	gaa Glu 755	cag Gln	gat Asp	cgc Arg	aac Asn	aca Thr 760	2368
	cgt	gaa	gtg	ttc	gct	caa	gtc	aaa	caa	atg	tac	aaa	acc	cca	act	ttg	2416



EP 1 694 829 B1

	Arg	Glu	Val	Phe	Ala 765	Gln	Val	Lys	Gln	Met 770	Tyr	Lys	Thr	Pro	Thr 775	Leu	
5	aaa Lys	tat Tyr	ttt Phe	ggt Gly 780	ggt Gly	ttt Phe	aat Asn	ttt Phe	tca Ser 785	caa Gln	ata Ile	tta Leu	cct Pro	gac Asp 790	cct Pro	cta Leu	2464
	aag Lys	cca Pro	act Thr 795	aag Lys	agg Arg	tct Ser	ttt Phe	att Ile 800	gag Glu	gac Asp	ttg Leu	ctc Leu	ttt Phe 805	aat Asn	aag Lys	gtg Val	2512
10	aca Thr	ctc Leu 810	gct Ala	gat Asp	gct Ala	ggc Gly	ttc Phe 815	atg Met	aag Lys	caa Gln	tat Tyr	ggc Gly 820	gaa Glu	tgc Cys	cta Leu	ggt Gly	2560
15	gat Asp 825	att Ile	aat Asn	gct Ala	aga Arg	gat Asp 830	ctc Leu	att Ile	tgt Cys	gcg Ala	cag Gln 835	aag Lys	ttc Phe	aat Asn	gga Gly	ctt Leu 840	2608
	aca Thr	gtg Val	ttg Leu	cca Pro	cct Pro 845	ctg Leu	ctc Leu	act Thr	gat Asp	gat Asp 850	atg Met	att Ile	gct Ala	gcc Ala	tac Tyr 855	act Thr	2656
20	gct Ala	gct Ala	cta Leu	gtt Val 860	agt Ser	ggc Gly	act Thr	gcc Ala	act Thr 865	gct Ala	gga Gly	tgg Trp	aca Thr	ttt Phe 870	ggt Gly	gct Ala	2704
25	ggc Gly	gct Ala	gct Ala 875	ctt Leu	caa Gln	ata Ile	cct Pro	ttt Phe 880	gct Ala	atg Met	caa Gln	atg Met	gca Ala 885	tat Tyr	agg Arg	ttc Phe	2752
	aat Asn	ggc Gly 890	att Ile	gga Gly	gtt Val	acc Thr	caa Gln 895	aat Asn	gtt Val	ctc Leu	tat Tyr	gag Glu 900	aac Asn	caa Gln	aaa Lys	caa Gln	2800
30	atc Ile 905	gcc Ala	aac Asn	caa Gln	ttt Phe	aac Asn 910	aag Lys	gcg Ala	att Ile	agt Ser	caa Gln 915	att Ile	caa Gln	gaa Glu	tca Ser	ctt Leu 920	2848
35	aca Thr	aca Thr	aca Thr	tca Ser	act Thr 925	gca Ala	ttg Leu	ggc Gly	aag Lys	ctg Leu 930	caa Gln	gac Asp	gtt Val	gtt Val	aac Asn 935	cag Gln	2896
	aat Asn	gct Ala	caa Gln	gca Ala 940	tta Leu	aac Asn	aca Thr	ctt Leu	gtt Val 945	aaa Lys	caa Gln	ctt Leu	agc Ser	tct Ser 950	aat Asn	ttt Phe	2944
40	ggt Gly	gca Ala	att Ile 955	tca Ser	agt Ser	gtg Val	cta Leu	aat Asn 960	gat Asp	atc Ile	ctt Leu	tgc Ser	cga Arg 965	ctt Leu	gat Asp	aaa Lys	2992
45	gtc Val	gag Glu 970	gcg Ala	gag Glu	gta Val	caa Gln	att Ile 975	gac Asp	agg Arg	tta Leu	att Ile	aca Thr 980	ggc Gly	aga Arg	ctt Leu	caa Gln	3040
	agc Ser 985	ctt Leu	caa Gln	acc Thr	tat Tyr	gta Val 990	aca Thr	caa Gln	caa Gln	cta Leu	atc Ile 995	agg Arg	gct Ala	gct Ala	gaa Glu	atc Ile 1000	3088
50	agg Arg	gct Ala	tct Ser	gct Ala	aat Asn 1005	ctt Leu	gct Ala	gct Ala	act Thr	aaa Lys 1010	atg Met	tct Ser	gag Glu	tgt Cys	gtt Val 1015		3133
	ctt Leu	gga Gly	caa Gln	tca Ser	aaa Lys 1020	aga Arg	gtt Val	gac Asp	ttt Phe	tgt Cys 1025	gga Gly	aag Lys	ggc Gly	tac Tyr	cac His 1030		3178
55	ctt	atg	tcc	ttc	cca	caa	gca	gcc	ccg	cat	ggt	gtt	gtc	ttc	cta		3223

EP 1 694 829 B1

	Leu	Met	Ser	Phe	Pro 1035	Gln	Ala	Ala	Pro	His 1040	Gly	Val	Val	Phe	Leu 1045	
5	cat	gtc	acg	tat	gtg Val 1050	cca	tcc	cag	gag	agg Arg 1055	aac	ttc	acc	aca	gcg Ala 1060	3268
	His	Val	Thr	Tyr		Pro	Ser	Gln	Glu		Asn	Phe	Thr	Thr		
10	cca	gca	att	tgt	cat His 1065	gaa	ggc	aaa	gca	tac Tyr 1070	ttc	cct	cgt	gaa	ggc Gly 1075	3313
	Pro	Ala	Ile	Cys		Glu	Gly	Lys	Ala		Phe	Pro	Arg	Glu		
	gtt	ttt	gtg	ttt	aat Asn 1080	ggc	act	tct	tgg	ttt Phe 1085	att	aca	cag	agg	aac Asn 1090	3358
	Val	Phe	Val	Phe		Gly	Thr	Ser	Trp		Ile	Thr	Gln	Arg		
15	ttc	ttt	tct	cca	caa Gln 1095	ata	att	act	aca	gac Asp 1100	aat	aca	ttt	gtc	tca Ser 1105	3403
	Phe	Phe	Ser	Pro		Ile	Ile	Thr	Thr		Asn	Thr	Phe	Val		
20	gga	aat	tgt	gat	gtc Val 1110	gtt	att	ggc	atc	att Ile 1115	aac	aac	aca	gtt	tat Tyr 1120	3448
	Gly	Asn	Cys	Asp		Val	Ile	Gly	Ile		Asn	Asn	Thr	Val		
	gat	cct	ctg	caa	cct Pro 1125	gag	ctt	gac	tca	ttc Phe 1130	aaa	gaa	gag	ctg	gac Asp 1135	3493
	Asp	Pro	Leu	Gln		Glu	Leu	Asp	Ser		Lys	Glu	Glu	Leu		
25	aag	tac	ttc	aaa	aat Asn 1140	cat	aca	tca	cca	gat Asp 1145	gtt	gat	ctt	ggc	gac Asp 1150	3538
	Lys	Tyr	Phe	Lys		His	Thr	Ser	Pro		Val	Asp	Leu	Gly		
30	att	tca	ggc	att	aac Asn 1155	gct	tct	gtc	gtc	aac Asn 1160	att	caa	aaa	gaa	att Ile 1165	3583
	Ile	Ser	Gly	Ile		Ala	Ser	Val	Val		Ile	Gln	Lys	Glu		
	gac	cgc	ctc	aat	gag Glu 1170	gtc	gct	aaa	aat	tta Leu 1175	aat	gaa	tca	ctc	att Ile 1180	3628
	Asp	Arg	Leu	Asn		Val	Ala	Lys	Asn		Asn	Glu	Ser	Leu		
35	gac	ctt	caa	gaa	ttg Leu 1185	gga	aaa	tat	gag	caa Gln 1190	tat	att	aaa	tgg	cct Pro 1195	3673
	Asp	Leu	Gln	Glu		Gly	Lys	Tyr	Glu		Tyr	Ile	Lys	Trp		
40	tgg	tat	gtt	tgg	ctc Leu 1200	ggc	ttc	att	gct	gga Gly 1205	cta	att	gcc	atc	gtc Val 1210	3718
	Trp	Tyr	Val	Trp		Gly	Phe	Ile	Ala		Leu	Ile	Ala	Ile		
	atg	gtt	aca	atc	ttg Leu 1215	ctt	tgt	tgc	atg	act Thr 1220	agt	tgt	tgc	agt	tgc Cys 1225	3763
	Met	Val	Thr	Ile		Leu	Cys	Cys	Met		Ser	Cys	Cys	Ser		
45	ctc	aag	ggc	gca	tgc Cys 1230	tct	tgt	ggc	tct	tgc Cys 1235	tgc	aag	ttt	gat	gag Glu 1240	3808
	Leu	Lys	Gly	Ala		Ser	Cys	Gly	Ser		Cys	Lys	Phe	Asp		
50	gat	gac	tct	gag	cca Pro 1245	gtt	ctc	aag	ggc	gtc Val 1250	aaa	tta	cat	tac	aca Thr 1255	3853
	Asp	Asp	Ser	Glu		Val	Leu	Lys	Gly		Lys	Leu	His	Tyr		
	taaacgaact tatggatttg tttatgagat tttttactct tggatcaatt actgcacagc															3913
55	cagtaaaaat tgacaatgct tctcctgcaa gt															3945

<210> 3

<211> 1255

EP 1 694 829 B1

<212> PRT

<213> CORONAVIRUS

<400> 3

5

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu  
1 5 10 15

10

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln  
20 25 30

15

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg  
35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser  
50 55 60

20

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val  
65 70 75 80

25

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn  
85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln  
100 105 110

30

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys  
115 120 125

35

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met  
130 135 140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr  
145 150 155 160

40

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser  
165 170 175

45

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly  
180 185 190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp  
195 200 205

50

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu  
210 215 220

55

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro  
225 230 235 240

EP 1 694 829 B1

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Tyr Phe Val Gly Tyr  
245 250 255

5 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile  
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys  
10 275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn  
290 295 300

15 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr  
305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser  
20 325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr  
340 345 350

25 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly  
355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala  
30 370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly  
385 390 395 400

35 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe  
405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser  
40 420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu  
435 440 445

45 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly  
450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp  
50 465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val  
485 490 495

55 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly  
500 505 510

EP 1 694 829 B1

Pro Lys Leu<sub>515</sub> Ser Thr Asp Leu Ile<sub>520</sub> Lys Asn Gln Cys Val<sub>525</sub> Asn Phe Asn

5 Phe Asn<sub>530</sub> Gly Leu Thr Gly Thr<sub>535</sub> Gly Val Leu Thr Pro<sub>540</sub> Ser Ser Lys Arg

10 Phe Gln Pro Phe Gln Gln<sub>550</sub> Phe Gly Arg Asp Val<sub>555</sub> Ser Asp Phe Thr Asp<sub>560</sub>

Ser Val Arg Asp Pro<sub>565</sub> Lys Thr Ser Glu Ile<sub>570</sub> Leu Asp Ile Ser Pro<sub>575</sub> Cys

15 Ser Phe Gly Gly<sub>580</sub> Val Ser Val Ile Thr<sub>585</sub> Pro Gly Thr Asn Ala<sub>590</sub> Ser Ser

Glu Val Ala<sub>595</sub> Val Leu Tyr Gln Asp<sub>600</sub> Val Asn Cys Thr Asp<sub>605</sub> Val Ser Thr

20 Ala Ile<sub>610</sub> His Ala Asp Gln Leu<sub>615</sub> Thr Pro Ala Trp Arg<sub>620</sub> Ile Tyr Ser Thr

25 Gly Asn Asn Val Phe Gln<sub>630</sub> Thr Gln Ala Gly Cys<sub>635</sub> Leu Ile Gly Ala Glu<sub>640</sub>

His Val Asp Thr Ser<sub>645</sub> Tyr Glu Cys Asp Ile<sub>650</sub> Pro Ile Gly Ala Gly<sub>655</sub> Ile

30 Cys Ala Ser Tyr<sub>660</sub> His Thr Val Ser Leu<sub>665</sub> Leu Arg Ser Thr Ser<sub>670</sub> Gln Lys

35 Ser Ile Val<sub>675</sub> Ala Tyr Thr Met Ser<sub>680</sub> Leu Gly Ala Asp Ser<sub>685</sub> Ser Ile Ala

Tyr Ser<sub>690</sub> Asn Asn Thr Ile Ala<sub>695</sub> Ile Pro Thr Asn Phe<sub>700</sub> Ser Ile Ser Ile

40 Thr Thr Glu Val Met Pro<sub>710</sub> Val Ser Met Ala Lys<sub>715</sub> Thr Ser Val Asp Cys<sub>720</sub>

45 Asn Met Tyr Ile Cys<sub>725</sub> Gly Asp Ser Thr Glu<sub>730</sub> Cys Ala Asn Leu Leu<sub>735</sub> Leu

Gln Tyr Gly Ser<sub>740</sub> Phe Cys Thr Gln Leu<sub>745</sub> Asn Arg Ala Leu Ser<sub>750</sub> Gly Ile

50 Ala Ala Glu<sub>755</sub> Gln Asp Arg Asn Thr<sub>760</sub> Arg Glu Val Phe Ala<sub>765</sub> Gln Val Lys

55 Gln Met<sub>770</sub> Tyr Lys Thr Pro Thr<sub>775</sub> Leu Lys Tyr Phe Gly<sub>780</sub> Gly Phe Asn Phe

EP 1 694 829 B1

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile  
 785 790 795 800  
 5  
 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met  
 805 810 815  
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile  
 820 825 830  
 10  
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr  
 835 840 845  
 Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala  
 850 855 860  
 15  
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe  
 865 870 875 880  
 20  
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn  
 885 890 895  
 Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala  
 900 905 910  
 25  
 Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly  
 915 920 925  
 30  
 Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu  
 930 935 940  
 Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn  
 945 950 955 960  
 35  
 Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp  
 965 970 975  
 40  
 Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln  
 980 985 990  
 Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala  
 995 1000 1005  
 45  
 Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp  
 1010 1015 1020  
 50  
 Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala  
 1025 1030 1035  
 Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln  
 1040 1045 1050  
 55

EP 1 694 829 B1

	Glu	Arg	Asn	Phe	Thr	Thr	Ala	Pro	Ala	Ile	Cys	His	Glu	Gly	Lys
		1055					1060					1065			
5	Ala	Tyr	Phe	Pro	Arg	Glu	Gly	Val	Phe	Val	Phe	Asn	Gly	Thr	Ser
		1070					1075					1080			
10	Trp	Phe	Ile	Thr	Gln	Arg	Asn	Phe	Phe	Ser	Pro	Gln	Ile	Ile	Thr
		1085					1090					1095			
	Thr	Asp	Asn	Thr	Phe	Val	Ser	Gly	Asn	Cys	Asp	Val	Val	Ile	Gly
		1100					1105					1110			
15	Ile	Ile	Asn	Asn	Thr	Val	Tyr	Asp	Pro	Leu	Gln	Pro	Glu	Leu	Asp
		1115					1120					1125			
20	Ser	Phe	Lys	Glu	Glu	Leu	Asp	Lys	Tyr	Phe	Lys	Asn	His	Thr	Ser
		1130					1135					1140			
	Pro	Asp	Val	Asp	Leu	Gly	Asp	Ile	Ser	Gly	Ile	Asn	Ala	Ser	Val
		1145					1150					1155			
25	Val	Asn	Ile	Gln	Lys	Glu	Ile	Asp	Arg	Leu	Asn	Glu	Val	Ala	Lys
		1160					1165					1170			
30	Asn	Leu	Asn	Glu	Ser	Leu	Ile	Asp	Leu	Gln	Glu	Leu	Gly	Lys	Tyr
		1175					1180					1185			
	Glu	Gln	Tyr	Ile	Lys	Trp	Pro	Trp	Tyr	Val	Trp	Leu	Gly	Phe	Ile
		1190					1195					1200			
35	Ala	Gly	Leu	Ile	Ala	Ile	Val	Met	Val	Thr	Ile	Leu	Leu	Cys	Cys
		1205					1210					1215			
40	Met	Thr	Ser	Cys	Cys	Ser	Cys	Leu	Lys	Gly	Ala	Cys	Ser	Cys	Gly
		1220					1225					1230			
	Ser	Cys	Cys	Lys	Phe	Asp	Glu	Asp	Asp	Ser	Glu	Pro	Val	Leu	Lys
		1235					1240					1245			
45	Gly	Val	Lys	Leu	His	Tyr	Thr								
		1250					1255								
50	<210> 4														
	<211> 3943														
	<212> DNA														
	<213> CORONAVIRUS														
55	<400> 4														

EP 1 694 829 B1

ctcttctgga aaaaggtagg cttatcatta gagaaaaca cagagttgtg gtttcaagtg 60

5

10

15

20

25

30

35

40

45

50

55



EP 1 694 829 B1

	atattcttgt taacaactaa acgaacatgt ttattttctt attatttctt actctcacta	120
	gtggtagtga ccttgaccgg tgcaccactt ttgatgatgt tcaagctcct aattacactc	180
5	aacatacttc atctatgagg ggggtttact atcctgatga aatttttaga tcagacactc	240
	tttatttaac tcaggattta tttcttccat tttattctaa tgttacaggg tttcatacta	300
	ttaatcatac gtttggaac cctgtcatac cttttaagga tggatttat tttgctgcca	360
10	cagagaaatc aaatgttgtc cgtgggtggg tttttggttc taccatgaac aacaagtcac	420
	agtcggtgat tattattaac aattctacta atgttggtat acgagcatgt aactttgaat	480
	tgtgtgacaa ccttttctt gctgtttcta aacccatggg tacacagaca catactatga	540
15	tattcgataa tgcatttaac tgcactttcg agtacatatc tgatgccttt tcgcttgatg	600
	tttcagaaaa gtcaggtaac tttaaacact tacgagagtt tgtgtttaa aataaagatg	660
	ggtttctcta tgtttataag ggctatcaac ctatagatgt agttcgtgat ctaccttctg	720
20	gttttaacac tttgaaacct atttttaagt tgcctcttgg tattaacatt acaaatttta	780
	gagccattct tacagccttt tcacctgctc aagacatttg gggcacgtca gctgcagcct	840
	attttggttg ctattttaag ccaactacat ttatgctcaa gtatgatgaa aatggtacaa	900
25	tcacagatgc tgttgattgt tctcaaaatc cacttgctga actcaaatgc tctgttaaga	960
	gctttgagat tgacaaagga atttaccaga cctctaattt cagggttgtt ccctcaggag	1020
	atgttgtagag attccctaac attacaaact tgtgtccttt tggagaggtt tttaatgcta	1080
30	ctaaattccc ttctgtctat gcatgggaga gaaaaaaaaat ttctaattgt gttgctgatt	1140
	actctgtgct ctacaactca acatTTTTTT caacctttaa gtgctatggc gtttctgcca	1200
	ctaagttgaa tgatctttgc ttctccaatg tctatgcaga ttcttttgta gtcaagggag	1260
35	atgatgtaag acaaatagcg ccaggacaaa ctggtgttat tgctgattat aattataaat	1320
	tgccagatga ttcatgggt tgtgtccttg cttggaatac taggaacatt gatgctactt	1380
	caactggtaa ttataattat aaatataggt atcttagaca tggcaagctt aggccctttg	1440
40	agagagacat atctaattgt cctttctccc ctgatggcaa acctgcacc ccacctgctc	1500
	ttaatgtta ttggccatta aatgattatg gtttttacac cactactggc attggctacc	1560
	aaccttacag agttgtagta ctttcttttg aacttttaaa tgcaccggcc acggtttgtg	1620
45	gaccaaatt atccactgac cttattaaga accagtgtgt caattttaat tttaatggac	1680
	tcactggtac tgggtgtgta actccttctt caaagagatt tcaaccattt caacaatttg	1740
	gccgtgatgt ctctgatttc actgattccg ttcgagatcc taaaacatct gaaatattag	1800
50	acatttcacc ttgctctttt gggggtgtaa gtgtaattac acctggaaca aatgcttcat	1860
	ctgaagttgc tgttctatat caagatgtta actgcactga tgtttctaca gcaatccatg	1920
	cagatcaact cacaccagct tggcgcatat attctactgg aaacaatgta ttccagactc	1980
55	aagcaggctg tcttatagga gctgagcatg tcgacacttc ttatgagtg gacattccta	2040
	ttggagctgg catttgctgct agttaceata cagtttcttt attacgtagt actagccaaa	2100

	aatctattgt	ggcttatact	atgtctttag	gtgctgatag	ttcaattgct	tactctaata	2160
	acaccattgc	tatacctact	aacttttcaa	ttagcattac	tacagaagta	atgcctgttt	2220
5	ctatggctaa	aacctccgta	gattgtaata	tgtacatctg	cggagattct	actgaatgtg	2280
	ctaatttgct	tctccaatat	ggtagctttt	gcacacaact	aaatcgtgca	ctctcaggta	2340
	ttgctgctga	acaggatcgc	aacacacgtg	aagtgttcgc	tcaagtcaaa	caaatgtaca	2400
10	aaaccccaac	tttgaaatat	tttggtggtt	ttaatttttc	acaaatatta	cctgaccctc	2460
	taaagccaac	taagaggtct	tttattgagg	acttgctctt	taataaggtg	acactcgtcg	2520
	atgctggctt	catgaagcaa	tatggcgaat	gcctagggtga	tattaatgct	agagatctca	2580
15	tttgtgcgca	gaagttcaat	gggcttacag	tgttgccacc	tctgctcact	gatgatatga	2640
	ttgctgccta	cactgctgct	ctagttagtg	gtactgccac	tgctggatgg	acatttggtg	2700
	ctggcgctgc	tcttcaaata	ccttttgcta	tgcaaatggc	atatagggtc	aatggcattg	2760
20	gagttacca	aaatgttctc	tatgagaacc	aaaaacaaat	cgccaaccaa	tttaacaagg	2820
	cgattagtca	aattcaagaa	tcacttacia	caacatcaac	tgcatggggc	aagctgcaag	2880
	acgttgtaa	ccagaatgct	caagcattaa	acacacttgt	taaacaactt	agctctaatt	2940
25	ttggtgcaat	ttcaagtgtg	ctaaatgata	tcctttcgcg	acttgataaa	gtcgaggcgg	3000
	aggtacaaat	tgacaggcta	attacaggca	gacttcaaag	ccttcaaacc	tatgtaacac	3060
	aacaactaat	cagggctgct	gaaatcaggg	cttctgctaa	tcttgctgct	actaaaatgt	3120
30	ctgagtgtgt	tcttggaaca	tcaaaaagag	ttgacttttg	tggaaagggc	taccacctta	3180
	tgtccttccc	acaagcagcc	ccgcatgggtg	ttgtcttcct	acatgtcacg	tatgtgccat	3240
	cccaggagag	gaacttcacc	acagcgccag	caatttgtca	tgaaggcaaa	gcatacttcc	3300
35	ctcgtgaagg	tgtttttgtg	tttaatggca	cttcttggtt	tattacacag	aggaacttct	3360
	tttctccaca	aataattact	acagacaata	catttgtctc	aggaaattgt	gatgtcgtta	3420
	ttggcatcat	taacaacaca	gtttatgatc	ctctgcaacc	tgagcttgac	tcattcaaag	3480
40	aagagctgga	caagtacttc	aaaaatcata	catcaccaga	tgttgatctt	ggcgacattt	3540
	caggcattaa	cgcttctgtc	gtcaacattc	aaaaagaaat	tgaccgcctc	aatgaggctc	3600
	ctaaaaattt	aaatgaatca	ctcattgacc	ttcaagaatt	gggaaaatat	gagcaatata	3660
45	ttaaatggcc	ttggtatggt	tggtcggct	tcattgctgg	actaattgcc	atcgatcatg	3720
	ttacaatctt	gctttgttgc	atgactagtt	gttgagttg	cctcaagggt	gcatgctctt	3780
	gtggttcttg	ctgcaagttt	gatgaggatg	actctgagcc	agttctcaag	ggtgtcaaat	3840
50	tacattacac	ataaacgaac	ttatggattt	gtttatgaga	ttttttactc	ttggatcaat	3900
	tactgcacag	ccagtaaaaa	ttgacaatgc	ttctcctgca	agt		3943
55	<210> 5						
	<211> 2049						
	<212> DNA						
	<213> CORONAVIRUS						

&lt;400&gt; 5

	ctcttctgga	aaaaggtagg	cttatcatta	gagaaaacaa	cagagttgtg	gtttcaagtg	60
5	atattcttgt	taacaactaa	acgaacatgt	ttattttctt	attatttctt	actctcacta	120
	gtggtagtga	ccttgaccgg	tgcaccactt	ttgatgatgt	tcaagctcct	aattacactc	180
	aacatacttc	atctatgagg	ggggtttact	atcctgatga	aattttttaga	tcagacactc	240
10	tttatttaac	tcaggattta	tttcttccat	tttattctaa	tgttacaggg	tttcatacta	300
	ttaatcatac	gtttggcaac	cctgtcatac	cttttaagga	tggtatttat	tttgctgcca	360
	cagagaaatc	aaatgttgct	cgtggttggg	tttttggttc	taccatgaac	aacaagtcac	420
15	agtcggtgat	tattattaac	aattctacta	atgttggtat	acgagcatgt	aactttgaat	480
	tgtgtgacaa	ccctttcttt	gctgtttcta	aacccatggg	tacacagaca	catactatga	540
	tattcgataa	tgcatttaat	tgcactttcg	agtacatatc	tgatgccttt	tcgcttgatg	600
20	tttcagaaaa	gtcaggtaat	tttaaacact	tacgagagtt	tgtgtttaaa	aataaagatg	660
	ggtttctcta	tgtttataag	ggctatcaac	ctatagatgt	agttcgtgat	ctaccttctg	720
	gttttaacac	tttgaaacct	atttttaagt	tgcttcttgg	tattaacatt	acaaatttta	780
25	gagccattct	tacagccttt	tcacctgctc	aagacatttg	gggcacgtca	gctgcagcct	840
	attttgttgg	ctattttaaag	ccaactacat	ttatgctcaa	gtatgatgaa	aatggtacaa	900
	tcacagatgc	tggtgattgt	tctcaaaatc	cacttgctga	actcaaatgc	tctgttaaga	960
30	gctttgagat	tgacaaagga	atttaccaga	cctctaattt	cagggttggt	ccctcaggag	1020
	atgttggtgag	attccctaata	attacaaact	tgtgtccttt	tgagagaggt	tttaattgcta	1080
	ctaaattccc	ttctgtctat	gcatgggaga	gaaaaaaaaat	ttctaattgt	gttgctgatt	1140
35	actctgtgct	ctacaactca	acatTTTTTT	caacctttta	gtgctatggc	gtttctgcca	1200
	ctaagttgaa	tgatctttgc	ttctccaatg	tctatgcaga	ttcttttgta	gtcaagggag	1260
	atgatgtaag	acaaatagcg	ccaggacaaa	ctgggtgttat	tgctgattat	aattataaat	1320
40	tgccagatga	tttcatgggt	tgtgtccttg	cttggaaatac	taggaacatt	gatgctactt	1380
	caactggtaa	ttataattat	aaatataggt	atcttagaca	tggaagcctt	aggccctttg	1440
	agagagacat	atctaattgt	cctttctccc	ctgatggcaa	accttgccac	ccacctgctc	1500
45	tttaattgtta	ttggccatta	aatgattatg	gtttttacac	cactactggc	attggctacc	1560
	aaccttacag	agttgtagta	ctttcttttg	aactttttaa	tgcaccggcc	acggtttgtg	1620
	gacaaaaatt	atccactgac	cttattaaga	accagtgtgt	caattttaat	tttaattggac	1680
50	tcactggtac	tggtgtgtta	actccttctt	caaagagatt	tcaaccattt	caacaatttg	1740
	gccgtgatgt	ctctgatttc	actgattccg	ttcgagatcc	taaaacatct	gaaatattag	1800
55	acatttcacc	ttgctctttt	gggggtgtaa	gtgtaattac	acctggaaca	aatgcttcat	1860

EP 1 694 829 B1

	ctgaagttgc tgttctatat caagatgtta actgcactga tgtttctaca gcaatccatg	1920
	cagatcaact cacaccagct tggcgcatat attctactgg aaacaatgta ttccagactc	1980
5	aagcaggctg tcttatagga gctgagcatg tcgacacttc ttatgagtgc gacattccta	2040
	ttggagctg	2049
10	<210> 6 <211> 2027 <212> DNA <213> CORONAVIRUS	
15	<400> 6	
	catgcagatc aactcacacc agcttggcgc atatattcta ctggaaacaa tgtattccag	60
20	actcaagcag gctgtcttat aggagctgag catgtcgaca cttcttatga gtgcgacatt	120
	cctattggag ctggcatttg tgctagttac catacagttt ctttattacg tagtactagc	180
	caaaaatcta ttgtggctta tactatgtct ttaggtgctg atagttcaat tgcttactct	240
25	aataacacca ttgctatacc tactaacttt tcaattagca ttactacaga agtaatgcct	300
	gtttctatgg ctaaaacctc cgtagattgt aatatgtaca tctgcggaga ttctactgaa	360
	tgtgctaatt tgcttctcca atatggtagc ttttgcacac aactaaatcg tgcactctca	420
30	ggtattgctg ctgaacagga tcgcaacaca cgtgaagtgt tcgctcaagt caaacaatg	480
	tacaaaacc caactttgaa atattttggt ggttttaatt tttcacaaat attacctgac	540
	cctctaaagc caactaagag gtcttttatt gaggacttgc tctttaataa ggtgacactc	600
35	gctgatgctg gcttcatgaa gcaatatggc gaatgcctag gtgatattaa tgctagagat	660
	ctcatttgctg cgcagaagtt caatgggctt acagtgttgc cacctctgct cactgatgat	720
	atgattgctg cctacactgc tgctctagtt agtggtactg ccactgctgg atggacattt	780
40	ggtgctggcg ctgctcttca aatacctttt gctatgcaaa tggcatatag gttcaatggc	840
	attggagtta cccaaaatgt tctctatgag aaccaaaaac aaatcgccaa ccaatttaac	900
	aaggcgatta gtcaaattca agaatacatt acaacaacat caactgcatt gggcaagctg	960
45	caagacgttg ttaaccagaa tgctcaagca ttaaacacac ttgttaaaca acttagctct	1020
	aattttggtg caatttcaag tgtgctaaat gatatccttt cgcgacttga taaagtcgag	1080
	gcggagggtac aaattgacag gttaattaca ggcagacttc aaagccttca aacctatgta	1140
50	acacaacaac taatcagggc tgctgaaatc agggcttctg ctaatcttgc tgctactaaa	1200
	atgtctgagt gtgttcttgg acaatcaaaa agagttgact tttgtggaaa gggctaccac	1260
	cttatgtcct tcccacaagc agccccgcat ggtgttgtct tcctacatgt cacgtatgtg	1320
55	ccatcccagg agaggaactt caccacagcg ccagcaattt gtcatgaagg caaagcatac	1380
	ttccctcgtg aagggtgttt tgtgtttaat ggcacttctt ggtttattac acagaggaac	1440

EP 1 694 829 B1

	ttcttttctc cacaataat tactacagac aatacatttg tctcaggaaa ttgtgatgtc	1500
	gttattggcg tcattaacaa cacagtttat gatcctctgc aacctgagct tgactcattc	1560
5	aaagaagagc tggacaagta cttcaaaaat catacatcac cagatggtga tcttggcgac	1620
	atttcaggca ttaacgcttc tgtcgtcaac attcaaaaag aaattgaccg cctcaatgag	1680
	gtcgctaaaa atttaaatga atcactcatt gaccttcaag aattgggaaa atatgagcaa	1740
10	tatattaaat ggccttggtg tgtttggtc ggcttcattg ctggactaat tgccatcgtc	1800
	atgggttaca tcttgctttg ttgcatgact agttgttgca gttgcctcaa ggggtgcatgc	1860
	tcttgtggtt cttgctgcaa gtttcatgag gatgactctg agccagttct caaggggtgc	1920
15	aaattacatt acacataaac gaacttatgg atttgtttat gagatttttt actcttggtg	1980
	caattactgc acagccagta aaaattgaca atgcttctcc tgcaagt	2027
20	<210> 7	
	<211> 1096	
	<212> DNA	
	<213> CORONAVIRUS	
25	<400> 7	
	tcttgctttg ttgcatgact agttgttgca gttgcctcaa ggggtgcatgc tcttgtggtt	60
	cttgcctgcaa gtttcatgag gatgactctg agccagttct caaggggtgc aaattacatt	120
30	acacataaac gaacttatgg atttgtttat gagatttttt actcttggtg caattactgc	180
	acagccagta aaaattgaca atgcttctcc tgcaagtact gttcatgcta cagcaacgat	240
	accgctacaa gcctcactcc ctttcggatg gcttggtatt ggcgttgcat tcttgcctgt	300
35	ttttcagagc gctacacaaa taattgcgct caataaaaga tggcagctag ccctttataa	360
	gggcttccag ttcatctgca atttactgct gctatttggt accatctatt cacatctttt	420
	gcttgcctgc gcaggtatgg aggcgcaatt tttgtacctc tatgccttga tataatttct	480
40	acaatgcctc aacgcatgta gaattattat gagatgttgg ctttgttgga agtgcaaatc	540
	caagaaccca ttactttatg atgccaacta ctttgtttgc tggcacacac ataactatga	600
	ctactgtata ccatataaca gtgtcacaga tacaattgtc gttactgaag gtgacggcat	660
45	ttcaacacca aaactcaaag aagactacca aattgggtgg tattctgagg ataggcactc	720
	agggtgttaa gactatgtcg ttgtacatgg ctatttcacc gaagtttact accagcttga	780
	gtctacacaa attactacag aactgggtat tgaaaatgct acattcttca tctttaacaa	840
50	gcttgtttaa gaccaccga atgtgcaaat acacacaatc gacggctctt caggagtgc	900
	taatccagca atggatccaa tttatgatga gccgacgacg actactagcg tgcccttgta	960
	agcacaagaa agtgagtacg aacttatgta ctattcgtt tcggaagaaa cagggtacgtt	1020
55	aatagttaat agcgtacttc ttttcttgc tttcgtggta ttcttgctag tcacactagc	1080

catccttact gcgctt

1096

&lt;210&gt; 8

&lt;211&gt; 1135

&lt;212&gt; DNA

&lt;213&gt; CORONAVIRUS

&lt;400&gt; 8

```

attgccatcg tcatggttac aatcttgctt tgttgcata ctagttggtg cagttgcctc    60
aaggggtgcat gctcttggtg ttcttgctgc aagtttgatg aggatgactc tgagccagtt    120
ctcaaggggtg tcaaattaca ttacacataa acgaacttat ggatttggtt atgagatttt    180
ttactcttggtg atcaattact gcacagccag taaaaattga caatgcttct cctgcaagta    240
ctgttcacatg tacagcaacg ataccgctac aagcctcact ccttttcgga tggcttggtta    300
ttggcggttgc atttcttgct gtttttcaga gcgctaccaa aataattgcg ctcaataaaa    360
gatggcagct agccctttat aagggcttcc agttcatttg caatttactg ctgctatttg    420
ttaccatcta ttcacatctt ttgcttgctg ctgcaggtat ggaggcgcaa tttttgtacc    480
tctatgcctt gatataatttt ctacaatgca tcaacgcatg tagaattatt atgagatggt    540
ggcttttggtg gaagtgcaca tccaagaacc cattacttta tgatgccaac tactttgttt    600
gctggcacac acataactat gactactgta taccatataa cagtgtcaca gatacaattg    660
tcgttactga aggtgacggc atttcaacac caaaactcaa agaagactac caaattgggtg    720
gttattctga ggataggcac tcagggtgta aagactatgt cgttgtacat ggctatttca    780
ccgaagttaa ctaccagctt gagtctacac aaattactac agacactggg attgaaaatg    840
ctacattctt catctttaac aagcttggtt aagaccacc gaatgtgcaa atacacacaa    900
tcgacggctc ttcaggaggt gctaattccag caatggatcc aatttatgat gagccgacga    960
cgactactag cgtgcctttg taagcacaag aaagtgagta cgaacttatg tactcattcg   1020
tttcggaaga aacagggtacg ttaatagtta atagcgctact tctttttctt gctttcgtgg   1080
tattcttgct agtcacacta gccatcctta ctgcgcttcg attgtgtgcg tactg       1135

```

&lt;210&gt; 9

&lt;211&gt; 1096

&lt;212&gt; DNA

&lt;213&gt; CORONAVIRUS

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (137)..(958)

&lt;223&gt;

&lt;400&gt; 9

EP 1 694 829 B1

	tcttgctttg ttgcatgact agttgttgca gttgcctcaa ggggtgcatgc tcttgtgggt	60
	cttgctgcaa gtttgatgag gatgactctg agccagttct caagggtgtc aaattacatt	120
5	acacataaac gaactt atg gat ttg ttt atg aga ttt ttt act ctt gga tca Met Asp Leu Phe Met Arg Phe Phe Thr Leu Gly Ser 1 5 10	172
10	att act gca cag cca gta aaa att gac aat gct tct cct gca agt act Ile Thr Ala Gln Pro Val Lys Ile Asp Asn Ala Ser Pro Ala Ser Thr 15 20 25	220
	gtt cat gct aca gca acg ata ccg cta caa gcc tca ctc cct ttc gga Val His Ala Thr Ala Thr Ile Pro Leu Gln Ala Ser Leu Pro Phe Gly 30 35 40	268
15	tgg ctt gtt att ggc gtt gca ttt ctt gct gtt ttt cag agc gct acc Trp Leu Val Ile Gly Val Ala Phe Leu Ala Val Phe Gln Ser Ala Thr 45 50 55 60	316
20	aaa ata att gcg ctc aat aaa aga tgg cag cta gcc ctt tat aag ggc Lys Ile Ile Ala Leu Asn Lys Arg Trp Gln Leu Ala Leu Tyr Lys Gly 65 70 75	364
	ttc cag ttc att tgc aat tta ctg ctg cta ttt gtt acc atc tat tca Phe Gln Phe Ile Cys Asn Leu Leu Leu Phe Val Thr Ile Tyr Ser 80 85 90	412
25	cat ctt ttg ctt gtc gct gca ggt atg gag gcg caa ttt ttg tac ctc His Leu Leu Val Ala Ala Gly Met Glu Ala Gln Phe Leu Tyr Leu 95 100 105	460
30	tat gcc ttg ata tat ttt cta caa tgc atc aac gca tgt aga att att Tyr Ala Leu Ile Tyr Phe Leu Gln Cys Ile Asn Ala Cys Arg Ile Ile 110 115 120	508
	atg aga tgt tgg ctt tgt tgg aag tgc aaa tcc aag aac cca tta ctt Met Arg Cys Trp Leu Cys Trp Lys Cys Lys Ser Lys Asn Pro Leu Leu 125 130 135 140	556
35	tat gat gcc aac tac ttt gtt tgc tgg cac aca cat aac tat gac tac Tyr Asp Ala Asn Tyr Phe Val Cys Trp His Thr His Asn Tyr Asp Tyr 145 150 155	604
40	tgt ata cca tat aac agt gtc aca gat aca att gtc gtt act gaa ggt Cys Ile Pro Tyr Asn Ser Val Thr Asp Thr Ile Val Val Thr Glu Gly 160 165 170	652
	gac ggc att tca aca cca aaa ctc aaa gaa gac tac caa att ggt ggt Asp Gly Ile Ser Thr Pro Lys Leu Lys Glu Asp Tyr Gln Ile Gly Gly 175 180 185	700
45	tat tct gag gat agg cac tca ggt gtt aaa gac tat gtc gtt gta cat Tyr Ser Glu Asp Arg His Ser Ser Gly Val Lys Asp Tyr Val Val Val His 190 195 200	748
50	ggc tat ttc acc gaa gtt tac tac cag ctt gag tct aca caa att act Gly Tyr Phe Thr Glu Val Tyr Tyr Gln Leu Glu Ser Thr Gln Ile Thr 205 210 215 220	796
	aca gac act ggt att gaa aat gct aca ttc ttc atc ttt aac aag ctt	844

55

EP 1 694 829 B1

Thr Asp Thr Gly Ile Glu Asn Ala Thr Phe Phe Ile Phe Asn Lys Leu  
225 230 235

5 gtt aaa gac cca ccg aat gtg caa ata cac aca atc gac ggc tct tca 892  
Val Lys Asp Pro Pro Asn Val Gln Ile His Thr Ile Asp Gly Ser Ser  
240 245 250

gga gtt gct aat cca gca atg gat cca att tat gat gag ccg acg acg 940  
Gly Val Ala Asn Pro Ala Met Asp Pro Ile Tyr Asp Glu Pro Thr Thr  
255 260 265

10 act act agc gtg cct ttg taagcacaag aaagtgagta cgaacttatg 988  
Thr Thr Ser Val Pro Leu  
270

15 tactcattcg tttcgggaaga aacaggtagc ttaatagtta atagcgtact tcttttttctt 1048  
gcttttcgtgg tattcttgct agtcacacta gccatcctta ctgctgctt 1096

<210> 10  
20 <211> 274  
<212> PRT  
<213> CORONAVIRUS

<400> 10  
25

Met Asp Leu Phe Met Arg Phe Phe Thr Leu Gly Ser Ile Thr Ala Gln  
1 5 10 15

30 Pro Val Lys Ile Asp Asn Ala Ser Pro Ala Ser Thr Val His Ala Thr  
20 25 30

Ala Thr Ile Pro Leu Gln Ala Ser Leu Pro Phe Gly Trp Leu Val Ile  
35 35 40 45

Gly Val Ala Phe Leu Ala Val Phe Gln Ser Ala Thr Lys Ile Ile Ala  
50 55 60

40 Leu Asn Lys Arg Trp Gln Leu Ala Leu Tyr Lys Gly Phe Gln Phe Ile  
65 70 75 80

Cys Asn Leu Leu Leu Leu Phe Val Thr Ile Tyr Ser His Leu Leu Leu  
45 85 90 95

Val Ala Ala Gly Met Glu Ala Gln Phe Leu Tyr Leu Tyr Ala Leu Ile  
100 105 110

50 Tyr Phe Leu Gln Cys Ile Asn Ala Cys Arg Ile Ile Met Arg Cys Trp  
115 120 125

Leu Cys Trp Lys Cys Lys Ser Lys Asn Pro Leu Leu Tyr Asp Ala Asn  
55 130 135 140



EP 1 694 829 B1

Tyr Phe Val Cys Trp His Thr His Asn Tyr Asp Tyr Cys Ile Pro Tyr  
 145 150 155 160  
 5 Asn Ser Val Thr Asp Thr Ile Val Val Thr Glu Gly Asp Gly Ile Ser  
 165 170 175  
 Thr Pro Lys Leu Lys Glu Asp Tyr Gln Ile Gly Gly Tyr Ser Glu Asp  
 10 180 185 190  
 Arg His Ser Gly Val Lys Asp Tyr Val Val Val His Gly Tyr Phe Thr  
 195 200 205  
 15 Glu Val Tyr Tyr Gln Leu Glu Ser Thr Gln Ile Thr Thr Asp Thr Gly  
 210 215 220  
 Ile Glu Asn Ala Thr Phe Phe Ile Phe Asn Lys Leu Val Lys Asp Pro  
 20 225 230 235 240  
 Pro Asn Val Gln Ile His Thr Ile Asp Gly Ser Ser Gly Val Ala Asn  
 245 250 255  
 25 Pro Ala Met Asp Pro Ile Tyr Asp Glu Pro Thr Thr Thr Thr Ser Val  
 260 265 270  
 Pro Leu  
 30  
 <210> 11  
 <211> 1096  
 <212> DNA  
 <213> CORONAVIRUS  
 35  
 <220>  
 <221> CDS  
 <222> (558)..(1019)  
 <223>  
 40  
 <400> 11  
 tcttgctttg ttgcatgact agttgttgca gttgcctcaa ggggtgcatgc tcttgtggtt 60  
 45 cttgctgcaa gtttgatgag gatgactctg agccagttct caaggggtgtc aaattacatt 120  
 acacataaac gaacttatgg atttgtttat gagatTTTTT actcttggat caattactgc 180  
 acagccagta aaaattgaca atgcttctcc tgcaagtact gttcatgcta cagcaacgat 240  
 50 accgctacaa gcctcactcc ctttcggatg gcttgttatt ggcgttgcat ttcttgctgt 300  
 ttttcagagc gctacaaaaa taattgcgct caataaaaga tggcagctag ccctttataa 360  
 55

	gggcttccag ttcatttgca atttactgct gctatttggt accatctatt cacatctttt	420
	gcttgtcgct gcaggatagg aggcgcaatt tttgtacctc tatgccttga tatattttct	480
5	acaatgcata aacgcata gaattattat gagatgttgg ctttgttggga agtgcaaatc	540
	caagaaccca ttacttt atg atg cca act act ttg ttt gct ggc aca cac Met Met Pro Thr Thr Leu Phe Ala Gly Thr His 1 5 10	590
10	ata act atg act act gta tac cat ata aca gtg tca cag ata caa ttg Ile Thr Met Thr Thr Val Tyr His Ile Thr Val Ser Gln Ile Gln Leu 15 20 25	638
	tcg tta ctg aag gtg acg gca ttt caa cac caa aac tca aag aag act Ser Leu Leu Lys Val Thr Ala Phe Gln His Gln Asn Ser Lys Lys Thr 30 35 40	686
15	acc aaa ttg gtg gtt att ctg agg ata ggc act cag gtg tta aag act Thr Lys Leu Val Val Ile Leu Arg Ile Gly Thr Gln Val Leu Lys Thr 45 50 55	734
20	atg tcg ttg tac atg gct att tca ccg aag ttt act acc agc ttg agt Met Ser Leu Tyr Met Ala Ile Ser Pro Lys Phe Thr Thr Ser Leu Ser 60 65 70 75	782
	cta cac aaa tta cta cag aca ctg gta ttg aaa atg cta cat tct tca Leu His Lys Leu Leu Gln Thr Leu Val Leu Lys Met Leu His Ser Ser 80 85 90	830
25	tct tta aca agc ttg tta aag acc cac cga atg tgc aaa tac aca caa Ser Leu Thr Ser Leu Leu Lys Thr His Arg Met Cys Lys Tyr Thr Gln 95 100 105	878
30	tcg acg gct ctt cag gag ttg cta atc cag caa tgg atc caa ttt atg Ser Thr Ala Leu Gln Glu Leu Leu Ile Gln Gln Trp Ile Gln Phe Met 110 115 120	926
	atg agc cga cga cga cta cta gcg tgc ctt tgt aag cac aag aaa gtg Met Ser Arg Arg Arg Leu Leu Ala Cys Leu Cys Lys His Lys Lys Val 125 130 135	974
35	agt acg aac tta tgt act cat tcg ttt cgg aag aaa cag gta cgt Ser Thr Asn Leu Cys Thr His Ser Phe Arg Lys Lys Gln Val Arg 140 145 150	1019
40	taatagttaa tagcgtactt ctttttcttg ctttcgtggt attcttgcta gtcacactag	1079
	ccatccttac tgcgctt	1096
45	<210> 12 <211> 154 <212> PRT <213> CORONAVIRUS	
50	<400> 12	
	Met Met Pro Thr Thr Leu Phe Ala Gly Thr His Ile Thr Met Thr Thr 1 5 10 15	
55		

EP 1 694 829 B1

val Tyr His Ile Thr Val Ser Gln Ile Gln Leu Ser Leu Leu Lys Val  
 20 25 30  
 5 Thr Ala Phe Gln His Gln Asn Ser Lys Lys Thr Thr Lys Leu Val Val  
 35 40 45  
 10 Ile Leu Arg Ile Gly Thr Gln Val Leu Lys Thr Met Ser Leu Tyr Met  
 50 55 60  
 Ala Ile Ser Pro Lys Phe Thr Thr Ser Leu Ser Leu His Lys Leu Leu  
 65 70 75 80  
 15 Gln Thr Leu Val Leu Lys Met Leu His Ser Ser Ser Leu Thr Ser Leu  
 85 90 95  
 20 Leu Lys Thr His Arg Met Cys Lys Tyr Thr Gln Ser Thr Ala Leu Gln  
 100 105 110  
 Glu Leu Leu Ile Gln Gln Trp Ile Gln Phe Met Met Ser Arg Arg Arg  
 115 120 125  
 25 Leu Leu Ala Cys Leu Cys Lys His Lys Lys Val Ser Thr Asn Leu Cys  
 130 135 140  
 30 Thr His Ser Phe Arg Lys Lys Gln Val Arg  
 145 150  
 <210> 13  
 <211> 332  
 <212> DNA  
 35 <213> CORONAVIRUS  
 <220>  
 <221> CDS  
 <222> (36)..(263)  
 40 <223>  
 <400> 13  
 45 tgcctttgta agcacaagaa agtgagtacg aactt atg tac tca ttc gtt tcg 53  
 Met Tyr Ser Phe Val Ser  
 1 5  
 gaa gaa aca ggt acg tta ata gtt aat agc gta ctt ctt ttt ctt gct 101  
 Glu Glu Thr Gly Thr Leu Ile Val Asn Ser Val Leu Leu Phe Leu Ala  
 10 15 20  
 50 ttc gtg gta ttc ttg cta gtc aca cta gcc atc ctt act gcg ctt cga 149  
 Phe Val Val Phe Leu Leu Val Thr Leu Ala Ile Leu Thr Ala Leu Arg  
 25 30 35  
 55

**EP 1 694 829 B1**

[illegible]

<210> 16  
 <211> 708  
 <212> DNA  
 <213> CORONAVIRUS

<220>  
 <221> CDS  
 <222> (41)..(703)  
 <223>

<400> 16

5	tattattatt attctgtttg gaactttaac attgcttatac	atg gca gac aac ggt	55
15		Met Ala Asp Asn Gly 1 5	
	act att acc gtt gag gag ctt aaa caa ctc ctg gaa caa tgg aac cta		103
	Thr Ile Thr Val Glu Glu Leu Lys Gln Leu Leu Glu Gln Trp Asn Leu		
		10 15 20	
20	gta ata ggt ttc cta ttc cta gcc tgg att atg tta cta caa ttt gcc		151
	Val Ile Gly Phe Leu Phe Leu Ala Trp Ile Met Leu Leu Gln Phe Ala		
		25 30 35	
25	tat tct aat cgg aac agg ttt ttg tac ata ata aag ctt gtt ttc ctc		199
	Tyr Ser Asn Arg Asn Arg Phe Leu Tyr Ile Ile Lys Leu Val Phe Leu		
		40 45 50	
	tgg ctc ttg tgg cca gta aca ctt gct tgt ttt gtg ctt gct gct gtc		247
	Trp Leu Leu Trp Pro Val Thr Leu Ala Cys Phe Val Leu Ala Ala Val		
		55 60 65	
30	tac aga att aat tgg gtg act ggc ggg att gcg att gca atg gct tgt		295
	Tyr Arg Ile Asn Trp Val Thr Gly Gly Ile Ala Ile Ala Met Ala Cys		
		70 75 80 85	
35	att gta ggc ttg atg tgg ctt agc tac ttc gtt gct tcc ttc agg ctg		343
	Ile Val Gly Leu Met Trp Leu Ser Tyr Phe Val Ala Ser Phe Arg Leu		
		90 95 100	
	ttt gct cgt acc cgc tca atg tgg tca ttc aac cca gaa aca aac att		391
	Phe Ala Arg Thr Arg Ser Met Trp Ser Phe Asn Pro Glu Thr Asn Ile		
		105 110 115	
40	ctt ctc aat gtg cct ctc cgg ggg aca att gtg acc aga ccg ctc atg		439
	Leu Leu Asn Val Pro Leu Arg Gly Thr Ile Val Thr Arg Pro Leu Met		
		120 125 130	
45	gaa agt gaa ctt gtc att ggt gct gtg atc att cgt ggt cac ttg cga		487
	Glu Ser Glu Leu Val Ile Gly Ala Val Ile Ile Arg Gly His Leu Arg		
		135 140 145	
50	atg gcc gga cac tcc cta ggg cgc tgt gac att aag gac ctg cca aaa		535
	Met Ala Gly His Ser Leu Gly Arg Cys Asp Ile Lys Asp Leu Pro Lys		
		150 155 160 165	

EP 1 694 829 B1

gag atc act gtg gct aca tca cga acg ctt tct tat tac aaa tta gga 583  
 Glu Ile Thr Val Ala Thr Ser Arg Thr Leu Ser Tyr Tyr Lys Leu Gly 170 175 180

5 gcg tcg cag cgt gta ggc act gat tca ggt ttt gct gca tac aac cgc 631  
 Ala Ser Gln Arg Val Gly Thr Asp Ser Gly Phe Ala Ala Tyr Asn Arg 185 190 195

10 tac cgt att gga aac tat aaa tta aat aca gac cac gcc ggt agc aac 679  
 Tyr Arg Ile Gly Asn Tyr Lys Leu Asn Thr Asp His Ala Gly Ser Asn 200 205 210

gac aat att gct ttg cta gta cag taagt 708  
 Asp Asn Ile Ala Leu Leu Val Gln 215 220

15 <210> 17  
 <211> 221  
 <212> PRT  
 <213> CORONAVIRUS

20 <400> 17

25 Met Ala Asp Asn Gly Thr Ile Thr Val Glu Glu Leu Lys Gln Leu Leu 1 5 10 15

Glu Gln Trp Asn Leu Val Ile Gly Phe Leu Phe Leu Ala Trp Ile Met 20 25 30

30 Leu Leu Gln Phe Ala Tyr Ser Asn Arg Asn Arg Phe Leu Tyr Ile Ile 35 40 45

35 Lys Leu Val Phe Leu Trp Leu Leu Trp Pro Val Thr Leu Ala Cys Phe 50 55 60

Val Leu Ala Ala Val Tyr Arg Ile Asn Trp Val Thr Gly Gly Ile Ala 65 70 75 80

40 Ile Ala Met Ala Cys Ile Val Gly Leu Met Trp Leu Ser Tyr Phe Val 85 90 95

45 Ala Ser Phe Arg Leu Phe Ala Arg Thr Arg Ser Met Trp Ser Phe Asn 100 105 110

Pro Glu Thr Asn Ile Leu Leu Asn Val Pro Leu Arg Gly Thr Ile Val 115 120 125

50 Thr Arg Pro Leu Met Glu Ser Glu Leu Val Ile Gly Ala Val Ile Ile 130 135 140

55 Arg Gly His Leu Arg Met Ala Gly His Ser Leu Gly Arg Cys Asp Ile 145 150 155 160

EP 1 694 829 B1

Lys Asp Leu Pro Lys Glu Ile Thr Val Ala Thr Ser Arg Thr Leu Ser  
165 170 175

Tyr Tyr Lys Leu Gly Ala Ser Gln Arg Val Gly Thr Asp Ser Gly Phe  
180 185 190

Ala Ala Tyr Asn Arg Tyr Arg Ile Gly Asn Tyr Lys Leu Asn Thr Asp  
195 200 205

His Ala Gly Ser Asn Asp Asn Ile Ala Leu Leu Val Gln  
210 215 220

<210> 18

<211> 769

<212> DNA

<213> CORONAVIRUS

<400> 18

cctgatcttc tggctctaac gaactaacta ttattattat tctgtttgga actttaacat 60  
tgcttatcat ggcagacaac ggtactatta ccgttgagga gcttaaaca ctcctggaac 120  
aatggaacct agtaataggt ttcctattcc tagcctggat tatgttacta caatttgcct 180  
attctaactg gaacagggtt ttgtacataa taaagcttgt tttcctctgg ctcttggtggc 240  
cagtaacact tgcttggttt gtgcttgctg ctgtctacag aattaattgg gtgactggcg 300  
ggattgcatg tgcaatggct tgtattgtag gcttgatgtg gcttagctac ttcggttgctt 360  
ccttcaggct gtttgctcgt acccgctcaa tgtggtcatt caaccagaa acaaacattc 420  
ttctcaatgt gcctctccgg gggacaattg tgaccagacc gctcatggaa agtgaacttg 480  
tcattggtgc tgtgatcatt cgtggtcact tgcgaatggc cggacactcc ctagggcgct 540  
gtgacattaa ggacctgcca aaagagatca ctgtggctac atcacgaacg ctttcttatt 600  
acaaattagg agcgtcgcag cgtgtaggca ctgattcagg ttttgctgca tacaaccgct 660  
accgtattgg aaactataaa ttaaatacag accacgccgg tagcaacgac aatattgctt 720  
tgctagtaca gtaagtgaca acagatgttt catcttggtg acttccagg 769

<210> 19

<211> 1231

<212> DNA

<213> CORONAVIRUS

<400> 19

taccgtattg gaaactataa attaaatata gaccacgccg gtagcaacga caatattgct 60

	gctagtagtac agtaagtgac aacagatggt tcattcttggt gacttccagg ttacaatagc	120
	agagatattg attatcatta tgaggacttt caggattgct atttggaatc ttgacgttat	180
5	aataagttca atagtgagac aattatttaa gcctctaact aagaagaatt attcggagtt	240
	agatgatgaa gaacctatgg agttagatta tccataaaac gaacatgaaa attattctct	300
	tcctgacatt gattgtattt acatcttgcg agctatatca ctatcaggag tgtgttagag	360
10	gtacgactgt actactaaaa gaaccttgcc catcaggaac atacgagggc aattcaccat	420
	ttcacctctt tgctgacaat aaatttgac taacttgac tagcacacac tttgcttttg	480
	cttggtgctga cggtagctga catacctatc agctgcgtgc aagatcagtt tcaccaaacc	540
15	ttttcatcag acaagaggag gttcaacaag agctctactc gccacttttt ctattgttg	600
	ctgctctagt atttttaata ctttgcttca ccattaagag aaagacagaa tgaatgagct	660
	cactttaatt gacttctatt tgtgcttttt agcctttctg ctattccttg ttttaataat	720
20	gcttattata ttttggtttt cactcgaaat ccaggatcta gaagaacctt gtaccaaagt	780
	ctaaacgaac atgaaacttc tcattgtttt gacttgtatt tctctatgca gttgcatatg	840
	cactgtagta cagcgctgtg catctaataa acctcatgtg cttgaagatc cttgtaaggt	900
25	acaacactag gggtaatact tatagcactg cttggctttg tgctctagga aagggtttac	960
	cttttcatag atggcacact atgggtcaaa catgcacacc taatgttact atcaactgtc	1020
	aagatccagc tgggtggtgcg cttatagcta ggtgttggtg cttcatgaa ggtcaccaaa	1080
30	ctgctgcatt tagagacgta cttgttgttt taaataaacg aacaaattaa aatgtctgat	1140
	aatggacccc aatcaaacca acgtagtgcc ccccgcatca catttggtgg acccacagat	1200
	tcaactgaca ataaccagaa tggaggacgc a	1231
35	<210> 20	
	<211> 1242	
	<212> DNA	
	<213> CORONAVIRUS	
40	<400> 20	
	gcatacaacc gctaccgtat tggaaactat aaattaaata cagaccacgc cggtagcaac	60
45	gacaatattg ctttgctagt acagtaagtg acaacagatg tttcatcttg ttgacttcca	120
	ggttacaata gcagagatat tgattatcat tatgaggact ttcaggattg ctatttgga	180
	tcttgacgtt ataataagtt caatagtgag acagttattt aagcctctaa ctaagaagaa	240
50	ttattcggag ttagatgatg aagaacctat ggagttagat tatccataaa acgaacatga	300
	aaattattct ctccctgaca ttgattgtat ttacatcttg cgagctatat cactatcagg	360
	agtgtgttag aggtacgact gtactactaa aagaaccttg cccatcagga acatacgagg	420
55	gcaattcacc atttcaccct cttgctgaca ataaatttgc actaacttgc actagcacac	480



EP 1 694 829 B1

actttgcttt tgcttgtgct gacgggtactc gacataccta tcagctgcgt gcaagatcag 540  
 tttcaccaaa acttttcatc agacaagagg aggttcaaca agagctctac tcgccacttt 600  
 5 tttctattgt tgctgctcta gtatttttaa tactttgctt caccattaag agaaagacag 660  
 aatgaatgag ctcaacttaa ttgacttcta tttgtgcttt ttagcctttc tgctattcct 720  
 tgttttaata atgcttatta tatttttggtt ttcactcgaa atccaggatc tagaagaacc 780  
 10 ttgtaccaaa gtctaaacga acatgaaact tctcattggt ttgacttgta tttctctatg 840  
 cagttgcata tgcactgtag tacagcgctg tgcactaat aaacctcatg tgcttgaaga 900  
 tccttgtaag gtacaacact aggggtaata cttatagcac tgcttggtt tgctgcttag 960  
 15 gaaaggtttt accttttcat agatggcaca ctatggttca aacatgcaca cctaagtgtta 1020  
 ctatcaactg tcaagatcca gctgggtggtg cgcttatagc taggtgttgg taccttcatg 1080  
 aaggtcacca aactgctgca tttagagacg tacttggttg tttaaataaaa cgaacgaatt 1140  
 20 aaaatgtctg ataatggacc ccaatcaaac caacgtagtg ccccccgcac tacatttggt 1200  
 ggaccacag attcaactga caataaccag aatggaggac gc 1242

25 <210> 21  
 <211> 1231  
 <212> DNA  
 <213> CORONAVIRUS

30 <220>  
 <221> CDS  
 <222> (86)..(274)  
 <223>

35 <400> 21

taccgtattg gaaactataa attaaatata gaccacgccg gtagcaacga caatattgct 60  
 ttgctagtac agtaagtac aacag atg ttt cat ctt gtt gac ttc cag gtt 112  
 40 Met Phe His Leu Val Asp Phe Gln Val  
 1 5  
 aca ata gca gag ata ttg att atc att atg agg act ttc agg att gct 160  
 Thr Ile Ala Glu Ile Leu Ile Ile Ile Met Arg Thr Phe Arg Ile Ala  
 10 15 20 25  
 45 att tgg aat ctt gac gtt ata ata agt tca ata gtg aga caa tta ttt 208  
 Ile Trp Asn Leu Asp Val Ile Ile Ser Ser Ile Val Arg Gln Leu Phe  
 30 35 40  
 50 aag cct cta act aag aag aat tat tcg gag tta gat gat gaa gaa cct 256  
 Lys Pro Leu Thr Lys Lys Asn Tyr Ser Glu Leu Asp Asp Glu Glu Pro  
 45 50 55  
 atg gag tta gat tat cca taaaacgaac atgaaaatta ttctcttctt 304

55

Met Glu Leu Asp Tyr Pro  
60

5 gacattgatt gtatttacat cttgcgagct atatcactat caggagtgtg ttagaggtac 364  
gactgtacta ctaaaagaac cttgcccac aggaacatac gagggcaatt caccatttca 424  
ccctcttgct gacaataaat ttgcactaac ttgcactagc acacactttg cttttgcttg 484  
10 tgctgacggt actcgacata cctatcagct gcgtgcaaga tcagtttcac caaaactttt 544  
catcagacaa gaggagggtt aacaagagct ctactcgcca ctttttctca ttgttgctgc 604  
tctagtattt ttaatacttt gcttcacat taagagaaag acagaatgaa tgagctcact 664  
15 ttaattgact tctatttgct ctttttagcc tttctgctat tccttgtttt aataatgctt 724  
attatatattt ggttttcact cgaaatccag gatctagaag aaccttgtag caaagtctaa 784  
acgaacatga aacttctcat tgttttgact tgtatttctc tatgcagttg catatgcact 844  
20 gtagtacagc gctgtgcatc taataaacct catgtgcttg aagatccttg taaggtagaa 904  
cactaggggt aatacttata gcactgcttg gctttgtgct ctaggaaagg ttttaccttt 964  
tcatagatgg cacactatgg ttcaaacaat cacacctaatt gttactatca actgtcaaga 1024  
25 tccagctggt ggtgcgctta tagctaggtg ttggtacctt catgaagggt accaaactgc 1084  
tgcatttaga gacgtacttg ttgttttaaa taaacgaaca aattaaaatg tctgataatg 1144  
gacccaatc aaaccaacgt agtgccccc gcattacatt tggtaggacc acagattcaa 1204  
30 ctgacaataa ccagaatgga ggacgca 1231

<210> 22

<211> 63

<212> PRT

35 <213> CORONAVIRUS

<400> 22

40 Met Phe His Leu Val Asp Phe Gln Val Thr Ile Ala Glu Ile Leu Ile  
1 5 10 15

45 Ile Ile Met Arg Thr Phe Arg Ile Ala Ile Trp Asn Leu Asp Val Ile  
20 25 30

Ile Ser Ser Ile Val Arg Gln Leu Phe Lys Pro Leu Thr Lys Lys Asn  
35 40 45

50 Tyr Ser Glu Leu Asp Asp Glu Glu Pro Met Glu Leu Asp Tyr Pro  
50 55 60

<210> 23

55 <211> 1231

<212> DNA

<213> CORONAVIRUS

<220>  
 <221> CDS  
 <222> (285)..(650)  
 <223>

5 <400> 23

10 taccgtattg gaaactataa attaaataca gaccacgccg gtagcaacga caatattgct 60

10 ttgctagtac agtaagtac aacagatggt tcatcttggt gacttccagg ttacaatagc 120

agagatatgg attatcatta tgaggacttt caggattgct atttggaatc ttgacgttat 180

aataagttca atagtggagac aattatttaa gcctctaact aagaagaatt attcggagtt 240

15 agatgatgaa gaacctatgg agttagatta tccataaaac gaac atg aaa att att 296  
 Met Lys Ile Ile  
 1

20 ctc ttc ctg aca ttg att gta ttt aca tct tgc gag cta tat cac tat 344  
 Leu Phe Leu Thr Leu Ile Val Phe Thr Ser Cys Glu Leu Tyr His Tyr  
 5 10 15 20

cag gag tgt gtt aga ggt acg act gta cta cta aaa gaa cct tgc cca 392  
 Gln Glu Cys Val Arg Gly Thr Thr Val Leu Leu Lys Glu Pro Cys Pro  
 25

25 tca gga aca tac gag ggc aat tca cca ttt cac cct ctt gct gac aat 440  
 Ser Gly Thr Tyr Glu Gly Asn Ser Pro Phe His Pro Leu Ala Asp Asn  
 40 45 50

30 aaa ttt gca cta act tgc act agc aca cac ttt gct ttt gct tgt gct 488  
 Lys Phe Ala Leu Thr Cys Thr Ser Thr His Phe Ala Phe Ala Cys Ala  
 55 60 65

gac ggt act cga cat acc tat cag ctg cgt gca aga tca gtt tca cca 536  
 Asp Gly Thr Arg His Thr Tyr Gln Leu Arg Ala Arg Ser Val Ser Pro  
 70 75 80

35 aaa ctt ttc atc aga caa gag gag gtt caa caa gag ctc tac tcg cca 584  
 Lys Leu Phe Ile Arg Gln Glu Glu Val Gln Gln Glu Leu Tyr Ser Pro  
 85 90 95 100

40 ctt ttt ctc att gtt gct gct cta gta ttt tta ata ctt tgc ttc acc 632  
 Leu Phe Leu Ile Val Ala Ala Leu Val Phe Leu Ile Leu Cys Phe Thr  
 105 110 115

att aag aga aag aca gaa tgaatgagct cactttaatt gacttctatt 680  
 Ile Lys Arg Lys Thr Glu  
 120

45 tgtgcttttt agcctttctg ctattccttg ttttaataat gcttattata ttttggtttt 740

cactcgaaat ccaggatcta gaagaacctt gtaccaaagt cttaaactgaac atgaaacttc 800

50 tcattgtttt gacttgtatt tctctatgca gttgcatatg cactgtagta cagcgctgtg 860

catctaataa acctcatgtg cttgaagatc cttgtaaggt acaacactag gggttaatact 920

tatagcactg cttggctttg tgctctagga aaggttttac cttttcatag atggcacact 980  
 atgggttcaaa catgcacacc taatgttact atcaactgtc aagatccagc tgggtggtgcg 1040  
 5 cttatagcta ggtgttggtgta ctttcatgaa ggtcacacaaa ctgctgcatt tagagacgta 1100  
 cttgttggtt taaataaacg aacaaattaa aatgtctgat aatggacccc aatcaaacca 1160  
 acgtagtgcc ccccgcattha catttggtgg acccacagat tcaactgaca ataaccagaa 1220  
 10 tggaggacgc a 1231

<210> 24

<211> 122

15 <212> PRT

<213> CORONAVIRUS

<400> 24

20

Met Lys Ile Ile Leu Phe Leu Thr Leu Ile Val Phe Thr Ser Cys Glu  
 1 5 10 15

25

Leu Tyr His Tyr Gln Glu Cys Val Arg Gly Thr Thr Val Leu Leu Lys  
 20 25 30

Glu Pro Cys Pro Ser Gly Thr Tyr Glu Gly Asn Ser Pro Phe His Pro  
 35 40 45

30

Leu Ala Asp Asn Lys Phe Ala Leu Thr Cys Thr Ser Thr His Phe Ala  
 50 55 60

35

Phe Ala Cys Ala Asp Gly Thr Arg His Thr Tyr Gln Leu Arg Ala Arg  
 65 70 75 80

Ser Val Ser Pro Lys Leu Phe Ile Arg Gln Glu Glu Val Gln Gln Glu  
 85 90 95

40

Leu Tyr Ser Pro Leu Phe Leu Ile Val Ala Ala Leu Val Phe Leu Ile  
 100 105 110

45

Leu Cys Phe Thr Ile Lys Arg Lys Thr Glu  
 115 120

<210> 25

<211> 1231

50 <212> DNA

<213> CORONAVIRUS

<220>

<221> CDS

55 <222> (650)..(781)

<223>

<400> 25

EP 1 694 829 B1

	taccgtattg gaaactataa attaaataca gaccacgccg gtagcaacga caatattgct	60
	ttgctagtac agtaagtgac aacagatggt tcattcttggt gacttccagg ttacaatagc	120
5	agagatattg attatcatta tgaggacttt caggattgct atttggaatc ttgacgttat	180
	aataagttca atagtgagac aattatttaa gcctctaact aagaagaatt attcggagtt	240
	agatgatgaa gaacctatgg agttagatta tccataaaac gaacatgaaa attattctct	300
10	tcctgacatt gattgtattt acatcttgcg agctatatca ctatcaggag tgtgttagag	360
	gtacgactgt actactaaaa gaaccttgcc catcaggaac atacgagggc aattcaccat	420
	ttcaccctct tgctgacaat aaatttgcac taacttgcac tagcacacac tttgcttttg	480
15	cttgtgctga cgggtactcga catacctatc agctgcgtgc aagatcagtt tcaccaaacc	540
	ttttcatcag acaagaggag gttcaacaag agctctactc gccacttttt ctcatgttg	600
	ctgctctagt atttttaata ctttgcttca ccattaagag aaagacaga atg aat gag	658
20		Met Asn Glu 1
	ctc act tta att gac ttc tat ttg tgc ttt tta gcc ttt ctg cta ttc	706
	Leu Thr Leu Ile Asp Phe Tyr Leu Cys Phe Leu Ala Phe Leu Leu Phe	
	5 10 15	
25	ctt gtt tta ata atg ctt att ata ttt tgg ttt tca ctc gaa atc cag	754
	Leu Val Leu Ile Met Leu Ile Ile Phe Trp Phe Ser Leu Glu Ile Gln	
	20 25 30 35	
	gat cta gaa gaa cct tgt acc aaa gtc taaacgaaca tgaaacttct	801
30	Asp Leu Glu Glu Pro Cys Thr Lys Val	
	40	
	cattgttttg acttgtattt ctctatgcag ttgcatatgc actgtagtac agcgctgtgc	861
	atctaataaa cctcatgtgc ttgaagatcc ttgtaaggta caacactagg ggtaataactt	921
35	atagcactgc ttggctttgt gctctaggaa aggttttacc ttttcataga tggcacacta	981
	tggttcaaac atgcacacct aatgttacta tcaactgtca agatccagct ggtggtgcgc	1041
	ttatagctag gtgttggtac cttcatgaag gtcaccaaac tgctgcattt agagacgtac	1101
40	ttgttgtttt aaataaacga acaaattaaa atgtctgata atggacccca atcaaaccaa	1161
	cgtagtgccc cccgcattac atttggtgga cccacagatt caactgacaa taaccagaat	1221
	ggaggacgca	1231
45		
	<210> 26	
	<211> 44	
	<212> PRT	
50	<213> CORONAVIRUS	
	<400> 26	
55		

# EP 1 694 829 B1

Met Asn Glu Leu Thr Leu Ile Asp Phe Tyr Leu Cys Phe Leu Ala Phe  
1 5 10 15

Leu Leu Phe Leu Val Leu Ile Met Leu Ile Ile Phe Trp Phe Ser Leu  
20 25 30

Glu Ile Gln Asp Leu Glu Glu Pro Cys Thr Lys Val  
35 40

<210> 27

<211> 1231

<212> DNA

<213> CORONAVIRUS

<220>

<221> CDS

<222> (791)..(907)

<223>

<400> 27

taccgtattg	gaaactataa	attaaatata	gaccacgccg	gtagcaacga	caatattgct	60
ttgctagtag	agtaagtgac	aacagatggt	tcattcttgt	gacttccagg	ttacaatagc	120
agagatattg	attatcatta	tgaggacttt	caggattgct	atttggaatc	ttgacgttat	180
aataagttca	atagtgagac	aattatttaa	gcctctaact	aagaagaatt	attcggagtt	240
agatgatgaa	gaacctatgg	agttagatta	tccataaaac	gaacatgaaa	attattctct	300
tcctgacatt	gattgtattt	acattcttgc	agctatatca	ctatcaggag	tgtgttagag	360
gtacgactgt	actactaaaa	gaaccttgcc	catcaggaac	atacgagggc	aattcaccat	420
ttcacctctt	tgctgacaat	aaatttgcac	taacttgcac	tagcacacac	tttgcttttg	480
cttgctgctga	cggtactcga	catacctatc	agctgcgtgc	aagatcagtt	tcaccaaaac	540
ttttcatcag	acaagaggag	gttcaacaag	agctctactc	gccacttttt	ctcattgttg	600
ctgctctagt	atttttaata	ctttgcttca	ccattaagag	aaagacagaa	tgaatgagct	660
cactttaatt	gacttctatt	tgtgcttttt	agcctttctg	ctattccttg	ttttaataat	720
gcttattata	ttttggtttt	cactcgaaat	ccaggatcta	gaagaacctt	gtaccaaagt	780
ctaaacgaac	atg aaa ctt	ctc att gtt	ttg act tgt	att tct cta	tgc	829

EP 1 694 829 B1

Met Lys Leu Leu Ile Val Leu Thr Cys Ile Ser Leu Cys  
1 5 10

5 agt tgc ata tgc act gta gta cag cgc tgt gca tct aat aaa cct cat 877  
Ser Cys Ile Cys Thr Val Val Gln Arg Cys Ala Ser Asn Lys Pro His  
15 20 25

gtg ctt gaa gat cct tgt aag gta caa cac taggggtaat acttatagca 927  
Val Leu Glu Asp Pro Cys Lys Val Gln His  
30

10 ctgcttggct ttgtgctcta ggaaagggtt taccttttca tagatggcac actatgggttc 987  
aaacatgcac acctaattgtt actatcaact gtcaagatcc agctgggtggt gcgcttatag 1047  
ctagggtgttg gtaccttcat gaagggtcacc aaactgctgc atttagagac gtacttgttg 1107  
15 ttttaaataa acgaacaaat taaaatgtct gataatggac cccaatcaaa ccaacgtagt 1167  
gccccccgca ttacatttgg tggacccaca gattcaactg acaataacca gaatggagga 1227  
cgca 1231

20

<210> 28  
<211> 39  
<212> PRT  
25 <213> CORONAVIRUS  
<400> 28

30 Met Lys Leu Leu Ile Val Leu Thr Cys Ile Ser Leu Cys Ser Cys Ile  
1 5 10 15

Cys Thr Val Val Gln Arg Cys Ala Ser Asn Lys Pro His Val Leu Glu  
20 25 30

35 Asp Pro Cys Lys Val Gln His  
35

40 <210> 29  
<211> 1231  
<212> DNA  
<213> CORONAVIRUS

45 <220>  
<221> CDS  
<222> (876)..(1127)  
<223>

50 <400> 29

55

EP 1 694 829 B1

	taccgtattg gaaactataa attaaataca gaccacgccg gtagcaacga caatattgct	60
	ttgctagtag agtaagtac aacagatggt tcatcttggt gacttccagg ttacaatagc	120
5	agagatattg attatcatta tgaggacttt caggattgct atttggaatc ttgacggtat	180
	aataagttca atagttagac aattatttaa gcctctaact aagaagaatt attcggaggt	240
	agatgatgaa gaacctatgg agttagatta tccataaaac gaacatgaaa attattctct	300
10	tcctgacatt gattgtattt acatcttgcg agctatatca ctatcaggag tgtgttagag	360
	gtacgactgt actactaaaa gaaccttgcc catcaggaac atacgagggc aattcaccat	420
	ttcacctctt tgctgacaat aaatttgcac taacttgcac tagcacacac tttgcttttg	480
15	cttggtgctga cggtagctga catacctatc agctgcgtgc aagatcagtt tcacaaaaac	540
	ttttcatcag acaagaggag gttcaacaag agctctactc gccacttttt ctcatgttg	600
	ctgctctagt atttttaata ctttgcttca ccattaagag aaagacagaa tgaatgagct	660
20	cactttaatt gacttctatt tgtgcttttt agcctttctg ctattccttg ttttaataat	720
	gcttattata ttttggtttt cactcgaaat ccaggatcta gaagaacctt gtaccaaagt	780
	ctaaacgaac atgaaacttc tcattgtttt gacttgtatt tctctatgca gttgcatatg	840
25	cactgtagta cagcgtgtg catctaataa acctc atg tgc ttg aag atc ctt Met Cys Leu Lys Ile Leu 1 5	893
	gta agg tac aac act agg ggt aat act tat agc act gct tgg ctt tgt Val Arg Tyr Asn Thr Arg Gly Asn Thr Tyr Ser Thr Ala Trp Leu Cys 10 15 20	941
30	gct cta gga aag gtt tta cct ttt cat aga tgg cac act atg gtt caa Ala Leu Gly Lys Val Leu Pro Phe His Arg Trp His Thr Met Val Gln 25 30 35	989
35	aca tgc aca cct aat gtt act atc aac tgt caa gat cca gct ggt ggt Thr Cys Thr Pro Asn Val Thr Ile Asn Cys Gln Asp Pro Ala Gly Gly 40 45 50	1037
40	gcg ctt ata gct agg tgt tgg tac ctt cat gaa ggt cac caa act gct Ala Leu Ile Ala Arg Cys Trp Tyr Leu His Glu Gly His Gln Thr Ala 55 60 65 70	1085
	gca ttt aga gac gta ctt gtt gtt tta aat aaa cga aca aat Ala Phe Arg Asp Val Leu Val Val Leu Asn Lys Arg Thr Asn 75 80	1127
45	taaaatgtct gataatggac cccaatcaaa ccaacgtagt gcccccgca ttacatttg	1187
	tggaccacaca gattcaactg acaataacca gaatggagga cgca	1231
50	<210> 30 <211> 84 <212> PRT <213> CORONAVIRUS	
55	<400> 30	



# EP 1 694 829 B1

Met Cys Leu Lys Ile Leu Val Arg Tyr Asn Thr Arg Gly Asn Thr Tyr  
1 5 10 15

5 ser Thr Ala Trp Leu Cys Ala Leu Gly Lys Val Leu Pro Phe His Arg  
20 25 30

Trp His Thr Met Val Gln Thr Cys Thr Pro Asn Val Thr Ile Asn Cys  
10 35 40 45

Gln Asp Pro Ala Gly Gly Ala Leu Ile Ala Arg Cys Trp Tyr Leu His  
50 55 60

15 Glu Gly His Gln Thr Ala Ala Phe Arg Asp Val Leu Val Val Leu Asn  
65 70 75 80

Lys Arg Thr Asn

20 <210> 31  
<211> 21221  
<212> DNA  
<213> CORONAVIRUS

25 <400> 31

atggagagcc ttgttcttgg tgtcaacgag aaaacacacg tccaactcag tttgcctgtc 60  
30 cttcagggtta gagacgtgct agtgcgtggc ttcggggact ctgtggaaga ggccttatcg 120  
gaggcacgtg aacacctcaa aaatggcact tgtgggtctag tagagctgga aaaaggcgta 180  
ctgccccagc ttgaacagcc ctatgtgttc attaaacggt ctgatgcctt aagcaccaat 240  
35 cacggccaca aggtcgttga gctgggttgca gaaatggacg gcattcagta cggtcgtagc 300  
gggtataacac tgggagtact cgtgccacat gtgggcgaaa cccaattgc ataccgcaat 360  
gttctttcttc gtaagaacgg taataaggga gccgggtggc atagctatgg catcgatcta 420  
40 aagtcttatg acttaggtga cgagcttggc actgatcca ttgaagatta tgaacaaaac 480  
tggaacacta agcatggcag tgggtgcactc cgtgaactca ctctgtgagct caatggaggt 540  
gcagtcactc gctatgtcga caacaatttc tgtggcccag atgggtaccc tcttgattgc 600  
45 atcaaagatt ttctcgacg cgcgggcaag tcaatgtgca ctctttccga acaacttgat 660  
tacatcgagt cgaagagagg tgtctactgc tgccgtgacc atgagcatga aattgcctgg 720  
ttcactgagc gctctgataa gagctacgag caccagacac ccttcgaaat taagagtgcc 780  
50 aagaaatttg acactttcaa aggggaatgc ccaaagtttg tgtttcctct taactcaaaa 840  
gtcaaagtca ttcaaccacg tgttgaaaag aaaaagactg aggggtttcat ggggcgtata 900  
cgctctgtgt accctgttgc atctccacag gagtgtaaca atatgcactt gtctaccttg 960  
55

EP 1 694 829 B1

	atgaaatgta atcattgcga tgaagtttca tggcagacgt gcgactttct gaaagccact	1020
	tgtgaacatt gtggcactga aaatttagtt attgaaggac ctactacatg tgggtaccta	1080
5	cctactaatg ctgtagtgaa aatgccatgt cctgcctgtc aagaccaga gattggacct	1140
	gagcatagtg ttgcagatta tcacaaccac tcaaacattg aaactcgact ccgcaaggga	1200
	ggtaggacta gatgttttgg aggctgtgtg tttgcctatg ttggctgcta taataagcgt	1260
10	gcctactggg ttcctcgtgc tagtgctgat attggctcag gccatactgg cattactggg	1320
	gacaatgtgg agaccttgaa tgaggatctc cttgagatac tgagtcgtga acgtgttaac	1380
	attaacattg ttggcgattt tcatttgaat gaagaggttg ccatcatttt ggcatctttc	1440
15	tctgcttcta caagtgcctt tattgacact ataaagagtc ttgattacaa gtctttcaaa	1500
	accattgttg agtcctgcgg taactataaa gttaccaagg gaaagccgt aaaagggtgct	1560
	tggaacattg gacaacagag atcagtttta acaccactgt gtggttttcc ctcacaggct	1620
20	gctggtgtta tcagatcaat ttttgcgcgc acacttgatg cagcaaacca ctcaattcct	1680
	gatttgcaaa gagcagctgt caccatactt gatggtatth ctgaacagtc attacgtctt	1740
	gtcgacgcca tgggtttatac ttcagacctg ctcaccaaca gtgtcattat tatggcatat	1800
25	gtaactgggtg gtcttgtaca acagacttct cagtgggtgt ctaatctttt gggcactact	1860
	gttgaaaaac tcaggcctat ctttgaatgg attgaggcga aacttagtgc aggagttgaa	1920
	tttctcaagg atgcttgga gattctcaaa tttctcatta cagggtgtttt tgacatcgtc	1980
30	aagggtcaaa tacagggtgc ttcagataac atcaaggatt gtgtaaaatg cttcattgat	2040
	gttggttaaca aggcactcga aatgtgcatt gatcaagtca ctatcgctgg cgcaaagttg	2100
	cgatcactca acttaggtga agtcttcacg gctcaaagca agggacttta ccgtcagtgt	2160
35	atacgtggca aggagcagct gcaactactc atgcctctta aggcaccaa agaagtaacc	2220
	tttcttgaag gtgattcaca tgacacagta cttacctctg aggaggttgt tctcaagaac	2280
	gggtgaactcg aagcactcga gacgcccgtt gatagcttca caaatggagc tatcgttggc	2340
40	acaccagtct gtgtaaatgg cctcatgctc ttagagatta aggacaaaga acaatactgc	2400
	gcattgtctc ctggtttact ggctacaaac aatgtctttc gcttaaaagg ggggtgcacca	2460
	attaaagggtg taaccttttg agaagatact gtttggaag ttcaagggtta caagaatgtg	2520
45	agaatcacat ttgagcttga tgaacgtgtt gacaaagtgc ttaatgaaaa gtgctctgtc	2580
	tacactgttg aatccggtac cgaagttact gagtttgcag gtgttgtagc agaggctgtt	2640
	gtgaagactt tacaaccagt ttctgatctc cttaccaaca tgggtattga tcttgatgag	2700
50	tggagtgtag ctacattcta cttatttgat gatgctgggtg aagaaaactt ttcacacgt	2760
	atgtattgtt ctttttacct tccagatgag gaagaagagg acgatgcaga gtgtgaggaa	2820
	gaagaaattg atgaaacctg tgaacatgag tacgggtacag aggatgatta tcaaggctctc	2880
55	cctctggaat ttgggtgcctc agctgaaaca gttcgagttg aggaagaaga agaggaagac	2940
	tggctggatg atactactga gcaatcagag attgagccag aaccagaacc tacacctgaa	3000

## EP 1 694 829 B1

	gaaccagtta atcagtttac tggttattta aaacttactg acaatgttgc cattaaatgt	3060
	gttgacatcg ttaaggaggc acaaagtgc t aatcctatgg tgattgtaaa tgctgctaac	3120
5	atacacctga aacatggtgg tgggtgtagca ggtgcactca acaaggcaac caatggtgcc	3180
	atgcaaaagg agagtgatga ttacattaag ctaaatggcc ctcttacagt aggagggctct	3240
	tgtttgcttt ctggacataa tcttgctaag aagtgtctgc atgttgttgg acctaacct	3300
10	aatgcagggtg aggacatcca gcttcttaag gcagcatatg aaaatttcaa ttcacaggac	3360
	atcttacttg caccattggt gtcagcaggc atatttggtg ct aaaccact tcagtcttta	3420
	caagtgtgcg tgcagacggt tcgtacacag gtttatattg cagtcaatga caaagctctt	3480
15	tatgagcagg ttgtcatgga ttatcttgat aacctgaagc ctagagtggg agcacctaaa	3540
	caagaggagc caccaaacac agaagattcc aaaactgagg agaaatctgt cgtacagaag	3600
	cctgtcgatg tgaagccaaa aattaaggcc tgcattgatg aggttaccac aacactggaa	3660
20	gaaactaagt ttcttaccaa taagttactc ttgtttgctg atatcaatgg taagctttac	3720
	catgattctc agaacatgct tagagggtgaa gatatgtctt tccttgagaa ggatgcacct	3780
	tacatggtag gtgatgttat cactagtggg gatatcactt gtgttgtaat accctccaaa	3840
25	aaggctggtg gcactactga gatgctctca agagctttga agaaagtgcc agttgatgag	3900
	tatataacca cgtaccctgg acaaggatgt gctgggtata cacttgagga agctaagact	3960
	gctcttaaga aatgcaaadc tgcattttat gtactacctt cagaagcacc taatgctaag	4020
30	gaagagattc taggaactgt atcctggaat ttgagagaaa tgcttgctca tgctgaagag	4080
	acaagaaaat taatgcctat atgcatggat gttagagcca taatggcaac catccaacgt	4140
	aagtataaag gaattaaaat tcaagagggc atcggtgact atgggtgtccg attcttcttt	4200
35	tatactagta aagagcctgt agcttctatt attacgaagc tgaactctct aaatgagccg	4260
	cttgtcacia tgccaattgg ttatgtgaca catgggtttta atcttgaaga ggctgcgcgc	4320
	tgtatgctgt ctcttaaaagc tcctgccgta gtgtcagtat catcaccaga tgctgttact	4380
40	acataaatg gatactcac ttcgtcatca aagacatctg aggagcactt ttagaaaaca	4440
	gtttcttttg ctggctctta cagagattgg tcctattcag gacagcgta agagttaggt	4500
	gttgaatttc ttaagcgtgg tgacaaaatt gtgtaccaca ctctggagag ccccgctgag	4560
45	tttcatcttg acggtgaggt tctttcactt gacaaactaa agagtctctt atccctgcgg	4620
	gaggttaaga ctataaaagt gttcacaact gtggacaaca ctaatctcca cacacagctt	4680
	gtggatatgt ctatgacata tggacagcag tttgggtcaa catacttgga tgggtgctgat	4740
50	gttacaaaaa ttaaacctca tgtaaatcat gagggttaaga ctttctttgt actacctagt	4800
	gatgacacac tacgtagtga agctttcgag tactaccata ctcttgatga gagttttctt	4860
	ggtaggtaca tgtctgcttt aaaccacaca aagaaatgga aatttcctca agttggtggt	4920
55	ttaacttcaa ttaaatgggc tgataacaat tgttatttgt ctagtgtttt attagcactt	4980
	caacagcttg aagtcaaatt caatgcacca gcacttcaag aggcttatta tagagcccgt	5040

EP 1 694 829 B1

	gctggtgatg ctgctaactt ttgtgcactc atactcgctt acagtaataa aactgttggc	5100
	gagcttggtg atgtcagaga aactatgacc catcttctac agcatgctaa tttggaatct	5160
5	gcaaagcgag ttcttaatgt ggtgtgtaaa cattgtggtc agaaaactac taccttaacg	5220
	ggtgtagaag ctgtgatgta tatgggtact ctatcttatg ataatcttaa gacaggtggt	5280
	tccattccat gtgtgtgtgg tcgtgatgct acacaatatc tagtacaaca agagtcttct	5340
10	tttgttatga tgtctgcacc acctgctgag tataaattac agcaaggtag attcttatgt	5400
	gcgaatgagt aacttggtaa ctatcagtgt gggtcattaca ctcatataac tgctaaggag	5460
	accctctatc gtattgacgg agctcacctt acaaagatgt cagagtacaa aggaccagtg	5520
15	actgatgttt tctacaagga aacatcttac actacaacca tcaagcctgt gtcgtataaa	5580
	ctcgtatggag ttactttacac agagattgaa ccaaaattgg atgggtatta taaaaaggat	5640
	aatgcttact atacagagca gcctatagac cttgtacca ctcaaccatt accaaatgcg	5700
20	agttttgata atttcaaact cacatgttct aacacaaaat ttgctgatga tttaaatcaa	5760
	atgacaggct tcacaaagcc agcttcacga gagctatctg tcacattctt cccagacttg	5820
	aatggcgatg tagtggctat tgactataga cactattcag cgagtttcaa gaaagggtgct	5880
25	aaattactgc ataagccaat tgtttggcac attaaccagg ctacaacca gacaacgttc	5940
	aaaccaaaaca cttggtgttt acgttgtctt tggagtacaa agccagtaga tactttcaat	6000
	tcattttgaag ttctggcagt agaagacaca caaggaatgg acaatcttgc ttgtgaaagt	6060
30	caacaaccca cctctgaaga agtagtgga aatcctacca tacagaagga agtcatagag	6120
	tgtgacgtga aaactaccga agttgtaggc aatgtcatac ttaaaccatc agatgaagg	6180
	gttaaagtaa cacaagagtt aggtcatgag gatcttatgg ctgcttatgt ggaaaacaca	6240
35	agcattacca ttaagaaacc taatgagctt tcactagcct taggtttaaa aacaattgcc	6300
	actcatggta ttgctgcaat taatagtgtt ccttggagta aaattttggc ttatgtcaaa	6360
	ccattcttag gacaagcagc aattacaaca tcaaattgag ctaagagatt agcacaacgt	6420
40	gtgtttaaca atttatatgcc ttatgtgttt acattattgt tccaattgtg tacttttact	6480
	aaaagtacca attctagaat tagagcttca ctacctaca ctattgctaa aaatagtgtt	6540
	aagagtgttg ctaaattatg tttggatgcc ggcattaatt atgtgaagtc acccaaattt	6600
45	tctaaattgt tcacaatcgc tatgtggcta ttgttgtaa gtatttgctt aggttctcta	6660
	atctgtgtaa ctgctgcttt tgggtgactc ttatctaatt ttggtgctcc ttcttattgt	6720
	aatggcgtaa gagaattgta tcttaattcg tctaacgtta ctactatgga tttctgtgaa	6780
50	ggttcttttc cttgcagcat ttgtttaagt ggattagact cccttgattc ttatccagct	6840
	cttgaaacca ttcagggtgac gatttcacgc tacaagctag acttgacaat tttagggtctg	6900
	gccgctgagt gggttttggc atatatgttg ttcacaaaat tcttttattt attaggctct	6960
55	tcagctataa tgcagggtgtt ctttggctat tttgctagtc atttcatcag caattcttgg	7020
	ctcatgtggt ttatcattag tattgtacaa atggcacccg tttctgcaat ggtaggatg	7080

## EP 1 694 829 B1

	tacatcttct ttgcttcttt ctactacata tggaagagct atgttcatat catggatggg	7140
	tgcacctctt cgacttgcat gatgtgctat aagcgcaatc gtgccacacg cgttgagtgt	7200
5	acaactattg ttaatggcat gaagagatct ttctatgtct atgcaaattg aggccgtggc	7260
	ttctgcaaga ctcaaatg gaattgtctc aattgtgaca cattttgcac tggtagtaca	7320
	ttcattagt atgaagttgc tcgtgatttg tcactccagt ttaaaagacc aatcaaccct	7380
10	actgaccagt catcgtatat tgttgatagt gttgctgtga aaaatggcgc gcttcacctc	7440
	tactttgaca aggctgggtca aaagacctat gagagacatc cgctctccca ttttgtcaat	7500
	ttagacaatt tgagagctaa caacactaaa gggttactgc ctattaatgt catagttttt	7560
15	gatggcaagt ccaaatgcga cgagtctgct tctaagtctg cttctgtgta ctacagtcag	7620
	ctgatgtgcc aacctattct gttgcttgac caagctcttg tatcagacgt tggagatagt	7680
	actgaagttt ccgttaagat gtttgatgct tatgtcgaca ctttttcagc aacttttagt	7740
20	gttcctatgg aaaaacttaa ggcacttggt gctacagctc acagcgagtt agcaaagggt	7800
	gtagcttttag atggtgtcct ttctacattc gtgtcagctg cccgacaagg tgttgttgat	7860
	accgatgttg acacaaagga tgttattgaa tgtctcaaac tttcacatca ctctgactta	7920
25	gaagtgacag gtgacagttg taacaatttc atgctcacct ataataaggt tgaaaacatg	7980
	acgcccagag atcttggcgc atgtattgac tgtaatgcaa ggcataatcaa tgcccaagta	8040
	gcaaaaagtc acaatgtttc actcatctgg aatgtaaaag actacatgtc tttatctgaa	8100
30	cagctgcgta aacaaattcg tagtgctgcc aagaagaaca acataccttt tagactaact	8160
	tgtgctacaa ctagacaggt tgtcaatgtc ataactacta aaatctcact caaggggtgg	8220
	aagattgtta gtacttgttt taaacttatg cttaaggcca cattattgtg cgttcttgct	8280
35	gcattgggtt gttatatcgt tatgccagta catacattgt caatccatga tggttacaca	8340
	aatgaaatca ttgggttaca agccattcag gatgggtgtc ctcgtgacat catttctact	8400
	gatgattgtt ttgcaaataa acatgctggg tttgacgcat ggtttagcca gcgtgggtgg	8460
40	tcatacaaaa atgacaaaag ctgccctgta gtagctgcta tcattacaag agagattggg	8520
	ttcatagtgc ctggcttacc ggggtactgtg ctgagagcaa tcaatgggtga cttcttgcat	8580
	tttctacctc gtgttttttag tgctgttggc aacatttgct acacaccttc caaactcatt	8640
45	gagtatagt attttgctac ctctgcttgc gttcttgctg ctgagtgtac aatttttaag	8700
	gatgctatgg gcaaacctgt gccatattgt tatgacacta atttgctaga gggttctatt	8760
	tcttatagt agcttcgtcc agacactcgt tatgtgctta tggatgggtc catcatacag	8820
50	tttcctaaca cttacctgga gggttctgtt agagtagtaa caacttttga tgctgagtac	8880
	tgtagacatg gtacatgcga aaggtcagaa gtaggtattt gcctatctac cagtggtaga	8940
	tgggttctta ataatgagca ttacagagct ctatcaggag ttttctgtgg tgttgatgcg	9000
55	atgaatctca tagctaacat ctttactcct cttgtgcaac ctgtgggtgc tttagatgtg	9060
	tctgcttcag tagtggtggg tgggtattatt gccatattgg tgacttgtgc tgcctactac	9120

EP 1 694 829 B1

	tttatgaaat tcagacgtgt ttttggtgag tacaaccatg ttgttgctgc taatgcactt	9180
	ttgtttttga tgtctttcac tatactctgt ctggtaccag cttacagctt tctgccggga	9240
5	gtctactcag tcttttactt gtacttgaca ttctatttca ccaatgatgt ttcatcttg	9300
	gctcaccttc aatgggttgc catgttttct cctattgtgc ctttttggat aacagcaatc	9360
	tatgtattct gtatttctct gaagcactgc cattggttct ttaacaacta tcttaggaaa	9420
10	agagtcatgt ttaatggagt tacatttagt accttcgagg aggctgcttt gtgtaccttt	9480
	ttgctcaaca aggaaatgta cctaaaattg cgtagcgaga cactgttgcc acttacacag	9540
	tataacaggt atcttgctct atataacaag tacaagtatt tcagtggagc cttagatact	9600
15	accagctatc gtgaagcagc ttgctgccac ttagcaaagg ctctaaatga ctttagcaac	9660
	tcagggtctg atgttctcta ccaaccacca cagacatcaa tcacttctgc tgttctgcag	9720
	agtggtttta ggaaaatggc attcccgtca ggcaaagttg aagggtgcat ggtacaagta	9780
20	acctgtggaa ctacaactct taatggattg tggttggatg acacagtata ctgtccaaga	9840
	catgtcattt gcacagcaga agacatgctt aatcctaact atgaagatct gctcattcgc	9900
	aaatccaacc atagcttttct tgttcaggct ggcaatgttc aacttcgtgt tattggccat	9960
25	tctatgcaaa attgtctgct taggcttaaa gttgatactt ctaaccctaa gacacccaag	10020
	tataaatttg tccgtatcca acctggtcaa acattttcag ttctagcatg ctacaatggg	10080
	tcaccatctg gtgtttatca gtgtgccatg agacctaatc ataccattaa aggttctttc	10140
30	cttaatggat catgtggtag tgttggtttt aacattgatt atgattgctg gtctttctgc	10200
	tatatgcatac atatggagct tccaacagga gtacacgctg gtactgactt agaaggtaaa	10260
	ttctatggtc catttggtga cagacaaact gcacaggctg caggtagaga cacaaccata	10320
35	acattaaatg ttttggtcatg gctgtatgct gctgttatca atggtgatag gtggtttctt	10380
	aatagattca cactactttt gaatgacttt aaccttggtg caatgaagta caactatgaa	10440
	cctttgacac aagatcatgt tgacatattg ggacctcttt ctgctcaaac aggaattgcc	10500
40	gtcttagata tgtgtgctgc tttgaaagag ctgctgcaga atggtatgaa tggtcgtact	10560
	atccttggtgta gcactatttt agaagatgag ttacaccat ttgatgttgt tagacaatgc	10620
	tctggtgtta ctttccaagg taagttcaag aaaattgtta agggcactca tcattggatg	10680
45	cttttaactt tcttgacatc actattgatt cttgttcaaa gtacacagtg gtcactgttt	10740
	ttctttgttt acgagaatgc tttcttgcca ttactcttg gtattatggc aattgctgca	10800
	tgtgctatgc tgcttggtta gcataagcac gcattcttgt gcttgtttct gttaccttct	10860
50	cttgcaacag ttgcttactt taatatgggc tacatgcctg ctagctgggt gatgcgtatc	10920
	atgacatggc ttgaattggc tgacactagc ttgtctggtt ataggcttaa ggattgtgtt	10980
	atgtatgctt cagcttttagt tttgcttatt ctcatgacag ctgcactgt ttatgatgat	11040
55	gctgctagac gtgtttggac actgatgaat gtcattacac ttgtttacaa agtctactat	11100
	ggtaatgctt tagatcaagc tatttccatg tgggccttag ttatttctgt aacctctaac	11160

	tattctggtg	tcgttacgac	tatcatgttt	ttagctagag	ctatagtgtt	tgtgtgtgtt	11220
	gagtattacc	cattgttatt	tattactggc	aacaccttac	agtgtatcat	gcttgtttat	11280
5	tgtttcttag	gctattgttg	ctgctgctac	tttggccttt	tctgtttact	caaccgttac	11340
	ttcaggctta	ctcttggtgt	ttatgactac	ttggtctcta	cacaagaatt	taggtatatg	11400
	aactcccagg	ggcttttgcc	tcctaagagt	agtattgatg	ctttcaagct	taacattaag	11460
10	ttgttgggta	ttggaggtaa	accatgtatc	aagggtgcta	ctgtacagtc	taaaatgtct	11520
	gacgtaaagt	gcacatctgt	ggtagctgtc	tcggttcttc	aacaacttag	agtagagtca	11580
	tcttctaaat	tgtgggcaca	atgtgtacaa	ctccacaatg	atattcttct	tgcaaaagac	11640
15	acaactgaag	ctttcgagaa	gatggtttct	cttttgtctg	ttttgctatc	catgcagggt	11700
	gctgtagaca	ttaatagggt	gtgcgaggaa	atgctcgata	accgtgctac	tcttcaggct	11760
	attgcttcag	aatttagttc	tttaccatca	tatgccgctt	atgccactgc	ccaggaggcc	11820
20	tatgagcagg	ctgtagctaa	tgggtgattct	gaagtcgttc	tcaaaaagtt	aaagaaatct	11880
	ttgaatgtgg	ctaaatctga	gtttgaccgt	gatgctgcca	tgcaacgcaa	gttggaaaag	11940
	atggcagatc	aggctatgac	ccaaatgtac	aaacaggcaa	gatctgagga	caagagggca	12000
25	aaagtaacta	gtgctatgca	aacaatgtc	ttcactatgc	ttaggaagct	tgataatgat	12060
	gcacttaaca	acattatcaa	caatgcgcgt	gatggttgtg	ttccactcaa	catcatacca	12120
	ttgactacag	cagccaaact	catggttggt	gtccctgatt	atggtaccta	caagaacact	12180
30	tgtgatggta	acacctttac	atatgcatct	gcactctggg	aaatccagca	agttgttgat	12240
	gcggatagca	agattgttca	acttagtgaa	attaacatgg	acaattcacc	aaatttggct	12300
	tggcctctta	ttgttacagc	tctaagagcc	aactcagctg	ttaaactaca	gaataatgaa	12360
35	ctgagtccag	tagcactacg	acagatgtcc	tgtgcggctg	gtaccacaca	aacagcttgt	12420
	actgatgaca	atgcacttgc	ctactataac	aattcgaagg	gaggtagggt	tgtgctggca	12480
	ttactatcag	accaccaaga	tctcaaattg	gctagattcc	ctaagagtga	tggtagagggt	12540
40	acaatttaca	cagaactgga	accaccttgt	aggtttggtta	cagacacacc	aaaagggcct	12600
	aaagtgaat	acttgtactt	catcaaaggc	ttaaacaacc	taaatagagg	tatggtgctg	12660
	ggcagtttag	ctgctacagt	acgtcttcag	gctggaaatg	ctacagaagt	acctgccaat	12720
45	tcaactgtgc	tttccttctg	tgcttttgca	gtagaccctg	ctaaagcata	taaggattac	12780
	ctagcaagtg	gaggacaacc	aatcaccaac	tgtgtgaaga	tgttgtgtac	acacactgggt	12840
	acaggacagg	caattactgt	aacaccagaa	gctaacatgg	accaagagtc	ctttgggtgggt	12900
50	gcttcattgtt	gtctgtattg	tagatgccac	attgaccatc	caaataccta	aggattctgt	12960
	gacttgaaag	gtaagtacgt	ccaaatacct	accacttgtg	ctaatagccc	agtgggtttt	13020
	acacttagaa	acacagtctg	taccgtctgc	ggaatgtgga	aaggttatgg	ctgtagttgt	13080
55	gaccaactcc	gcgaaccctt	gatgcagtct	gcggatgcat	caacgttttt	aaacgggttt	13140
	gcggtgtaag	tgacagccgt	cttacaccgt	gcggcacagg	cactagtact	gatgtcgtct	13200

	acagggcttt	tgatatttac	aacgaaaaag	ttgctggttt	tgcaaagttc	ctaaaaacta	13260
	attgctgtcg	cttccaggag	aaggatgagg	aaggcaattt	attagactct	tactttgtag	13320
5	ttaagaggca	tactatgtct	aactaccaac	atgaagagac	tatttataac	ttgggttaaag	13380
	attgtccagc	ggttgctgtc	catgactttt	tcaagtttag	agtagatggg	gacatggtac	13440
	cacatatatc	acgtcagcgt	ctaactaaat	acacaatggc	tgatttagtc	tatgctctac	13500
10	gtcattttga	tgagggtaat	tgtgatacat	taaaagaaat	actcgtcaca	tacaattgct	13560
	gtgatgatga	ttatttcaat	aagaaggatt	ggatgactt	cgtagagaat	cctgacatct	13620
	tacgcgtata	tgctaactta	ggtagcgtg	tacgccaatc	attattaaag	actgtacaat	13680
15	tctgcgatgc	tatgcgtgat	gcaggcattg	taggcgtact	gacattagat	aatcaggatc	13740
	ttaatgggaa	ctggtacgat	ttcgggtgatt	tcgtacaagt	agcaccaggc	tgcgaggttc	13800
	ctattgtgga	ttcatattac	tcattgctga	tgcccatcct	cactttgact	agggcattgg	13860
20	ctgctgagtc	ccatatggat	gctgatctcg	caaaaccact	tattaagtgg	gatttgctga	13920
	aatatgattt	tacggaagag	agactttgtc	tcttcgaccg	ttattttaaa	tattgggacc	13980
	agacatacca	tccaattgt	attaactgtt	tggatgatag	gtgtatcctt	cattgtgcaa	14040
25	actttaatgt	gttattttct	actgtgtttc	cacctacaag	tttggacca	ctagtaagaa	14100
	aaatatttgt	agatgggtgt	ccttttggtg	tttcaactgg	ataccatttt	cgtgagttag	14160
	gagtcgtaca	taatcaggat	gtaaacttac	atagctcgcg	tctcagtttc	aaggaaacttt	14220
30	tagtgtatgc	tgctgatcca	gctatgcatg	cagcttctgg	caatttattg	ctagataaac	14280
	gcactacatg	cttttcagta	gctgcactaa	caaacaatgt	tgcttttcaa	actgtcaaac	14340
	ccggtaat	taataaagac	ttttatgact	ttgctgtgtc	taaaggtttc	tttaagggaag	14400
35	gaagttctgt	tgaactaaaa	cacttcttct	ttgctcagga	tggcaacgct	gctatcagtg	14460
	attatgacta	ttatcgttat	aatctgccaa	caatgtgtga	tatcagacaa	ctcctattcg	14520
	tagttgaagt	tgttgataaa	tactttgatt	gttacgatgg	tggctgtatt	aatgccaacc	14580
40	aagtaatcgt	taacaatctg	gataaatcag	ctggtttccc	atttaataaa	tggggtaagg	14640
	ctagacttta	ttatgactca	atgagttatg	aggatcaaga	tgactttttc	gcgtatacta	14700
	agcgtaatgt	catccctact	ataactcaaa	tgaatcttaa	gtatgccatt	agtgcaaaga	14760
45	atagagctcg	caccgtagct	gggtgtctcta	tctgtagtac	tatgacaaat	agacagtttc	14820
	atcagaaatt	attgaagtca	atagccgcca	ctagaggagc	tactgtggta	attggaacaa	14880
	gcaagtttta	cgggtggctgg	cataatatgt	taaaaactgt	ttacagtgat	gtagaaactc	14940
50	cacaccttat	gggttgggat	tatccaaaat	gtgacagagc	catgcctaac	atgcttagga	15000
	taatggcctc	tcttggttctt	gctcgcaaac	ataacacttg	ctgtaactta	tcacaccggt	15060
	tctacagggt	agctaacgag	tgtgcgcaag	tattaagtga	gatgggtcatg	tgtggcggt	15120
55	cactatatgt	taaaccagggt	ggaacatcat	ccggtgatgc	tacaactgct	tatgctaata	15180
	gtgtctttta	catttggtcaa	gctgttacag	ccaatgtaaa	tgactttctt	tcaactgatg	15240



	gtaataagat agctgacaag tatgtccgca atctacaaca caggctctat gagtgtctct	15300
	atagaaatag ggatgttgat catgaattcg tggatgagtt ttacgcttac ctgcgtaaac	15360
5	atttctccat gatgattctt tctgatgatg ccgttgtgtg ctataacagt aactatgcgg	15420
	ctcaagggtt agtagctagc attaagaact ttaaggcagt tctttattat caaataatg	15480
	tgttcatgtc tgaggcaaaa tgttggactg agactgacct tactaaagga cctcacgaat	15540
10	tttgctcaca gcatacaatg ctagttaaac aaggagatga ttacgtgtac ctgccttacc	15600
	cagatccatc aagaatatta ggcgcaggct gttttgtcga tgatattgtc aaaacagatg	15660
	gtacacttat gattgaaagg ttcgtgtcac tggctattga tgcttaccac cttacaaaac	15720
15	atcctaatac ggagtatgct gatgtctttc acttgtattt acaatacatt agaaagttag	15780
	atgatgagct tactggccac atgttggaca tgtattccgt aatgctaact aatgataaca	15840
	cctcacggtc ctgggaacct gagttttatg aggctatgta cacaccacat acagtcttgc	15900
20	aggctgtagg tgcttgtgta ttgtgcaatt cacagacttc acttcgttgc ggtgcctgta	15960
	ttaggagacc attcctatgt tgcaagtgct gctatgacca tgtcatttca acatcacaca	16020
	aattagtgtt gtctgttaat ccctatgttt gcaatgcccc aggttgtgat gtcactgatg	16080
25	tgacacaact gtatctagga ggtatgagct attattgcaa gtcacataag cctccatta	16140
	gttttccatt atgtgctaatt ggtcagggtt ttggtttata caaaaacaca tgtgtaggca	16200
	gtgacaatgt cactgacttc aatgcgatag caacatgtga ttggactaat gctggcgatt	16260
30	acatacttgc caacacttgt actgagagac tcaagctttt cgcagcagaa acgctcaaag	16320
	ccactgagga aacatttaag ctgtcatatg gtattgccac tgtacgcgaa gtactctctg	16380
	acagagaatt gcatctttca tgggagggtg gaaaacctag accaccattg aacagaaact	16440
35	atgtctttac tggttaccgt gtaactaaaa atagtaaagt acagattgga gagtacacct	16500
	ttgaaaaagg tgactatggt gatgctgttg tgtacagagg tactacgaca tacaagttga	16560
	atgttggga ttactttgtg ttgacatctc aactgtaat gccacttagt gcacctactc	16620
40	tagtgccaca agagcactat gtgagaatta ctggcttgta cccaacactc aacatctcag	16680
	atgagttttc tagcaatgtt gcaaattatc aaaagggtcg catgcaaaag tactctacac	16740
	tccaaggacc acctgggtact ggtaagagtc attttgccat cggacttgct ctctattacc	16800
45	catctgctcg catagtgtat acggcatgct ctcatgcagc tgttgatgcc ctatgtgaaa	16860
	aggcattaaa atatttgccc atagataaat gtagtagaat catacctgcg cgtgcgcgcg	16920
	tagagtgttt tgataaattc aaagtgaatt caacactaga acagtatgtt ttctgcactg	16980
50	taaatgcatt gccagaaaca actgctgaca ttgtagtctt tgatgaaatc tctatggcta	17040
	ctaattatga cttgagtgtt gtcaatgcta gacttcgtgc aaaacactac gtctatattg	17100
	gcgatcctgc tcaattacca gcccccgca cattgctgac taaaggcaca ctagaaccag	17160
55	aatattttta ttcagtgtgc agacttatga aaacaatagg tccagacatg ttccttgtaa	17220
	cttgtcgccg ttgtcctgct gaaattgttg aactgtgag tgcttttagt tatgacaata	17280

## EP 1 694 829 B1

agctaaaagc acacaaggat aagtcagctc aatgcttcaa aatgtttctac aaaggtgtta 17340  
 ttacacatga tgtttcatct gcaatcaaca gacctcaaat aggcgttgta agagaatttc 17400  
 5 ttacacgcaa tcctgcttgg agaaaagctg tttttatctc accttataat tcacagaacg 17460  
 ctgtagcttc aaaaatctta ggattgccta cgcagactgt tgattcatca caggggttctg 17520  
 aatatgacta tgtcatattc acacaaacta ctgaaacagc acactcttgt aatgtcaacc 17580  
 10 gcttcaatgt ggctatcaca agggcaaaaa ttggcatttt gtgcataatg tctgatagag 17640  
 atctttatga caaactgcaa tttaacaagtc tagaaatacc acgtcgcaat gtggctacat 17700  
 tacaagcaga aaatgtaact ggacttttta aggactgtag taagatcatt actggctctc 17760  
 15 atcctacaca ggcacctaca cacctcagcg ttgatataaa gttcaagact gaaggattat 17820  
 gtgttgacat accaggcata ccaaaggaca tgacctaccg tagactcatc tctatgatgg 17880  
 gtttcaaaat gaattaccaa gtcaatgggt accctaatat gtttatcacc cgcgaagaag 17940  
 20 ctattcgtca cgttcgtgcg tggattggct ttgatgtaga gggctgtcat gcaactagag 18000  
 atgctgtggg tactaaccta cctctccagc taggattttc tacagggtgtt aacttagtag 18060  
 ctgtaccgac tggttatggt gacactgaaa ataacacaga attcaccaga gttaatgcaa 18120  
 25 aacctccacc aggtgaccag tttaaacatc ttataccact catgtataaa ggcttgccct 18180  
 ggaatgtagt gcgtattaag atagtacaaa tgctcagtga tacctgaaa ggattgtcag 18240  
 acagagtcgt gttcgtcctt tgggcgcatg gctttgagct tacatcaatg aagtactttg 18300  
 30 tcaagattgg acctgaaaga acgtgttgct tgtgtgacaa acgtgcaact tgcttttcta 18360  
 cttcatcaga tacttatgcc tgctggaatc attctgtggg ttttgactat gtctataacc 18420  
 cttttatgat tgatgttcag cagtggggct ttacgggtaa ctttcagagt aaccatgacc 18480  
 35 aacattgcca ggtacatgga aatgcacatg tggctagttg tgatgctatc atgactagat 18540  
 gtttagcagt ccatgagtg c ttgttaagc gcgttgattg gtctgttgaa taccctatta 18600  
 taggagatga actgaggggt aattctgctt gcagaaaagt acaacacatg gttgtgaagt 18660  
 40 ctgcattgct tgctgataag tttccagttc ttcattgacat tggaaatcca aaggctatca 18720  
 agtgtgtgcc tcaggctgaa gtagaatgga agttctacga tgctcagcca tgtagtga 18780  
 aagcttaca aatagaggaa ctcttctatt cttatgctac acatcacgat aaattcactg 18840  
 45 atggtgtttg tttgttttgg aattgtaacg ttgatcgtaa cccagccaat gcaattgtgt 18900  
 gtaggtttga cacaagagtc ttgtcaaact tgaacttacc aggcgtgat ggtggtagt 18960  
 tgtatgtgaa taagcatgca ttccacactc cagctttcga taaaagtgca tttactaatt 19020  
 50 taaagcaatt gcctttcttt tactattctg atagtccttg tgagtctcat ggcaaacaag 19080  
 tagtgtcgga tattgattat gttccactca aatctgctac gtgtattaca cgatgcaatt 19140  
 taggtggtgc tgtttgcaga caccatgcaa atgagtaccg acagtacttg gatgcatata 19200  
 55 atatgatgat ttctgctgga tttagcctat ggatttaca acaatttgat acttataacc 19260  
 tgtggaatac atttaccagg ttacagagtt tagaaaatgt ggcttataat gttgttaata 19320

EP 1 694 829 B1

	aaggacactt	tgatggacac	gccggcgaag	cacctgtttc	catcattaat	aatgctgttt	19380
	acacaaaggt	agatgggtatt	gatgtggaga	tctttgaaaa	taagacaaca	cttcctgtta	19440
5	atgttgcatt	tgagcttttg	gctaagcgta	acattaaacc	agtgccagag	attaagatac	19500
	tcaataatth	gggtgttgat	atcgctgcta	atactgtaat	ctgggactac	aaaagagaag	19560
	ccccagcaca	tgtatctaca	ataggtgtct	gcacaatgac	tgacattgcc	aagaaaccta	19620
10	ctgagagtgc	ttgttcttca	cttactgtct	tgtttgatgg	tagagtggaa	ggacaggtag	19680
	acctttttag	aaacgcccgt	aatgggtgtt	taataacaga	aggttcagtc	aaaggtctaa	19740
	caccttcaaa	gggaccagca	caagctagcg	tcaatggagt	cacattaatt	ggagaatcag	19800
15	taaaaacaca	gtttaactac	tttaagaaag	tagacggcat	tattcaacag	ttgcctgaaa	19860
	cctactttac	tcagagcaga	gacttagagg	attttaagcc	cagatcacia	atggaaactg	19920
	actttctcga	gctcgctatg	gatgaattca	tacagcgata	taagctcgag	ggctatgcct	19980
20	tcgaacacat	cgtttatgga	gatttcagtc	atggacaact	tggcggctct	catttaatga	20040
	taggccttagc	caagcgctca	caagattcac	cacttaaatt	agaggatttt	atccctatgg	20100
	acagcacagt	gaaaaattac	ttcataacag	atgcgcaaac	aggttcatca	aaatgtgtgt	20160
25	gttctgtgat	tgatctttta	cttgatgact	ttgtcgagat	aataaagtca	caagatttgt	20220
	cagtgatthc	aaaagtggtc	aaggttacaa	ttgactatgc	tgaaatttca	ttcatgcttt	20280
	gggtgaagga	tggacatgth	gaaaccttct	acccaaaact	acaagcaagt	caagcgtggc	20340
30	aaccaggtgt	tgcgatgcct	aacttgtaca	agatgcaaag	aatgcttctt	gaaaagtgtg	20400
	accttcagaa	ttatgggtgaa	aatgctgtta	taccaaaggg	aataatgatg	aatgtcgcaa	20460
	agtatactca	actgtgtcaa	tacttaataa	cacttacttt	agctgtaccc	tacaacatga	20520
35	gagttattca	ctttgggtgct	ggctctgata	aaggagttgc	accaggtaca	gctgtgctca	20580
	gacaatggth	gccaactggc	acactacttg	tcgattcaga	tcttaatgac	ttcgtctccg	20640
	acgcagattc	tacttttaatt	ggagactgtg	caacagtaca	tacggctaatt	aaatgggacc	20700
40	ttattattag	cgatatgtat	gaccctagga	ccaacatgt	gacaaaagag	aatgactcta	20760
	aagaagggth	tttcacttat	ctgtgtggat	ttataaagca	aaaactagcc	ctgggtggth	20820
	ctatagctgt	aaagataaca	gagcattctt	ggaatgctga	cctttacaag	cttatgggccc	20880
45	atttctcatg	gtggacagct	tttgttacaa	atgtaaatgc	atcatcatcg	gaagcattth	20940
	taattggggc	taactatctt	ggcaagccga	aggaacaaat	tgatggctat	accatgcatg	21000
	ctaactacat	tttctggagg	aacacaaatc	ctatccagth	gtcttcctat	tcactctthg	21060
50	acatgagcaa	atttcctctt	aaattaagag	gaactgctgt	aatgtctctt	aaggagaatc	21120
	aaatcaatga	tatgatttht	tctcttctgg	aaaaaggtag	gcttatcatt	agagaaaaca	21180
	acagagttgt	ggthtcaagt	gatattcttg	ttaacaacta	a		21221
55	<210> 32						
	<211> 297						
	<212> DNA						
	<213> CORONAVIRUS						

EP 1 694 829 B1

<400> 32

atggacccca atcaaacc aa cgtagtgccc cccgcattac atttggtgga cccacagatt 60  
5 caactgacaa taaccagaat ggaggacgca atggggcaag gccaaaacag cgccgacccc 120  
aaggtttacc caataatact gcgtcttggt tcacagctct cactcagcat ggcaaggagg 180  
aacttagatt cctcgcaggc cagggcggtc caatcaacac caatagtggg ccagatgacc 240  
10 aaattggcta ctaccgaaga gctacccgac gagttcgtgg tggtgacggc aaaatga 297

<210> 33

<211> 98

15 <212> PRT

<213> CORONAVIRUS

<400> 33

20 Met Asp Pro Asn Gln Thr Asn Val Val Pro Pro Ala Leu His Leu Val  
1 5 10 15  
25 Asp Pro Gln Ile Gln Leu Thr Ile Thr Arg Met Glu Asp Ala Met Gly  
20 25 30  
Gln Gly Gln Asn Ser Ala Asp Pro Lys Val Tyr Pro Ile Ile Leu Arg  
35 40 45  
30 Leu Gly Ser Gln Leu Ser Leu Ser Met Ala Arg Arg Asn Leu Asp Ser  
50 55 60  
35 Leu Glu Ala Arg Ala Phe Gln Ser Thr Pro Ile Val Val Gln Met Thr  
65 70 75 80  
Lys Leu Ala Thr Thr Glu Glu Leu Pro Asp Glu Phe Val Val Val Thr  
85 90 95  
40 Ala Lys

<210> 34

45 <211> 213

<212> DNA

<213> CORONAVIRUS

<400> 34

50 atgctgccac cgtgctacaa cttcctcaag gaacaacatt gccaaaaggc ttctacgcag 60  
aggggaagcag aggcggcagt caagcctctt ctcgctcctc atcacgtagt cgcggtaatt 120  
55 caagaaattc aactcctggc agcagtaggg gaaattctcc tgctcgaatg gctagcggag 180  
gtggtgaaac tgccctcgcg ctattgctgc tag 213

# EP 1 694 829 B1

<210> 35  
 <211> 70  
 <212> PRT  
 <213> CORONAVIRUS

5

<400> 35

10

Met Leu Pro Pro Cys Tyr Asn Phe Leu Lys Glu Gln His Cys Gln Lys  
 1 5 10 15

Ala Ser Thr Gln Arg Glu Ala Glu Ala Ala Val Lys Pro Leu Leu Ala  
 20 25 30

15

Pro His His Val Val Ala Val Ile Gln Glu Ile Gln Leu Leu Ala Ala  
 35 40 45

20

Val Gly Glu Ile Leu Leu Leu Glu Trp Leu Ala Glu Val Val Lys Leu  
 50 55 60

Pro Ser Arg Tyr Cys Cys  
 65 70

25

<210> 36  
 <211> 1377  
 <212> DNA  
 <213> CORONAVIRUS

30

<220>  
 <221> CDS  
 <222> (67)..(1335)  
 <223>

35

<400> 36

40

atgaagggtca ccaaactgct gcatttagag acgtacttgt tgtttttaaataaacgaacaa 60  
 attaaa atg tct gat aat gga ccc caa tca aac caa cgt agt gcc ccc 108

45

50

55

EP 1 694 829 B1

	Met	Ser	Asp	Asn	Gly	Pro	Gln	Ser	Asn	Gln	Arg	Ser	Ala	Pro			
	1				5					10							
5	cgc Arg 15	att Ile	aca Thr	ttt Phe	ggt Gly 20	gga Gly 20	ccc Pro	aca Thr	gat Asp	tca Ser 25	act Thr 25	gac Asp	aat Asn	aac Asn	cag Gln 30	156	
	gga Gly	gga Gly	cgc Arg	aat Asn	ggg Gly 35	gca Ala	agg Arg	cca Pro	aaa Lys	cag Gln 40	cgc Arg	cga Arg	ccc Pro	caa Gln	ggt Gly 45	204	
10	ccc Pro	aat Asn	aat Asn	act Thr 50	gcg Ala	tct Ser	tgg Trp	ttc Phe	aca Thr 55	gct Ala	ctc Leu	act Thr	cag Gln	cat His 60	ggc Gly	aag Lys	252
15	gag Glu	gaa Glu	ctt Leu 65	aga Arg	ttc Phe	cct Pro	cga Arg	ggc Gly 70	cag Gln	ggc Gly	gtt Val	cca Pro	atc Ile 75	aac Asn	acc Thr	aat Asn	300
	agt Ser	ggt Gly 80	cca Pro	gat Asp	gac Asp	caa Gln	att Ile 85	ggc Gly	tac Tyr	tac Tyr	cga Arg	aga Arg 90	gct Ala	acc Thr	cga Arg	cga Arg	348
20	gtt Val 95	cgt Arg	ggt Gly	ggt Gly	gac Asp	ggc Gly 100	aaa Lys	atg Met	aaa Lys	gag Glu	ctc Leu 105	agc Ser	ccc Pro	aga Arg	tgg Trp	tac Tyr 110	396
25	ttc Phe	tat Tyr	tac Tyr	cta Leu	gga Gly 115	act Thr	ggc Gly	cca Pro	gaa Glu	gct Ala 120	tca Ser	ctt Leu	ccc Pro	tac Tyr	ggc Gly 125	gct Ala	444
	aac Asn	aaa Lys	gaa Glu	ggc Gly 130	atc Ile	gta Val	tgg Trp	gtt Val	gca Ala 135	act Thr	gag Glu	gga Gly	gcc Ala	ttg Leu 140	aat Asn	aca Thr	492
30	ccc Pro	aaa Lys	gac Asp 145	cac His	att Ile	ggc Gly	acc Thr	cgc Arg 150	aat Asn	cct Pro	aat Asn	aac Asn	aat Asn 155	gct Ala	gcc Ala	acc Thr	540
35	gtg Val 160	cta Leu	caa Gln	ctt Leu	cct Pro	caa Gln	gga Gly 165	aca Thr	aca Thr	ttg Leu	cca Pro	aaa Lys 170	ggc Gly	ttc Phe	tac Tyr	gca Ala	588
	gag Glu 175	gga Gly	agc Ser	aga Arg	ggc Gly 180	ggc Gly	agt Ser	caa Gln	gcc Ala	tct Ser	tct Ser 185	cgc Arg	tcc Ser	tca Ser	tca Ser	cgt Arg 190	636
40	agt Ser	cgc Arg	ggt Gly	aat Asn	tca Ser 195	aga Arg	aat Asn	tca Ser	act Thr	cct Pro 200	ggc Gly	agc Ser	agt Ser	agg Arg	gga Gly 205	aat Asn	684
45	tct Ser	cct Pro	gct Ala	cga Arg 210	atg Met	gct Ala	agc Ser	gga Gly	ggt Gly 215	ggt Gly	gaa Glu	act Thr	gcc Ala	ctc Leu 220	gcg Ala	cta Leu	732
	ttg Leu	ctg Leu	cta Leu 225	gac Asp	aga Arg	ttg Leu	aac Asn	cag Gln 230	ctt Leu	gag Glu	agc Ser	aaa Lys	ggt Val 235	tct Ser	ggt Gly	aaa Lys	780
50	ggc Gly 240	caa Gln	caa Gln	caa Gln	caa Gln	ggc Gly 245	caa Gln	act Thr	gtc Val	act Thr	aag Lys	aaa Lys 250	tct Ser	gct Ala	gct Ala	gag Glu	828
55	gca Ala 255	tct Ser	aaa Lys	aag Lys	cct Pro	cgc Arg 260	caa Gln	aaa Lys	cgt Arg	act Thr	gcc Ala 265	aca Thr	aaa Lys	cag Gln	tac Tyr	aac Asn 270	876
	gtc	act	caa	gca	ttt	ggg	aga	cgt	ggt	cca	gaa	caa	acc	caa	gga	aat	924

EP 1 694 829 B1

	Val	Thr	Gln	Ala	Phe 275	Gly	Arg	Arg	Gly	Pro 280	Glu	Gln	Thr	Gln	Gly 285	Asn	
5	ttc	ggg	gac	caa	gac	cta	atc	aga	caa	gga	act	gat	tac	aaa	cat	tgg	972
	Phe	Gly	Asp	Gln 290	Asp	Leu	Ile	Arg	Gln 295	Gly	Thr	Asp	Tyr	Lys 300	His	Trp	
10	ccg	caa	att	gca	caa	ttt	gct	cca	agt	gcc	tct	gca	ttc	ttt	gga	atg	1020
	Pro	Gln	Ile	Ala	Gln	Phe	Ala	Pro 310	Ser	Ala	Ser	Ala	Phe 315	Phe	Gly	Met	
	tca	cgc	att	ggc	atg	gaa	gtc	aca	cct	tcg	gga	aca	tgg	ctg	act	tat	1068
	Ser	Arg	Ile	Gly	Met	Glu	Val 325	Thr	Pro	Ser	Gly	Thr 330	Trp	Leu	Thr	Tyr	
15	cat	gga	gcc	att	aaa	ttg	gat	gac	aaa	gat	cca	caa	ttc	aaa	gac	aac	1116
	His	Gly	Ala	Ile	Lys	Leu 340	Asp	Asp	Lys	Asp	Pro 345	Gln	Phe	Lys	Asp	Asn 350	
20	gtc	ata	ctg	ctg	aac	aag	cac	att	gac	gca	tac	aaa	aca	ttc	cca	cca	1164
	Val	Ile	Leu	Leu	Asn 355	Lys	His	Ile	Asp	Ala 360	Tyr	Lys	Thr	Phe	Pro 365	Pro	
	aca	gag	cct	aaa	aag	gac	aaa	aag	aaa	aag	act	gat	gaa	gct	cag	cct	1212
	Thr	Glu	Pro	Lys 370	Lys	Asp	Lys	Lys	Lys 375	Lys	Thr	Asp	Glu	Ala 380	Gln	Pro	
25	ttg	ccg	cag	aga	caa	aag	aag	cag	ccc	act	gtg	act	ctt	ctt	cct	gcg	1260
	Leu	Pro	Gln 385	Arg	Gln	Lys	Lys	Gln 390	Pro	Thr	Val	Thr	Leu 395	Leu	Pro	Ala	
30	gct	gac	atg	gat	gat	ttc	tcc	aga	caa	ctt	caa	aat	tcc	atg	agt	gga	1308
	Ala	Asp	Met	Asp	Asp	Phe 405	Ser	Arg	Gln	Leu	Gln	Asn 410	Ser	Met	Ser	Gly	
	gct	tct	gct	gat	tca	act	cag	gca	taa	acact	catga	tgaccacaca					1355
	Ala	Ser	Ala	Asp	Ser	Thr 420	Gln	Ala									
35	aggcagatgg	gctatgtaaa	cg														1377
40	<210>	37															
	<211>	422															
	<212>	PRT															
	<213>	CORONAVIRUS															
	<400>	37															
45	Met	Ser	Asp	Asn	Gly 5	Pro	Gln	Ser	Asn	Gln 10	Arg	Ser	Ala	Pro	Arg 15	Ile	
50	Thr	Phe	Gly	Gly 20	Pro	Thr	Asp	Ser	Thr 25	Asp	Asn	Asn	Gln	Asn 30	Gly	Gly	
	Arg	Asn	Gly 35	Ala	Arg	Pro	Lys	Gln 40	Arg	Arg	Pro	Gln	Gly 45	Leu	Pro	Asn	
55	Asn	Thr 50	Ala	Ser	Trp	Phe	Thr 55	Ala	Leu	Thr	Gln	His 60	Gly	Lys	Glu	Glu	

Leu Arg Phe Pro Arg Gly Gln Gly Val Pro Ile Asn Thr Asn Ser Gly  
 65 70 75 80  
 5 Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg Ala Thr Arg Arg Val Arg  
 85 90 95  
 10 Gly Gly Asp Gly Lys Met Lys Glu Leu Ser Pro Arg Trp Tyr Phe Tyr  
 100 105 110  
 Tyr Leu Gly Thr Gly Pro Glu Ala Ser Leu Pro Tyr Gly Ala Asn Lys  
 115 120 125  
 15 Glu Gly Ile Val Trp Val Ala Thr Glu Gly Ala Leu Asn Thr Pro Lys  
 130 135 140  
 20 Asp His Ile Gly Thr Arg Asn Pro Asn Asn Asn Ala Ala Thr Val Leu  
 145 150 155 160  
 Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys Gly Phe Tyr Ala Glu Gly  
 165 170 175  
 25 Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg Ser Ser Ser Arg Ser Arg  
 180 185 190  
 30 Gly Asn Ser Arg Asn Ser Thr Pro Gly Ser Ser Arg Gly Asn Ser Pro  
 195 200 205  
 Ala Arg Met Ala Ser Gly Gly Gly Glu Thr Ala Leu Ala Leu Leu Leu  
 210 215 220  
 35 Leu Asp Arg Leu Asn Gln Leu Glu Ser Lys Val Ser Gly Lys Gly Gln  
 225 230 235 240  
 40 Gln Gln Gln Gly Gln Thr Val Thr Lys Lys Ser Ala Ala Glu Ala Ser  
 245 250 255  
 Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr Lys Gln Tyr Asn Val Thr  
 260 265 270  
 45 Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly  
 275 280 285  
 50 Asp Gln Asp Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln  
 290 295 300  
 Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg  
 305 310 315 320  
 55 Ile Gly Met Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr His Gly  
 325 330 335



EP 1 694 829 B1

Ala Ile Lys Leu Asp Asp Lys Asp Pro Gln Phe Lys Asp Asn Val Ile  
340 345 350

5 Leu Leu Asn Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu  
355 360 365

10 Pro Lys Lys Asp Lys Lys Lys Lys Thr Asp Glu Ala Gln Pro Leu Pro  
370 375 380

Gln Arg Gln Lys Lys Gln Pro Thr Val Thr Leu Leu Pro Ala Ala Asp  
385 390 395 400

15 Met Asp Asp Phe Ser Arg Gln Leu Gln Asn Ser Met Ser Gly Ala Ser  
405 410 415

20 Ala Asp Ser Thr Gln Ala  
420

<210> 38

<211> 1377

<212> DNA

<213> CORONAVIRUS

<400> 38

30 atgaagggtca ccaaactgct gcatttagag acgtacttgt tgttttaaat aaacgaacaa 60  
attaaaatgt ctgataatgg accccaatca aaccaacgta gtgccccccg cattacattt 120  
ggtggaccga cagattcaac tgacaataac cagaatggag gacgcaatgg ggcaaggcca 180  
35 aaacagcgcc gacccaagg tttacccaat aatactgcgt cttgggtcac agctctcact 240  
cagcatggca aggaggaact tagattccct cgaggccagg gcgttccaat caacaccaat 300  
agtgggtccag atgaccaaatt tggctactac cgaagagcta cccgacgagt tcgtgggtgt 360  
40 gacggcaaaa tgaaagagct cagccccaga tgggtacttct attacctagg aactggccca 420  
gaagcttcac ttccctacgg cgctaacaaa gaaggcatcg tatgggttgc aactgagggga 480  
gccttgaata caccacaaga ccacattggc acccgcaatc ctaataacaa tgctgccacc 540  
45 gtgctacaac ttcctcaagg aacaacattg ccaaaaggct tctacgcaga gggaagcaga 600  
ggcggcagtc aagcctcttc tcgctcctca tcacgtagtc gcggtaatc aagaaattca 660  
actcctggca gcagtagggg aaattctcct gctcgaatgg ctagcggagg tggtgaaact 720  
50 gccctcgcgc tattgctgct agacagattg aaccagcttg agagcaaagt ttctggtaaa 780  
ggccaacaac aacaaggcca aactgtcact aagaaatctg ctgctgaggc atctaaaaag 840  
cctcgccaaa aacgtactgc caaaaacag tacaacgtca ctcaagcatt tgggagacgt 900  
55 ggtccagaac aaaccaagg aaatttcggg gaccaagacc taatcagaca aggaactgat 960

EP 1 694 829 B1

	tacaaacatt ggccgcaa at tgcacaattt gctccaagt cctctgcatt ctttggaatg	1020
	tcacgcattg gcatggaagt cacaccttcg ggaacatggc tgacttatca tggagccatt	1080
5	aaattggatg acaaagatcc acaattcaaa gacaacgtca tactgctgaa caagcacatt	1140
	gacgcataca aaacattccc accaacagag cctaaaaagg acaaaaagaa aaagactgat	1200
	gaagctcagc ctttgccgca gagacaaaag aagcagccca ctgtgactct tcttcctgcg	1260
10	gctgacatgg atgatttctc cagacaactt caaaattcca tgagtggagc ttctgctgat	1320
	tcaactcagg cataaacact catgatgacc acacaaggca gatgggctat gtaaacg	1377
15	<210> 39 <211> 204 <212> DNA <213> CORONAVIRUS	
20	<400> 39	
	atattaggtt tttaacctacc caggaaaagc caaccaacct cgatctcttg tagatctgtt	60
	ctctaaacga actttaaaat ctgtgtagct gtcgctcggc tgcattgccta gtgcacctac	120
25	gcagtataaa caataataaa ttttactgtc gttgacaaga aacgagtaac tcgtccctct	180
	tctgcagact gcttacgggt tcgt	204
30	<210> 40 <211> 809 <212> DNA <213> CORONAVIRUS	
35	<400> 40	
	actcaagcat ttgggagacg tgggtccagaa caaacccaag gaaatttcgg ggaccaagac	60
	ctaatacagac aaggaactga ttacaaacat tggccgcaaa ttgcacaatt tgctccaagt	120
40	gcctctgcat tctttggaat gtcacgcatt ggcatggaag tcacaccttc gggacatgg	180
	ctgacttatc atggagccat taaattggat gacaaagatc cacaattcaa agacaacgctc	240
	atactgctga acaagcacat tgacgcatac aaaacattcc caccaacaga gcctaaaaag	300
45	gacaaaaaga aaaagactga tgaagctcag cctttgccgc agagacaaaa gaagcagccc	360
	actgtgactc tcttcctgc ggctgacatg gatgatttct ccagacaact tcaaaattcc	420
	atgagtggag cttctgctga ttcaactcag gcataaacac tcatgatgac cacacaaggc	480
50	agatgggcta tgtaaactgt ttcgcaattc cgtttacgat acatagtcta ctcttgctga	540
	gaatgaattc tcgtaactaa acagcacaag taggtttagt taactttaat ctcacatagc	600
55	aatctttaat caatgtgtaa cattagggag gacttgaaag agccaccaca ttttcatcga	660

ggccacgcgg agtacgatcg aggggtacagt gaataatgct agggagagct gcctatatgg 720  
aagagcccta atgtgtaaaa ttaatttttag tagtgctatc cccatgtgat tttaatagct 780  
5 tcttaggaga atgacaaaaa aaaaaaaaaa 809

<210> 41

<211> 448

10 <212> DNA

<213> CORONAVIRUS

<400> 41

15 aatgaacaca tagggctggt caagctgggg cagtacgcct ttttccagct ctactagacc 60  
acaagtgcc a tttttgaggt gttcacgtgc ctccgatagg gcctcttcca cagagtcccc 120  
gaagccacgc actagcacgt ctctaacctg aaggacaggc aaactgagtt ggacgtgtgt 180  
20 tttctcgttg acaccaagaa caaggctctc catcttacct ttcggtcaca cccggacgaa 240  
acctaggtat gctgatgatc gactgcaaca cggacgaaac cgtaagcagt ctgcagaaga 300  
gggacgagtt actcgtttct tgtcaacgac agtaaaattt attattgttt atactgcgta 360  
25 ggtgcactag gcatgcagcc gagcgacagc tacacagatt ttaaagtctg tttagagaac 420  
agatctacaa gagatcgagg ttggttgg 448

30 <210> 42

<211> 2033

<212> DNA

<213> CORONAVIRUS

35 <400> 42

atacctaggt ttcgtccggg tgtgaccgaa aggtaagatg gagagccttg ttcttggtgt 60  
caacgagaaa acacacgtcc aactcagttt gcctgtcctt cagggttagag acgtgctagt 120  
40 gcgtggcttc ggggactctg tggaagaggc cctatcggag gcacgtgaac acctcaaaaa 180  
tggcacttgt ggtctagtag agctggaaaa aggcgtactg cccagcttg aacagcccta 240  
45 tgtgttcatt aaacgttctg atgccttaag caccaatcac ggccacaagg tcgttgagct 300  
ggttgacgaa atggacggca ttcagtacgg tcgtagcggg ataacactgg gagtactcgt 360  
gccacatgtg ggcgaaaccc caattgcata ccgcaatgtt cttcttcgta agaacggtaa 420  
50 taagggagcc ggtgggcata gctatggcat cgatctaaag tcttatgact taggtgacga 480  
gcttggcact gatcccattg aagattatga acaaaactgg aacactaagc atggcagtg 540  
tgactccgt gaactcactc gtgagctcaa tggaggtgca gtcactcgt atgtcgacaa 600  
55 caatttctgt ggcccagatg ggtaccctct tgattgcata aaagattttc tcgcacgcgc 660

EP 1 694 829 B1

	gggcaagtca atgtgcactc tttccgaaca acttgattac atcgagtcga agagaggtgt	720
	ctactgctgc cgtgaccatg agcatgaaat tgcctggttc actgagcgct ctgataagag	780
5	ctacgagcac cagacaccct tcgaaattaa gagtgccaag aaatttgaca ctttcaaagg	840
	ggaatgcca aagtttgtgt ttcctcttaa ctcaaaagtc aaagtcattc aaccacgtgt	900
	tgaaaagaaa aagactgagg gtttcatggg gcgtatacgc tctgtgtacc ctgttgcac	960
10	tccacaggag tgtaacaata tgcacttgtc taccttgatg aaatgtaatc attgcatga	1020
	agtttcatgg cagacgtgcg actttctgaa agccacttgt gaacattgtg gcactgaaaa	1080
	tttagttatt gaaggaccta ctacatgtgg gtacctacct actaatgctg tagtgaaaat	1140
15	gccatgtcct gcctgtcaag acccagagat tggacctgag catagtgttg cagattatca	1200
	caaccactca aacattgaaa ctcgactccg caagggagggt aggactagat gttttggagg	1260
	ctgtgtgttt gcctatgttg gctgctataa taagcgtgcc tactgggttc ctcgtgctag	1320
20	tgctgatatt ggctcaggcc atactggcat tactggtgac aatgtggaga ccttgaatga	1380
	ggatctcctt gagatactga gtcgtgaacg tgttaacatt aacattgttg gcgattttca	1440
	tttgaatgaa gaggttgcca tcattttggc atctttctct gcttctacaa gtgcctttat	1500
25	tgacactata aagagtcttg attacaagtc tttcaaaacc attgttgagt cctgcggtaa	1560
	ctataaagtt accaagggaa agcccgtaaa aggtgcttgg aacattggac aacagagatc	1620
	agttttaaca ccactgtgtg gttttccctc acaggctgct ggtgttatca gatcaatttt	1680
30	tgcgcgca cttgatgcag caaaccactc aattcctgat ttgcaaagag cagctgtcac	1740
	catacttgat ggtatttctg aacagtcatt acgtcttgtc gacgccatgg tttatacttc	1800
	agacctgctc accaacagtg tcattattat ggcataatga actggtgggtc ttgtacaaca	1860
35	gacttctcag tggttgtcta atcttttggg cactactgtt gaaaaactca ggcctatctt	1920
	tgaatggatt gaggcgaaac ttagtgagg agttgaattt ctcaaggatg cttgggagat	1980
40	tctcaaattt ctcattacag gtgtttttga catcgtcaag ggtcaaatac agg	2033
	<210> 43	
	<211> 2018	
	<212> DNA	
45	<213> CORONAVIRUS	
	<400>	
50	ggattgagggc gaaacttagt gcaggagttg aatttctcaa ggatgcttgg gagattctca	60
	aatttctcat tacagggtgtt tttgacatcg tcaaggggtca aatacagggt gcttcagata	120
	acatcaagga ttgtgtaaaa tgcttcattg atgttgttaa caaggcactc gaaatgtgca	180
55	ttgatcaagt cactatcgct ggcgcaaagt tgcgatcact caacttaggt gaagtcttca	240
	tcgctcaaag caagggactt taccgtcagt gtatacgtgg caaggagcag ctgcaactac	300

tcatgcctct taaggcacca aaagaagtaa cctttcttga aggtgattca catgacacag 360  
 tacttacctc tgaggagggt gttctcaaga acggtgaact cgaagcactc gagacgcccg 420  
 5 ttgatagctt cacaaatgga gctatcggtg gcacaccagt ctgtgtaaataat ggcctcatgc 480  
 tcttagagat taaggacaaa gaacaatact gcgcattgtc tcctgggttta ctggctacaa 540  
 acaatgtctt tcgcttaaaa ggggggtgcac caattaaagg tgtaaccttt ggagaagata 600  
 10 ctgtttggga agttcaagggt tacaagaatg tgagaatcac atttgagctt gatgaacgtg 660  
 ttgacaaaagt gcttaatgaa aagtgcctctg tctacactgt tgaatccggt accgaagtta 720  
 ctgagtttgc atgtgttgta gcagaggctg ttgtgaagac tttaacaacca gtttctgac 780  
 15 tccttaccaa catgggtatt gatcttgatg agtgaggagt agctacattc tacttatttg 840  
 atgatgctgg tgaagaaaac ttttcatcac gtatgtattg ttctttttac cctccagatg 900  
 aggaagaaga ggacgatgca gagtgtgagg aagaagaaat tgatgaaacc tgtgaacatg 960  
 20 agtacgggtac agaggatgat tatcaagggtc tccctctgga atttggtgcc tcagctgaaa 1020  
 cagttcgagt tgaggaagaa gaagaggaag actggctgga tgatactact gagcaatcag 1080  
 agattgagcc agaaccagaa cctacacctg aagaaccagt taatcagttt actggttatt 1140  
 25 taaaacttac tgacaatggt gccattaaat gtgttgacat cgtaaggag gcacaaagtg 1200  
 ctaatcctat ggtgattgta aatgctgcta acatacacct gaaacatggt ggtggtgtag 1260  
 cagggtgcact caacaaggca accaatggtg ccatgcaaaa ggagagtgat gattacatta 1320  
 30 agctaaatgg cctctttaca gtaggagggt cttgtttgct ttctggacat aatcttgcta 1380  
 agaagtgtct gcatgttggt ggacctaac taaatgcagg tgaggacatc cagcttctta 1440  
 aggcagcata tgaataatttc aattcacagg acatcttact tgcaccattg ttgtcagcag 1500  
 35 gcatatttggt tgctaaacca cttcagtcct tacaagtgtg cgtgcagacg gttcgtacac 1560  
 aggtttatat tgcagtcaat gacaaagctc tttatgagca ggttgtcatg gattatcttg 1620  
 ataacctgaa gcctagagtg gaagcaccta aacaagagga gccaccaaac acagaagatt 1680  
 40 ccaaaactga ggagaaatct gtcgtacaga agcctgtcga tgtgaagcca aaaattaagg 1740  
 cctgcattga tgaggttacc acaacactgg aagaaactaa gtttcttacc aataagttac 1800  
 tcttgtttgc tgatatcaat ggtaagcttt accatgattc tcagaacatg cttagagggtg 1860  
 45 aagatatgtc tttccttgag aaggatgcac cttacatggt aggtgatgtt atcactagt 1920  
 gtgatatacac ttgtgttgta ataccctcca aaaaggctgg tggcactact gagatgctct 1980  
 caagagcttt gaagaaagtg ccagttgatg agtatata 2018

&lt;210&gt; 44

&lt;211&gt; 1442

&lt;212&gt; DNA

&lt;213&gt; CORONAVIRUS

&lt;400&gt; 44

ttgatgaggt taccacaaca ctggaagaaa ctaagtttct taccaataag ttactcttgt 60  
 ttgctgatat caatggtaag ctttaccatg attctcagaa catgcttaga ggtgaagata 120  
 5 tgtctttcct tgagaaggat gcaccttaca tggtaggtga tgttatcact agtggtgata 180  
 tcacttgtgt tgtaataccc tccaaaaagg ctggtggcac tactgagatg ctctcaagag 240  
 ctttgaagaa agtgccagtt gatgagtata taaccacgta ccctggacaa ggatgtgctg 300  
 10 gttatacact tgaggaagct aagactgctc ttaagaaatg caaatctgca ttttatgtac 360  
 taccttcaga agcacctaata gctaaggaag agattctagg aactgtatcc tggaatttga 420  
 gagaaatgct tgctcatgct gaagagacaa gaaaattaat gcctatatgc atggatgtta 480  
 15 gagccataat ggcaaccatc caacgtaagt ataaaggaat taaaattcaa gagggcatcg 540  
 ttgactatgg tgtccgattc ttcttttata ctagtaaaga gcctgtagct tctattatta 600  
 cgaagctgaa ctctctaaat gagccgcttg tcacaatgcc aattggttat gtgacacatg 660  
 20 gttttaatct tgaagaggct gcgcgctgta tgcgttctct taaagctcct gccgtagtgt 720  
 cagtatcatc accagatgct gttactacat ataatggata cctcacttcg tcatcaaaga 780  
 catctgagga gcactttgta gaaacagttt ctttggctgg ctcttacaga gattggtcct 840  
 25 attcaggaca gcgtacagag ttaggtgttg aatttcttaa gcgtggtgac aaaatttgtgt 900  
 accacactct ggagagcccc gtcgagtttc atcttgacgg tgaggttctt tcacttgaca 960  
 aactaaagag tctcttatcc ctgcgggagg ttaagactat aaaagtgttc acaactgtgg 1020  
 30 acaacactaa tctccacaca cagcttgttg atatgtctat gacatatgga cagcagtttg 1080  
 gtccaacata cttggatggg gctgatgtta caaaaattaa acctcatgta aatcatgagg 1140  
 gtaagacttt ctttgtacta cctagtgatg acacactacg tagtgaagct ttcgagtact 1200  
 35 accatactct tgatgagagt tttcttggtg ggtacatgtc tgctttaaac cacacaaaga 1260  
 aatggaaatt tcctcaagtt ggtggtttta cttcaattaa atgggctgat aacaattggt 1320  
 atttgtctag tgttttatta gcacttcaac agcttgaagt caaattcaat gcaccagcac 1380  
 40 ttcaagaggc ttattataga gcccgtgctg gtgatgctgc taacttttgt gcactcatac 1440  
 tc 1442

45 <210> 45  
 <211> 1050  
 <212> DNA  
 <213> CORONAVIRUS

50 <400> 45

atatgtctat gacatatgga cagcagtttg gtccaacata cttggatggg gctgatgtta 60  
 caaaaattaa acctcatgta aatcatgagg gtaagacttt ctttgtacta cctagtgatg 120  
 55

acacactacg tagtgaagct ttcgagtact accatactct tgatgagagt tttcttggtg 180  
 ggtacatgtc tgcttttaaac cacacaaaga aatggaaatt tcctcaagtt ggtggtttaa 240  
 5 cttcaattaa atgggctgat aacaattggt atttgtctag tgttttatta gcacttcaac 300  
 agcttgaagt caaattcaat gcaccagcac ttcaagaggc ttattataga gcccgtgctg 360  
 gtgatgctgc taacttttgt gcactcatac tcgcttacag taataaaact gttggcgagc 420  
 10 ttggtgatgt cagagaaact atgacccatc ttctacagca tgctaatttg gaatctgcaa 480  
 agcgagttct taatgtggtg tgtaaacatt gtggtcagaa aactactacc ttaacgggtg 540  
 tagaagctgt gatgtatatg ggtactctat cttatgataa tcttaagaca ggtgtttcca 600  
 15 ttccatgtgt gtgtggctgt gatgctacac aatatctagt acaacaagag tcttcttttg 660  
 ttatgatgtc tgcaccacct gctgagtata aattacagca aggtacattc ttatgtgcga 720  
 atgagtacac tggtaactat cagtgtggtc attacactca tataactgct aaggagaccc 780  
 20 tctatcgat tgacggagct caccttaca agatgtcaga gtacaaagga ccagtgactg 840  
 atgttttcta caaggaaaca tcttacta caaccatcaa gcctgtgtcg tataaactcg 900  
 atggagttac ttacacagag attgaaccaa aattggatgg gtattataaa aaggataatg 960  
 25 cttactatac agagcagcct atagaccttg taccaactca accattacca aatgcgagtt 1020  
 ttgataattt caaactcaca tgttctaaca 1050

30 <210> 46  
 <211> 1995  
 <212> DNA  
 <213> CORONAVIRUS

35 <400> 46

tttgtgcact catactcgct tacagtaata aaactgttgg cgagcttggt gatgtcagag 60  
 aaactatgac ccatcttcta cagcatgcta atttggaatc tgcaaagcga gttcttaatg 120  
 40 tgggtgtgtaa acattgtggt cagaaaacta ctaccttaac ggggtgtagaa gctgtgatgt 180  
 atatgggtac tctatcttat gataatctta agacaggtgt ttccattcca tgtgtgtgtg 240  
 gtcgtgatgc tacacaatat ctagtacaac aagagtcttc ttttgttatg atgtctgcac 300  
 45 cacctgctga gtataaatta cagcaaggta cattcttatg tgcgaatgag tacactggta 360  
 actatcagtg tggtcattac actcatataa ctgctaagga gaccctctat cgtattgacg 420  
 gagctcacct tacaagatg tcagagtaca aaggaccagt gactgatgtt ttctacaagg 480  
 50 aaacatctta cactacaacc atcaagcctg tgtcgtataa actcgatgga gttacttaca 540  
 cagagattga accaaaattg gatgggtatt ataaaaagga taatgcttac tatacagagc 600  
 55 agcctataga cttgtacca actcaaccat taccaaatgc gagttttgat aatttcaaac 660  
 tcacatgttc taacacaaaa tttgctgatg atttaaataa aatgacaggc ttcacaaagc 720

EP 1 694 829 B1

	cagcttcacg agagctatct gtcacattct tcccagactt gaatggcgat gtagtggcta	780
	ttgactatag acactattca gcgagtttca agaaagggtgc taaattactg cataagccaa	840
5	ttgtttggca cattaaccag gctacaacca agacaacgtt caaaccaaac acttggtggt	900
	tacgttgtct ttggagtaca aagccagtag atacttcaaa ttcatttgaa gttctggcag	960
	tagaagacac acaaggaatg gacaatcttg cttgtgaaag tcaacaaccc acctctgaag	1020
10	aagtagtgga aaatcctacc atacagaagg aagtcataga gtgtgacgtg aaaactaccg	1080
	aagttgtagg caatgtcata cttaaaccat cagatgaagg tgtaaagta acacaagagt	1140
	taggtcatga ggatcttatg gctgcttatg tggaaaacac aagcattacc attaagaaac	1200
15	ctaagagct ttcactagcc ttagggttaa aaacaattgc cactcatggt attgctgcaa	1260
	ttaatagtgt tccttggagt aaaatcttgg cttatgtcaa accattctta ggacaagcag	1320
	caattacaac atcaaattgc gctaagagat tagcacaacg tgtgtttaac aattatatgc	1380
20	cttatgtgtt tacattattg ttccaattgt gtacttttac taaaagtacc aattctagaa	1440
	ttagagcttc actacctaca actattgcta aaaatagtgt taagagtgtt gctaaattat	1500
	gtttggatgc cggcattaat tatgtgaagt cacccaaatt ttctaaattg ttcacaatcg	1560
25	ctatgtggct attgttggtta agtatttgct taggttctct aatctgtgta actgctgctt	1620
	ttggtgtact cttatctaatt tttggtgctc cttcttattg taatggcggt agagaattgt	1680
	atcttaattc gtctaacgtt actactatgg atttctgtga aggttctttt ccttgcagca	1740
30	tttgtttaag tggattagac tcccttgatt cttatccagc tcttgaaacc attcaggtga	1800
	cgatttcacg gtacaagcta gacttgacaa ttttaggtct ggccgctgag tgggttttgg	1860
	catatatgtt gttcacaata ttcttttatt tattaggtct ttcagctata atgcaggtgt	1920
35	tctttggcta ttttgctagt catttcatca gcaattcttg gctcatgtgg tttatcatta	1980
	gtattgtaca aatgg	1995
40	<210> 47 <211> 1884 <212> DNA <213> CORONAVIRUS	
45	<400> 47	
	aattcttggc tcatgtggtt tatcattagt attgtacaaa tggcaccgtt ttctgcaatg	60
50	gttaggatgt acatcttctt tgcttctttc tactacatat ggaagagcta tggtcatatc	120
	atggatgggt gcacctcttc gacttgcatg atgtgctata agcgcaatcg tgccacacgc	180
	gttgagtgt caactattgt taatggcatg aagagatctt tctatgtcta tgcaaattgga	240
55	ggccgtggct tctgcaagac tcacaattgg aattgtctca attgtgacac attttgcact	300
	ggtagtacat tcattagtga tgaagttgct cgtgatttgt cactccagtt taaaagacca	360



EP 1 694 829 B1

	atcaacccta ctgaccagtc atcgtatatatt gttgatagtg ttgctgtgaa aaatggcgcg	420
	cttcacctct acttttgacaa ggctgggtcaa aagacctatg agagacatcc gctctcccat	480
5	tttgtcaatt tagacaattt gagagctaac aacactaaag gttcactgcc tattaatgtc	540
	atagtttttg atggcaagtc caaatgcgac gagtctgctt ctaagtctgc ttctgtgtac	600
	tacagtcagc tgatgtgcca acctattctg ttgcttgacc aagctcttgt atcagacggt	660
10	ggagatagta ctgaagtttc cgtaagatg tttgatgctt atgtcgacac cttttcagca	720
	acttttagtg ttcctatgga aaaacttaag gcacttggtg ctacagctca cagcgagtta	780
	gcaaaggggtg tagcttttaga tgggtgtcctt tctacattcg tgtcagctgc cgcacaaggt	840
15	gttggttgata ccgatgttga cacaaaggat gttattgaat gtctcaaact ttcacatcac	900
	tctgacttag aagtgcagag tgacagttgt aacaatttca tgctcaccta taataagggt	960
	gaaaacatga cgcccagaga tcttggcgca tgtattgact gtaatgcaag gcatatcaat	1020
20	gcccagtag caaaaagtca caatgtttca ctcactctgga atgtaaaaga ctacatgtct	1080
	ttatctgaac agctgcgtaa acaaattcgt agtgctgcca agaagaacaa catacctttt	1140
	agactaactt gtgctacaac tagacagggt gtcaatgtca taactactaa aatctcactc	1200
25	aagggtggtg agattgttag tacttgtttt aaacttatgc ttaaggccac attattgtgc	1260
	gttcttgctg catttggtttg ttatatcggt atgccagtac atacattgtc aatccatgat	1320
	ggttacacaa atgaaatcat tggttacaaa gccattcagg atgggtgtcac tcgtgacatc	1380
30	atttctactg atgattgttt tgcaataaaa catgctgggt ttgacgcatg gtttagccag	1440
	cgtgggtggtt catacaaaaa tgacaaaagc tgccctgtag tagctgctat cattacaaga	1500
	gagattgggtt tcatagtgcc tggcctaccg ggtactgtgc tgagagcaat caatggtgac	1560
35	ttcttgcat tttctacctg tgtttttagt gctgttggtg acatttgcta cacaccttc	1620
	aaactcattg agtatagtga ttttgctacc tctgcttgcg ttcttgctgc tgagtgtaca	1680
	atttttaagg atgctatggg caaacctgtg ccataattgt atgacactaa tttgctagag	1740
40	ggttctat ttttatagtga gcttcgtcca gacactcgtt atgtgcttat ggatgggtcc	1800
	atcatacagt ttcctaacac ttacctggag ggttctgtta gagtagtaac aacttttgat	1860
	gctgagtact gtagacatgg taca	1884
45	<210> 48 <211> 2020 <212> DNA <213> CORONAVIRUS	
50	<400> 48	
	cactcgttat gtgcttatgg atggttccat catacagttt cctaacactt acctggaggg	60
55	ttctgttaga gtagtaacaa cttttgatgc tgagtactgt agacatggta catgcgaaag	120

## EP 1 694 829 B1

	gtcagaagta ggtatttgcc tatctaccag tggtagatgg gttcttaata atgagcatta	180
	cagagctcta tcaggagttt tctgtgggtg tgatgcatg aatctcatag ctaacatctt	240
5	tactcctctt gtgcaacctg tgggtgcttt agatgtgtct gcttcagtag tggctgggtg	300
	tattattgcc atattggtga cttgtgctgc ctactacttt atgaaattca gacgtgtttt	360
	tgggtgagtac aaccatgttg ttgctgctaa tgcacttttg tttttgatgt ctttcactat	420
10	actctgtctg gtaccagctt acagctttct gccgggagtc tactcagtct tttacttgta	480
	cttgacattc tatttcacca atgatgtttc attcttggct caccttcaat gggttgccat	540
	gttttctcct attgtgcctt tttggataac agcaatctat gtattctgta tttctctgaa	600
15	gcactgccat tggttcttta acaactatct taggaaaaga gtcatgttta atggagttac	660
	athtagtacc ttcgaggagg ctgctttgtg tacctttttg ctcaacaagg aaatgtacct	720
	aaaattgctg agcgagacac tgttgccact tacacagtat aacaggatc ttgctctata	780
20	taacaagtac aagtatttca gtggagcctt agatactacc agctatcgtg aagcagcttg	840
	ctgccactta gcaaaggctc taaatgactt tagcaactca ggtgctgatg ttctctacca	900
	accaccacag acatcaatca cttctgctgt tctgcagagt ggtttttagga aaatggcatt	960
25	cccgtcaggc aaagttgaag ggtgcatggt acaagtaacc tgtggaacta caactcttaa	1020
	tggattgtgg ttggatgaca cagtatactg tccaagacat gtcatttgca cagcagaaga	1080
	catgcttaat cctaactatg aagatctgct cattcgcaaa tccaaccata gctttcttgt	1140
30	tcaggctggc aatgttcaac ttcgtgttat tggccattct atgcaaaatt gtctgcttag	1200
	gcttaaagtt gatacttcta accctaagac acccaagtat aaatttgtcc gtatccaacc	1260
	tgggtcaaca ttttcagttc tagcatgcta caatggttca ccatctggtg tttatcagtg	1320
35	tgccatgaga cctaatacata ccattaaagg ttctttcctt aatggatcat gtggtagtgt	1380
	tggttttaac attgattatg attgctgtgc tttctgctat atgcatcata tggagcttcc	1440
	aacaggagta cacgctggta ctgacttaga aggtaaattc tatgggtccat ttgttgacag	1500
40	acaaactgca caggctgcag gtacagacac aaccataaca ttaaagtgtt tggcatggct	1560
	gtatgctgct gttatcaatg gtgatagggt gtttcttaat agattcacca ctactttgaa	1620
	tgactttaac cttgtggcaa tgaagtacaa ctatgaacct ttgacacaag atcatgttga	1680
45	catattggga cctctttctg ctcaaacagg aattgccgtc ttagatatgt gtgctgcttt	1740
	gaaagagctg ctgcagaatg gtatgaatgg tcgtactatc cttggtagca ctattttaga	1800
	agatgagttt acaccatttg atgttgtag acaatgctct ggtgttacct tccaaggtaa	1860
50	gttcaagaaa attgttaagg gcactcatca ttggatgctt ttaactttct tgacatcact	1920
	attgattctt gttcaaagta cacagtggtc actgttttct tttgtttacg agaatgcttt	1980
	cttgccattt actcttggtt ttatggcaat tgctgcatgt	2020

55

<210> 49  
 <211> 2040  
 <212> DNA

EP 1 694 829 B1

<213> CORONAVIRUS

<400> 49

5	agcattttcca gcctgaagac gtactgtagc agctaaactg cccagcacca tacctctatt	60
	taggttgttt aagcctttga tgaagtacaa gtatttcact ttaggccctt ttggtgtgtc	120
	tgtacaacaa ctacaagggt gttccagttc tgtgtaaatt gtacctgtac catcactctt	180
10	agggaatcta gcccatTTga gatcttggtg gtctgatagt aatgccagca caaacctacc	240
	tcccttcgaa ttgttatagt aggcaagtgc attgtcatca gtacaagctg tttgtgtggt	300
	accagccgca caggacatct gtcgtagtgc tactggactc agttcattat tctgtagttt	360
15	aacagctgag ttggctctta gagctgtaac aataagaggc caagccaaat ttggtgaatt	420
	gtccatgtta atttcactaa gttgaacaat cttgctatcc gcatcaacaa cttgctggat	480
	ttcccagagt gcagatgcat atgtaaagggt gttaccatca caagtgttct tgtagggtacc	540
20	ataatcaggg acaacaacca tgagtttggc tgctgtagtc aatgggtatga tgttgagtgg	600
	aacacaacca tcacgcgcat tgttgataat gttgttaagt gcatcattat caagcttcct	660
	aagcatagtg aagagcattg tttgcatagc actagttact tttgccctct tgtcctcaga	720
25	tcttgccctgt ttgtacattt ggggtcatagc ctgatctgcc atctttttcca acttgcgttg	780
	catggcagca tcacgggtcaa actcagattt agccacattc aaagatttct ttaacttttt	840
	gagaacgact tcagaatcac cattagctac agcctgctca taggcctcct gggcagtggtc	900
30	ataagcggca tatgatggta aagaactaaa ttctgaagca atagcctgaa gagtagcacg	960
	gttatcgagc atttcctcgc acaacctatt aatgtctaca gcaccctgca tggatagcaa	1020
	aacagacaaa agagaaacca tcttctcgaa agcttcagtt gtgtcttttg caagaagaat	1080
35	atcattgttg agttgtacac attgtgcca caatttagaa gatgactcta ctctaagttg	1140
	ttgaagaacc gagagcagta ccacagatgt gcactttacg tcagacattt tagactgtac	1200
	agtagcaacc ttgatacatg gtttacctcc aatacccaac aacttaatgt taagcttgaa	1260
40	agcatcaata ctactcttag gaggcaaaag cccctgggag ttcataatacc taaattcttg	1320
	tgtagagacc aagtagtcat aaacaccaag agtaagcctg aagtaacggg tgagtaaaca	1380
	gaaaaggcca aagtagcagc agcaacaata gcctaagaaa caataaaca gcatgataca	1440
45	ctgtaagggtg ttgccagtaa taaataacaa tgggtaatac tcaacacaca caaacactat	1500
	agctctagct aaaaacatga tagtcgtaac gacaccagaa tagttagagg ttacagaaat	1560
	aactaaggcc cacatggaaa tagcttgatc taaagcatta ccatagtaga ctttgtaaac	1620
50	aagtgtaatg acattcatca gtgtccaaac acgtctagca gcatcatcat aaacagtgcg	1680
	agctgtcatg agaataagca aaactaaagc tgaagcatac ataacacaat ccttaagcct	1740
55	ataaccagac aagctagtgt cagccaattc aagccatgtc atgatacgca tcaccagct	1800

EP 1 694 829 B1

	agcaggcatg tagaccatat taaagtaagc aactgttgca agagaaggta acagaaacaa	1860
	gcacaagaat gcgtgcttat gcttaacaag cagcatagca catgcagcaa ttgccataat	1920
5	accaagagta aatggcaaga aagcattctc gtaaacaag aaaaacagt accactgtgt	1980
	actttgaaca agaatcaata gtgatgtcaa gaaagttaaa agcatccaat gatgagtgc	2040
10	<210> 50	
	<211> 2012	
	<212> DNA	
	<213> CORONAVIRUS	
15	<400> 50	
	cttgtagggtt tgttacagac acaccaaaaag ggcctaaagt gaaatacttg tacttcatca	60
20	aaggcttaaa caacctaaat agaggatatg tgctgggcag tttagctgct acagtacgtc	120
	ttcaggctgg aaatgctaca gaagtacctg ccaattcaac tgtgctttcc ttctgtgctt	180
	ttgcagtaga ccctgctaaa gcatataagg attacctagc aagtggagga caaccaatca	240
25	ccaactgtgt gaagatgttg tgtacacaca ctggtacagg acaggcaatt actgtaacac	300
	cagaagctaa catggacca gagtcctttg gtgggtgctt atgttgctctg tattgtagat	360
	gccacattga ccatccaaat cctaaaggat tctgtgactt gaaaggtaag tacgtccaaa	420
30	tacctaccac ttgtgctaata gaccagtggt gttttacact tagaaacaca gtctgtaccg	480
	tctgcggaat gtggaaagggt tatggctgta gttgtgacca actccgcgaa cccttgatgc	540
	agtctgcgga tgcataacg tttttaaacg ggtttgcggt gtaagtgcag cccgtcttac	600
35	accgtgcggc acaggcacta gtactgatgt cgtctacagg gcttttgata tttaaacga	660
	aaaagttgct ggttttgcaa agttcctaaa aactaattgc tgtcgcttcc aggagaagga	720
	tgaggaaggc aatttattag actcttactt tgtagttaag aggcatacta tgtctaacta	780
40	ccaacatgaa gagactattt ataacttggt taaagattgt ccagcggttg ctgtccatga	840
	ctttttcaag tttagagtag atggtgacat ggtaccacat atatcacgtc agcgtctaac	900
	taaatacaca atggctgatt tagtctatgc tctacgtcat tttgatgagg gtaattgtga	960
45	tacattaaaa gaaatactcg tcacatacaa ttgctgtgat gatgattatt tcaataagaa	1020
	ggattggtat gacttcgtag agaatcctga catcttacgc gtatatgcta acttaggtga	1080
	gcgtgtacgc caatcattat taaagactgt acaattctgc gatgctatgc gtgatgcagg	1140
50	cattgtaggc gtactgacat tagataatca ggatcttaat gggaactggt acgatttcgg	1200
	tgatttcgta caagtagcac caggctgcgg agttcctatt gtggattcat attactcatt	1260
	gctgatgcc atcctcactt tgactagggc attggctgct gagtcccata tggatgctga	1320
55	tctcgcaaaa ccacttatta agtgggattt gctgaaatat gattttacgg aagagagact	1380
	ttgtctcttc gaccgttatt ttaaattatt ggaccagaca taccatccca attgtattaa	1440

EP 1 694 829 B1

	ctgtttggat gataggtgta tccttcattg tgcaaaacttt aatgtgttat tttctactgt	1500
	gtttccacct acaagttttg gaccactagt aagaaaaata tttgtagatg gtgttccttt	1560
5	tggtgtttca actggatacc attttcgtga gttaggagtc gtacataatc aggatgtaaa	1620
	cttacatagc tcgcgtctca gtttcaagga acttttagtg tatgctgctg atccagctat	1680
	gcatgcagct tctggcaatt tattgctaga taaacgcact acatgctttt cagtagctgc	1740
10	actaacaaac aatgttgctt ttcaaactgt caaaccgggt aattttaata aagactttta	1800
	tgactttgct gtgtctaaag gtttctttaa ggaaggaagt tctgttgaac taaaacactt	1860
	cttctttgct caggatggca acgctgctat cagtgattat gactattatc gttataatct	1920
15	gccaacaatg tgtgatatca gacaactcct attcgtagtt gaagttgttg ataaataactt	1980
	tgattgttac gatggtggct gtattaatgc ca	2012
20	<210> 51	
	<211> 1877	
	<212> DNA	
	<213> CORONAVIRUS	
25	<400> 51	
	gtacttcgcg tacagtggca ataccatatg acagcttaaa tgtttcctca gtggctttga	60
30	gcgtttctgc tgcgaaaagc ttgagtctct cagtacaagt gttggcaagt atgtaatcgc	120
	cagcattagt ccaatcacat gttgctatcg cattgaagtc agtgacattg tcaactgccta	180
	cacatgtggt tttgtataaa caaaaaacct gaccattagc acataatgga aaactaatgg	240
35	gaggcttatg tgacttgcaa taatagctca tacctcctag atacagttgt gtcacatcag	300
	tgacatcaca acctggggca ttgcaaacat agggattaac agacaacact aatttggtgtg	360
	atggtgaaat gacatgggtca tagcagcact tgcaacatag gaatgggtctc ctaatacagg	420
40	caccgcaacg aagtgaagtc tgtgaattgc acaatacaca agcacctaca gcctgcaaga	480
	ctgtatgtgg tgtgtacata gcctcataaa actcagggtc ccagtaccgt gaggtgttat	540
	cattagttag cattacggaa tacatgtcca acatgtggcc agtaagctca tcatgtaact	600
45	ttctaattgta ttgtaaatac aagtgaagaa catcagcata ctcttgatta ggatgttttg	660
	taagtgggta agcatcaata gccagtgaca cgaacctttc aatcataagt gtaccatctg	720
	ttttgacaat atcatcgaca aaacagcctg cgcctaatat tcttgatgga tctgggtaag	780
50	gcagggtacac gtaatcatct cttgttttaa ctagcattgt atgctgtgag caaaattcgt	840
	gaggtccttt agtaagggtca gtctcagtc aacattttgc ctcagacatg aacacattat	900
	tttgataata aagaactgcc ttaaagttct taatgctagc tactaaacct tgagccgcat	960
55	agttactgtt atagcacaca acggcatcat cagaaagaat catcatggag aaatgtttac	1020
	gcaggtaagc gtaaaactca tccacgaatt catgatcaac atccctatct ctatagagac	1080

EP 1 694 829 B1

	actcatagag cctgtgttgt agattgcgga catacttgtc agctatctta ttaccatcag	1140
	ttgaaagaag tgcatttaca ttggctgtaa cagcttgaca aatgttaaag acactattag	1200
5	cataagcagt tgtagcatca ccggatgatg ttccacctgg tttaacatat agtgagccgc	1260
	cacacatgac catctcactt aatacttgcg cacactcgtt agctaacctg tagaaacggt	1320
	gtgataagtt acagcaagtg ttatgtttgc gagcaagaac aagagaggcc attatcctaa	1380
10	gcatgttagg catggctctg tcacattttg gataatccca acccataagg tgtggagttt	1440
	ctacatcact gtaaacagtt tttaacatat tatgccagcc accgtaaac ttgcttgttc	1500
	caattaccac agtagctcct ctagtggcgg ctattgactt caataatttc tgatgaaact	1560
15	gtctattttgt catagtacta cagatagaga caccagctac ggtgcgagct ctattctttg	1620
	cactaatggc atacttaaga ttcatttgag ttatagtagg gatgacatta cgcttagtat	1680
	acgcgaaaag tgcattctga tcctcataac tcattgagtc ataataaagt ctagccttac	1740
20	cccatattatt aaatgggaaa ccagctgatt tatccagatt gttaacgatt acttggttgg	1800
	cattaataca gccaccatcg taacaatcaa agtattttatc aacaacttca actacgaata	1860
	ggagttgtct gatatca	1877
25		
	<210> 52	
	<211> 2051	
	<212> DNA	
30	<213> CORONAVIRUS	
	<400> 52	
35	tcagggtccaa tcttgacaaa gtacttcatt gatgtaagct caaagccatg cgcccaaagg	60
	acgaacacga ctctgtctga caatcctttc agtgatcac tgagcatttg tactatctta	120
	atacgcacta cattccaggg caagccttta tacatgagtg gtataagatg tttaaactgg	180
40	tcacctggtg gaggttttgc attaaactctg gtgaattctg tgttattttc agtgtcaaca	240
	taaccagtcg gtacagctac taagttaaca cctgtagaaa atcctagctg gagaggtagg	300
	ttagtaccba cagcatctct agttgcatga cagccctcta catcaaagcc aatccacgca	360
45	cgaacgtgac gaatagcttc ttcgcgggtg ataaacatat tagggtaacc attgacttgg	420
	taattcattt tgaaacccat catagagatg agtctacggt aggtcatgtc ctttggtatg	480
	cctgggtatgt caacacataa tccttcagtc ttgaacttta tatcaacgct gaggtgtgta	540
50	ggtgcctgtg taggatgaag accagtaatg atcttactac agtccttaaa aagtccagtt	600
	acattttctg cttgtaaatgt agccacattg cgacgtggta tttctagact tgtaaattgc	660
	agtttgtcat aaagatctct atcagacatt atgcacaaaa tgccaatttt tgcccttgtg	720
55	atagccacat tgaagcgggt gacattacaa gagtgtgctg tttcagtagt ttgtgtgaat	780
	atgacatagt catattcaga accctgtgat gaatcaacag tctgcgtagg caatcctaag	840

	atttttgaag ctacagcggt ctgtgaatta taagggtgaga taaaaacagc ttttctccaa	900
	gcaggattgc gtgtaagaaa ttctctttaca acgcctatatt gaggtctgtt gattgcagat	960
5	gaaacatcat gtgtaataac acctttgtag aacattttga agcattgagc tgactttatcc	1020
	ttgtgtgctt ttagcttatt gtcataaact aaagcactca cagtgtcaac aatttcagca	1080
	ggacaacggc gacaagttcc aaggaacatg tctggacctt ttgttttcat aagtctgcac	1140
10	actgaattaa aatatttctgg ttctagtgtg cttttagtca gcaatgtgcg gggggctggt	1200
	aattgagcag gatcgccaat atagacgtag tgttttgcac gaagtctagc attgacaaca	1260
	ctcaagtcac aattagtagc catagagatt tcatcaaaga ctacaatgtc agcagttggt	1320
15	tctggcaatg catttacagt gcagaaaaca tactgttcta gtgttgaatt cactttgaat	1380
	ttatcaaaac actctacgcg cgcacgcgca ggtatgattc tactacattt atctatgggc	1440
	aaatatttta atgccttttc acatagggca tcaacagctg catgagagca tgccgtatac	1500
20	actatgagcag cagatgggta atagagagca agtccgatgg caaaatgact cttaccagta	1560
	ccagggtggc cttggagtgt agagtacttt tgcattgccga ctttttgata atttgcaaca	1620
	ttgctagaaa actcatctga gatgttgagt gttgggtaca agccagtaat tctcacatag	1680
25	tgctcttggt gcactagagt aggtgcacta agtggcatta cagtgtgaga tgtcaacaca	1740
	aagtaatcac caacattcaa cttgtatgtc gtagtacctc tgtacacaac agcatcacca	1800
	tagtcacctt tttcaaagggt gtactctcca atctgtactt tactattttt agttacacgg	1860
30	taaccagtaa agacatagtt tctgttcaat ggtggtctag gttttccaac ctcccatgaa	1920
	agatgcaatt ctctgtcaga gagtacttcg cgtacagtgg caataccata tgacagctta	1980
	aatgtttcct cagtggcttt gagcgtttct gctgcgaaaa gcttgagtct ctcagtacaa	2040
35	gtgttggcaa g	2051
	<210> 53	
40	<211> 2075	
	<212> DNA	
	<213> CORONAVIRUS	
	<400> 53	
45		
	tgctttagt tttgggtaga aggtttcaac atgtccatcc ttacaccaaa gcatgaatga	60
	aatttcagca tagtcaattg taaccttgac cacttttgaa atcactgaca aatcttgtga	120
50	ctttattatc tcgacaaagt catcaagtaa aagatcaatc acagaacaca cacattttga	180
	tgaacctgtt tgcgcatctg ttatgaagta atttttcact gtgctgtcca tagggataaa	240
	atcctctaata ttaagtgggtg aatcttgtga gcgcttggct aagcctatca ttaaatagaag	300
55	accgccaagt tgtccatgac tgaaatctcc ataaacgatg tgttcgaagg catagccctc	360
	gagcttatat cgctgtatga attcatccat agcgagctcg agaaagtcag tttccatttg	420

EP 1 694 829 B1

	tgatctgggc	ttaaaatcct	ctaagtctct	gctctgagta	aagtaggttt	caggcaactg	480
	ttgaataatg	ccgtctactt	tcttaaagta	gttaaactgt	gtttttactg	attctccaat	540
5	taatgtgact	ccattgacgc	tagcttgtgc	tgggtcccttt	gaagggtgta	gacctttgac	600
	tgaaccttct	gttattaaaa	caccattacg	ggcgtttcta	aaaagggtcta	cctgtccttc	660
	cactctacca	tcaaacaaga	cagtaagtga	agaacaagca	ctctcagtag	gtttccttggc	720
10	aatgtcagtc	attgtgcaga	cacctattgt	agatacatgt	gctgggggctt	ctcttttgta	780
	gtcccagatt	acagtattag	cagcgatatc	aacacccaaa	ttattgagta	tcttaatctc	840
	tggcactggg	ttaatgttac	gcttagccca	aagctcaa	aatgcaacatt	caggaagtgt	900
15	tgtcttattt	tcaaagatct	ccacatcaat	accatctacc	tttgtgtaaa	cagcattatt	960
	aatgatggaa	acagggtgctt	cgccggcggt	tccatcaaag	tgtcctttat	taacaacatt	1020
	ataagccaca	ttttctaaac	tctgtaacct	ggtaaatgta	ttccacaggt	tataagtatc	1080
20	aaattgtttg	taaatccata	ggctaaatcc	agcagaaatc	atcatattat	atgcatccaa	1140
	gtactgtcgg	tactcatttg	catgggtgtc	gcaaacagca	ccacctaaat	tgcacgtgt	1200
	aatacacgta	gcagatttga	gtggaacata	atcaatatcc	gacactactt	gtttgccatg	1260
25	agactcaca	ggactatcag	aatagtaaaa	gaaaggcaat	tgctttaaat	tagtaaatgc	1320
	acttttatcg	aaagctggag	tgtggaatgc	atgcttattc	acatacaaac	taccaccatc	1380
	acagcctggg	aagttcaagt	ttgacaagac	tcttgtgtca	aacctacaca	caattgcatt	1440
30	ggctgggtaa	cgatcaacgt	tacaattcca	aaacaaaca	acaccatcag	tgaatttatc	1500
	gtgatgtgta	gcataagaat	agaagagttc	ctctattttg	taagctttgt	cactacatgg	1560
	ctgagcatcg	tagaacttcc	attctacttc	agcctgaggc	acacacttga	tagcctttgg	1620
35	atttccaatg	tcatgaagaa	ctggaaactt	atcagcaagc	aatgcagact	tcacaacat	1680
	gtgttgtact	tttctgcaag	cagaattaac	cctcagttca	tctcctataa	taggggtattc	1740
	aacagaccaa	tcaacgcgct	taacaaagca	ctcatggact	gctaaacatc	tagtcatgat	1800
40	agcatcaca	ctagccacat	gtgcatttcc	atgtacctgg	caatgttggt	catgggttact	1860
	ctgaagggtta	cccgtaaagc	cccactgctg	aacatcaatc	ataaatgggt	tatagacata	1920
	gtcaaaaccc	acagaatgat	tccagcaggc	ataagtatct	gatgaagtag	aaaagcaagt	1980
45	tgcacgtttg	tcacacagac	aacacgttct	ttcaggtcca	atcttgacaa	agtacttcat	2040
	tgatgtaagc	tcaaagccat	gcgcccgaag	gacga			2075
50	<210> 54						
	<211> 1891						
	<212> DNA						
	<213> CORONAVIRUS						
55	<400> 54						



EP 1 694 829 B1

	aagattcacc	acttaaatta	gaggatttta	tccctatgga	cagcacagt	aaaaattact	60
	tcataacaga	tgcgcaaaca	ggttcatcaa	aatgtgtgtg	ttctgtgatt	gatcttttac	120
5	ttgatgactt	tgtcgagata	ataaagtcac	aagatttgtc	agtgatttca	aaagtgggtca	180
	aggttacaat	tgactatgct	gaaatttcat	tcatgctttg	gtgtaaggat	ggacatgttg	240
	aaaccttcta	cccaaaacta	caagcaagtc	aagcgtggca	accaggtgtt	gcatgccta	300
10	acttgtacaa	gatgcaaaga	atgcttcttg	aaaagtgtga	ccttcagaat	tatggtgaaa	360
	atgctgttat	accaaaagga	ataatgatga	atgtcgcaaa	gtatactcaa	ctgtgtcaat	420
	acttaaatac	acttacttta	gctgtaccct	acaacatgag	agttattcac	tttgggtgctg	480
15	gctctgataa	aggagttgca	ccagggtacag	ctgtgctcag	acaatgggtg	ccaactggca	540
	cactacttgt	cgattcagat	cttaatgact	tcgtctccga	cgcagattct	actttaattg	600
	gagactgtgc	aacagtacat	acggctaata	aatgggacct	tattattagc	gatatgtatg	660
20	accctaggac	caaacatgtg	acaaaagaga	atgactctaa	agaagggttt	ttcacttatc	720
	tgtgtggatt	tataaagcaa	aaactagccc	tgggtgggtc	tatagctgta	aagataacag	780
	agcattcttg	gaatgctgac	ctttacaagc	ttatgggccca	tttctcatgg	tggacagctt	840
25	ttgtttacaaa	tgtaaatgca	tcatcatcgg	aagcattttt	aattggggct	aactatcttg	900
	gcaagccgaa	ggaacaaatt	gatggctata	ccatgcatgc	taactacatt	ttctggagga	960
	acacaaatcc	tatccagttg	tcttcctatt	cactctttga	catgagcaaa	tttcctctta	1020
30	aattaagagg	aactgctgta	atgtctctta	aggagaatca	aatcaatgat	atgattttatt	1080
	ctcttctgga	aaaaggtagg	cttatcatta	gagaaaacaa	cagagttgtg	gtttcaagtg	1140
	atattcttgt	taacaactaa	acgaacatgt	ttattttctt	attatttctt	actctcacta	1200
35	gtggtagtga	ccttgaccgg	tgcaccactt	ttgatgatgt	tcaagctcct	aattacactc	1260
	aacatacttc	atctatgagg	ggggtttact	atcctgatga	aatttttaga	tcagacactc	1320
	tttatttaac	tcaggattta	tttcttccat	tttattctaa	tgttacaggg	tttcatacta	1380
40	ttaatcatac	gtttggcaac	cctgtcatac	cttttaagga	tggtatttat	tttgctgcca	1440
	cagagaaatc	aaatgttgtc	cgtgggtggg	tttttggttc	taccatgaac	aacaagtcac	1500
	agtcggtgat	tattattaac	aattctacta	atgttggtat	acgagcatgt	aactttgaat	1560
45	tgtgtgacaa	ccctttcttt	gctgtttcta	aacctatggg	tacacagaca	catactatga	1620
	tattcgataa	tgcatttaat	tgacttttcg	agtacatatc	tgatgccttt	tcgcttgatg	1680
	tttcagaaaa	gtcaggtaat	tttaaacact	tacgagagtt	tgtgtttaaa	aataaagatg	1740
50	ggtttctcta	tgtttataag	ggctatcaac	ctatagatgt	agttcgtgat	ctaccttctg	1800
	gttttaacac	tttgaaacct	atttttaagt	tgcttcttgg	tattaacatt	acaaatttta	1860
	gagccattct	tacagccttt	tcacctgctc	a			1891

55

<210> 55  
 <211> 32  
 <212> DNA

&lt;213&gt; ARTIFICIAL SEQUENCE

&lt;220&gt;

&lt;223&gt; amorce N sens

5

&lt;400&gt; 55

cccatatgtc tgataatgga cccaatcaa ac 32

&lt;210&gt; 56

10

&lt;211&gt; 32

&lt;212&gt; DNA

&lt;213&gt; artificial sequence

&lt;220&gt;

15

&lt;223&gt; amorce N antisens

&lt;400&gt; 56

cccccggtg cctgagttga atcagcagaa gc 32

20

&lt;210&gt; 57

&lt;211&gt; 31

&lt;212&gt; DNA

&lt;213&gt; artificial sequence

25

&lt;220&gt;

&lt;223&gt; amorce Sc sens

&lt;400&gt; 57

30

cccatatgag tgacctgac cggcgaccaa c 31

&lt;210&gt; 58

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; artificial sequence

35

&lt;220&gt;

&lt;223&gt; amorce SL sens

&lt;400&gt; 58

40

cccatatgaa acctgcacc ccacctgctc 30

&lt;210&gt; 59

&lt;211&gt; 33

&lt;212&gt; DNA

45

&lt;213&gt; amorce Sc et SL antisens

&lt;400&gt; 59

cccccggtt taatatattg ctcatattt ccc 33

50

&lt;210&gt; 60

&lt;211&gt; 16

&lt;212&gt; DNA

&lt;213&gt; amorce sens série 1

55

&lt;400&gt; 60

ggcatcgat ggggtg 16

&lt;210&gt; 61

<211> 16  
 <212> DNA  
 <213> amorce antisens série 2 (28774-28759)

5      <400> 61  
       cagtttcacc acctcc            16

<210> 62  
 <211> 16  
 10      <212> DNA  
       <213> amorce sens série 2 (28375-28390)

<400> 62  
 15      ggctactacc gaagag            16

<210> 63  
 <211> 16  
 <212> DNA  
 20      <213> amorce antisens série 2 (28702-28687)

<400> 63  
       aattaccgcg actacg            16

<210> 64  
 25      <211> 26  
       <212> DNA  
       <213> sonde 1/série 1 (28561-28586)

<400> 64  
 30      ggcaccgcga atcctaataa caatgc            26

<210> 65  
 <211> 21  
 <212> DNA  
 35      <213> sonde 2/série 1 (28588-28608)

<400> 65  
       gccaccgtgc tacaacttc t            21

40      <210> 66  
       <211> 23  
       <212> DNA  
       <213> sonde 1/série 2 /sonde N/FL (28541-28563)

45      <400> 66  
       atacacccaa agaccacatt ggc            23

<210> 67  
 <211> 25  
 50      <212> DNA  
       <213> sonde 2/série 2/sonde SARS/N/LC705 (28565-28589)

<400> 67  
 55      cccgcaatcc taataacaat gctgc            25

<210> 68  
 <211> 30  
 <212> DNA

<213> artificial sequence

<220>

<223> amorce ancre 14T

5

<400> 68

agatgaattc ggtaccttt tttttttt 30

<210> 69

10

<211> 13

<212> PRT

<213> artificial sequence

<220>

15

<223> peptide M2-14

<400> 69

20

Ala Asp Asn Gly Thr Ile Thr Val Glu Glu Leu Lys Gln  
1 5 10

<210> 70

25

<211> 12

<212> PRT

<213> artificial sequence

<220>

30

<223> peptide E1-12

<400> 70

35

Met Tyr Ser Phe Val Ser Glu Glu Thr Gly Thr Leu  
1 5 10

<210> 71

40

<211> 24

<212> PRT

<213> artificial sequence

<220>

45

<223> peptide E53-72

<400> 71

50

Lys Pro Thr Val Tyr Val Tyr Ser Arg Val Lys Asn Leu Asn Ser Ser  
1 5 10 15

Glu Gly Val Pro Asp Leu Leu Val  
20

55

<210> 72

<211> 153

# EP 1 694 829 B1

<212> DNA  
<213> CORONAVIRUS

<400> 72

5 gatattaggt ttttacctac ccaggaaaag ccaaccaacc tcgatctctt gtagatctgt 60  
tctctaaacg aacttttaaaa tctgtgtagc tgtcgctcgg ctgcatgcct agtgcaccta 120  
10 cgcagtataa acaataataa attttactgt cgt 153

<210> 73  
<211> 410  
15 <212> DNA  
<213> CORONAVIRUS

<400> 73

20 ttctccagac aacttcaaaa ttccatgagt ggagcttctg ctgattcaac tcaggcataa 60  
acactcatga tgaccacaca aggcagatgg gctatgtaaa cgttttcgca attccgttta 120  
cgatacatag tctactcttg tgcagaatga attctcgtaa ctaaacagca caagtagggt 180  
25 tagttaactt taatctcaca tagcaatctt taatcaatgt gtaacattag ggaggacttg 240  
aaagagccac cacattttca tcgaggccac gcggagtagc atcgagggta cagtgaataa 300  
tgctagggag agctgcctat atggaagagc cctaattgtgt aaaattaatt ttagtagtgc 360  
30 tatcccatg tgattttaat agcttcttag gagaatgaca aaaaaaaaaa 410

<210> 74  
<211> 4382  
35 <212> PRT  
<213> CORONAVIRUS

<400> 74

40 Met Glu Ser Leu Val Leu Gly Val Asn Glu Lys Thr His Val Gln Leu  
1 5 10 15  
45 Ser Leu Pro Val Leu Gln Val Arg Asp Val Leu Val Arg Gly Phe Gly  
20 25 30

50

55

EP 1 694 829 B1

	Asp	Ser	Val	Glu	Glu	Ala	Leu	Ser	Glu	Ala	Arg	Glu	His	Leu	Lys	Asn
			35					40					45			
5	Gly	Thr	Cys	Gly	Leu	Val	Glu	Leu	Glu	Lys	Gly	Val	Leu	Pro	Gln	Leu
	50						55					60				
	Glu	Gln	Pro	Tyr	Val	Phe	Ile	Lys	Arg	Ser	Asp	Ala	Leu	Ser	Thr	Asn
10	65					70					75					80
	His	Gly	His	Lys	Val	Val	Glu	Leu	Val	Ala	Glu	Met	Asp	Gly	Ile	Gln
					85					90					95	
15	Tyr	Gly	Arg	Ser	Gly	Ile	Thr	Leu	Gly	Val	Leu	Val	Pro	His	Val	Gly
				100					105					110		
	Glu	Thr	Pro	Ile	Ala	Tyr	Arg	Asn	Val	Leu	Leu	Arg	Lys	Asn	Gly	Asn
20			115					120					125			
	Lys	Gly	Ala	Gly	Gly	His	Ser	Tyr	Gly	Ile	Asp	Leu	Lys	Ser	Tyr	Asp
		130					135					140				
25	Leu	Gly	Asp	Glu	Leu	Gly	Thr	Asp	Pro	Ile	Glu	Asp	Tyr	Glu	Gln	Asn
	145					150					155					160
	Trp	Asn	Thr	Lys	His	Gly	Ser	Gly	Ala	Leu	Arg	Glu	Leu	Thr	Arg	Glu
30					165					170					175	
	Leu	Asn	Gly	Gly	Ala	Val	Thr	Arg	Tyr	Val	Asp	Asn	Asn	Phe	Cys	Gly
				180					185					190		
35	Pro	Asp	Gly	Tyr	Pro	Leu	Asp	Cys	Ile	Lys	Asp	Phe	Leu	Ala	Arg	Ala
			195					200					205			
	Gly	Lys	Ser	Met	Cys	Thr	Leu	Ser	Glu	Gln	Leu	Asp	Tyr	Ile	Glu	Ser
40		210					215					220				
	Lys	Arg	Gly	Val	Tyr	Cys	Cys	Arg	Asp	His	Glu	His	Glu	Ile	Ala	Trp
	225					230					235				240	
45	Phe	Thr	Glu	Arg	Ser	Asp	Lys	Ser	Tyr	Glu	His	Gln	Thr	Pro	Phe	Glu
					245					250					255	
	Ile	Lys	Ser	Ala	Lys	Lys	Phe	Asp	Thr	Phe	Lys	Gly	Glu	Cys	Pro	Lys
50				260					265					270		
	Phe	Val	Phe	Pro	Leu	Asn	Ser	Lys	Val	Lys	Val	Ile	Gln	Pro	Arg	Val
			275					280					285			
55	Glu	Lys	Lys	Lys	Thr	Glu	Gly	Phe	Met	Gly	Arg	Ile	Arg	Ser	Val	Tyr
		290					295					300				

EP 1 694 829 B1

Pro Val Ala Ser Pro Gln Glu Cys Asn Asn Met His Leu Ser Thr Leu  
305 310 315 320

5 Met Lys Cys Asn His Cys Asp Glu Val Ser Trp Gln Thr Cys Asp Phe  
325 330 335

10 Leu Lys Ala Thr Cys Glu His Cys Gly Thr Glu Asn Leu Val Ile Glu  
340 345 350

Gly Pro Thr Thr Cys Gly Tyr Leu Pro Thr Asn Ala Val Val Lys Met  
355 360 365

15 Pro Cys Pro Ala Cys Gln Asp Pro Glu Ile Gly Pro Glu His Ser Val  
370 375 380

20 Ala Asp Tyr His Asn His Ser Asn Ile Glu Thr Arg Leu Arg Lys Gly  
385 390 395 400

Gly Arg Thr Arg Cys Phe Gly Gly Cys Val Phe Ala Tyr Val Gly Cys  
405 410 415

25 Tyr Asn Lys Arg Ala Tyr Trp Val Pro Arg Ala Ser Ala Asp Ile Gly  
420 425 430

30 Ser Gly His Thr Gly Ile Thr Gly Asp Asn Val Glu Thr Leu Asn Glu  
435 440 445

Asp Leu Leu Glu Ile Leu Ser Arg Glu Arg Val Asn Ile Asn Ile Val  
450 455 460

35 Gly Asp Phe His Leu Asn Glu Glu Val Ala Ile Ile Leu Ala Ser Phe  
465 470 475 480

40 Ser Ala Ser Thr Ser Ala Phe Ile Asp Thr Ile Lys Ser Leu Asp Tyr  
485 490 495

Lys Ser Phe Lys Thr Ile Val Glu Ser Cys Gly Asn Tyr Lys Val Thr  
500 505 510

45 Lys Gly Lys Pro Val Lys Gly Ala Trp Asn Ile Gly Gln Gln Arg Ser  
515 520 525

50 Val Leu Thr Pro Leu Cys Gly Phe Pro Ser Gln Ala Ala Gly Val Ile  
530 535 540

Arg Ser Ile Phe Ala Arg Thr Leu Asp Ala Ala Asn His Ser Ile Pro  
545 550 555 560

55 Asp Leu Gln Arg Ala Ala Val Thr Ile Leu Asp Gly Ile Ser Glu Gln  
565 570 575

EP 1 694 829 B1

Ser Leu Arg Leu Val Asp Ala Met Val Tyr Thr Ser Asp Leu Leu Thr  
 580 585 590  
 5  
 Asn Ser Val Ile Ile Met Ala Tyr Val Thr Gly Gly Leu Val Gln Gln  
 595 600 605  
 Thr Ser Gln Trp Leu Ser Asn Leu Leu Gly Thr Thr Val Glu Lys Leu  
 610 615 620  
 10  
 Arg Pro Ile Phe Glu Trp Ile Glu Ala Lys Leu Ser Ala Gly Val Glu  
 625 630 635 640  
 15  
 Phe Leu Lys Asp Ala Trp Glu Ile Leu Lys Phe Leu Ile Thr Gly Val  
 645 650 655  
 Phe Asp Ile Val Lys Gly Gln Ile Gln Val Ala Ser Asp Asn Ile Lys  
 660 665 670  
 20  
 Asp Cys Val Lys Cys Phe Ile Asp Val Val Asn Lys Ala Leu Glu Met  
 675 680 685  
 25  
 Cys Ile Asp Gln Val Thr Ile Ala Gly Ala Lys Leu Arg Ser Leu Asn  
 690 695 700  
 Leu Gly Glu Val Phe Ile Ala Gln Ser Lys Gly Leu Tyr Arg Gln Cys  
 705 710 715 720  
 30  
 Ile Arg Gly Lys Glu Gln Leu Gln Leu Leu Met Pro Leu Lys Ala Pro  
 725 730 735  
 35  
 Lys Glu Val Thr Phe Leu Glu Gly Asp Ser His Asp Thr Val Leu Thr  
 740 745 750  
 Ser Glu Glu Val Val Leu Lys Asn Gly Glu Leu Glu Ala Leu Glu Thr  
 755 760 765  
 40  
 Pro Val Asp Ser Phe Thr Asn Gly Ala Ile Val Gly Thr Pro Val Cys  
 770 775 780  
 45  
 Val Asn Gly Leu Met Leu Leu Glu Ile Lys Asp Lys Glu Gln Tyr Cys  
 785 790 795 800  
 Ala Leu Ser Pro Gly Leu Leu Ala Thr Asn Asn Val Phe Arg Leu Lys  
 805 810 815  
 50  
 Gly Gly Ala Pro Ile Lys Gly Val Thr Phe Gly Glu Asp Thr Val Trp  
 820 825 830  
 55  
 Glu Val Gln Gly Tyr Lys Asn Val Arg Ile Thr Phe Glu Leu Asp Glu  
 835 840 845



EP 1 694 829 B1

Arg Val Asp Lys Val Leu Asn Glu Lys Cys Ser Val Tyr Thr Val Glu  
 850 855 860  
 5 Ser Gly Thr Glu Val Thr Glu Phe Ala Cys Val Val Ala Glu Ala Val  
 865 870 875 880  
 Val Lys Thr Leu Gln Pro Val Ser Asp Leu Leu Thr Asn Met Gly Ile  
 885 890 895  
 10 Asp Leu Asp Glu Trp Ser Val Ala Thr Phe Tyr Leu Phe Asp Asp Ala  
 900 905 910  
 15 Gly Glu Glu Asn Phe Ser Ser Arg Met Tyr Cys Ser Phe Tyr Pro Pro  
 915 920 925  
 20 Asp Glu Glu Glu Glu Asp Asp Ala Glu Cys Glu Glu Glu Glu Ile Asp  
 930 935 940  
 Glu Thr Cys Glu His Glu Tyr Gly Thr Glu Asp Asp Tyr Gln Gly Leu  
 945 950 955 960  
 25 Pro Leu Glu Phe Gly Ala Ser Ala Glu Thr Val Arg Val Glu Glu Glu  
 965 970 975  
 Glu Glu Glu Asp Trp Leu Asp Asp Thr Thr Glu Gln Ser Glu Ile Glu  
 980 985 990  
 30 Pro Glu Pro Glu Pro Thr Pro Glu Glu Pro Val Asn Gln Phe Thr Gly  
 995 1000 1005  
 35 Tyr Leu Lys Leu Thr Asp Asn Val Ala Ile Lys Cys Val Asp Ile  
 1010 1015 1020  
 Val Lys Glu Ala Gln Ser Ala Asn Pro Met Val Ile Val Asn Ala  
 1025 1030 1035  
 40 Ala Asn Ile His Leu Lys His Gly Gly Gly Val Ala Gly Ala Leu  
 1040 1045 1050  
 45 Asn Lys Ala Thr Asn Gly Ala Met Gln Lys Glu Ser Asp Asp Tyr  
 1055 1060 1065  
 Ile Lys Leu Asn Gly Pro Leu Thr Val Gly Gly Ser Cys Leu Leu  
 1070 1075 1080  
 50 Ser Gly His Asn Leu Ala Lys Lys Cys Leu His Val Val Gly Pro  
 1085 1090 1095  
 55 Asn Leu Asn Ala Gly Glu Asp Ile Gln Leu Leu Lys Ala Ala Tyr  
 1100 1105 1110

EP 1 694 829 B1

	Glu	Asn 1115	Phe	Asn	Ser	Gln	Asp 1120	Ile	Leu	Leu	Ala	Pro 1125	Leu	Leu	Ser
5	Ala	Gly 1130	Ile	Phe	Gly	Ala	Lys 1135	Pro	Leu	Gln	Ser	Leu 1140	Gln	Val	Cys
10	Val	Gln 1145	Thr	Val	Arg	Thr	Gln 1150	Val	Tyr	Ile	Ala	Val 1155	Asn	Asp	Lys
	Ala	Leu 1160	Tyr	Glu	Gln	Val	Val 1165	Met	Asp	Tyr	Leu	Asp 1170	Asn	Leu	Lys
15	Pro	Arg 1175	Val	Glu	Ala	Pro	Lys 1180	Gln	Glu	Glu	Pro	Pro 1185	Asn	Thr	Glu
20	Asp	Ser 1190	Lys	Thr	Glu	Glu	Lys 1195	Ser	Val	Val	Gln	Lys 1200	Pro	Val	Asp
	Val	Lys 1205	Pro	Lys	Ile	Lys	Ala 1210	Cys	Ile	Asp	Glu	Val 1215	Thr	Thr	Thr
25	Leu	Glu 1220	Glu	Thr	Lys	Phe	Leu 1225	Thr	Asn	Lys	Leu	Leu 1230	Leu	Phe	Ala
30	Asp	Ile 1235	Asn	Gly	Lys	Leu	Tyr 1240	His	Asp	Ser	Gln	Asn 1245	Met	Leu	Arg
	Gly	Glu 1250	Asp	Met	Ser	Phe	Leu 1255	Glu	Lys	Asp	Ala	Pro 1260	Tyr	Met	Val
35	Gly	Asp 1265	Val	Ile	Thr	Ser	Gly 1270	Asp	Ile	Thr	Cys	Val 1275	Val	Ile	Pro
40	Ser	Lys 1280	Lys	Ala	Gly	Gly	Thr 1285	Thr	Glu	Met	Leu	Ser 1290	Arg	Ala	Leu
	Lys	Lys 1295	Val	Pro	Val	Asp	Glu 1300	Tyr	Ile	Thr	Thr	Tyr 1305	Pro	Gly	Gln
45	Gly	Cys 1310	Ala	Gly	Tyr	Thr	Leu 1315	Glu	Glu	Ala	Lys	Thr 1320	Ala	Leu	Lys
50	Lys	Cys 1325	Lys	Ser	Ala	Phe	Tyr 1330	Val	Leu	Pro	Ser	Glu 1335	Ala	Pro	Asn
	Ala	Lys 1340	Glu	Glu	Ile	Leu	Gly 1345	Thr	Val	Ser	Trp	Asn 1350	Leu	Arg	Glu
55	Met	Leu 1355	Ala	His	Ala	Glu	Glu 1360	Thr	Arg	Lys	Leu	Met 1365	Pro	Ile	Cys

EP 1 694 829 B1

	Met	Asp	Val	Arg	Ala	Ile	Met	Ala	Thr	Ile	Gln	Arg	Lys	Tyr	Lys
		1370					1375					1380			
5	Gly	Ile	Lys	Ile	Gln	Glu	Gly	Ile	Val	Asp	Tyr	Gly	Val	Arg	Phe
		1385					1390					1395			
	Phe	Phe	Tyr	Thr	Ser	Lys	Glu	Pro	Val	Ala	Ser	Ile	Ile	Thr	Lys
10		1400					1405					1410			
	Leu	Asn	Ser	Leu	Asn	Glu	Pro	Leu	Val	Thr	Met	Pro	Ile	Gly	Tyr
		1415					1420					1425			
15	Val	Thr	His	Gly	Phe	Asn	Leu	Glu	Glu	Ala	Ala	Arg	Cys	Met	Arg
		1430					1435					1440			
	Ser	Leu	Lys	Ala	Pro	Ala	Val	Val	Ser	Val	Ser	Ser	Pro	Asp	Ala
20		1445					1450					1455			
	Val	Thr	Thr	Tyr	Asn	Gly	Tyr	Leu	Thr	Ser	Ser	Ser	Lys	Thr	Ser
		1460					1465					1470			
25	Glu	Glu	His	Phe	Val	Glu	Thr	Val	Ser	Leu	Ala	Gly	Ser	Tyr	Arg
		1475					1480					1485			
	Asp	Trp	Ser	Tyr	Ser	Gly	Gln	Arg	Thr	Glu	Leu	Gly	Val	Glu	Phe
30		1490					1495					1500			
	Leu	Lys	Arg	Gly	Asp	Lys	Ile	Val	Tyr	His	Thr	Leu	Glu	Ser	Pro
		1505					1510					1515			
35	Val	Glu	Phe	His	Leu	Asp	Gly	Glu	Val	Leu	Ser	Leu	Asp	Lys	Leu
		1520					1525					1530			
	Lys	Ser	Leu	Leu	Ser	Leu	Arg	Glu	Val	Lys	Thr	Ile	Lys	Val	Phe
40		1535					1540					1545			
	Thr	Thr	Val	Asp	Asn	Thr	Asn	Leu	His	Thr	Gln	Leu	Val	Asp	Met
		1550					1555					1560			
45	Ser	Met	Thr	Tyr	Gly	Gln	Gln	Phe	Gly	Pro	Thr	Tyr	Leu	Asp	Gly
		1565					1570					1575			
	Ala	Asp	Val	Thr	Lys	Ile	Lys	Pro	His	Val	Asn	His	Glu	Gly	Lys
50		1580					1585					1590			
	Thr	Phe	Phe	Val	Leu	Pro	Ser	Asp	Asp	Thr	Leu	Arg	Ser	Glu	Ala
		1595					1600					1605			
55	Phe	Glu	Tyr	Tyr	His	Thr	Leu	Asp	Glu	Ser	Phe	Leu	Gly	Arg	Tyr
		1610					1615					1620			

EP 1 694 829 B1

	Met	Ser	Ala	Leu	Asn	His	Thr	Lys	Lys	Trp	Lys	Phe	Pro	Gln	Val
		1625					1630					1635			
5	Gly	Gly	Leu	Thr	Ser	Ile	Lys	Trp	Ala	Asp	Asn	Asn	Cys	Tyr	Leu
		1640					1645					1650			
10	Ser	Ser	Val	Leu	Leu	Ala	Leu	Gln	Gln	Leu	Glu	Val	Lys	Phe	Asn
		1655					1660					1665			
	Ala	Pro	Ala	Leu	Gln	Glu	Ala	Tyr	Tyr	Arg	Ala	Arg	Ala	Gly	Asp
		1670					1675					1680			
15	Ala	Ala	Asn	Phe	Cys	Ala	Leu	Ile	Leu	Ala	Tyr	Ser	Asn	Lys	Thr
		1685					1690					1695			
20	Val	Gly	Glu	Leu	Gly	Asp	Val	Arg	Glu	Thr	Met	Thr	His	Leu	Leu
		1700					1705					1710			
	Gln	His	Ala	Asn	Leu	Glu	Ser	Ala	Lys	Arg	Val	Leu	Asn	Val	Val
		1715					1720					1725			
25	Cys	Lys	His	Cys	Gly	Gln	Lys	Thr	Thr	Thr	Leu	Thr	Gly	Val	Glu
		1730					1735					1740			
30	Ala	Val	Met	Tyr	Met	Gly	Thr	Leu	Ser	Tyr	Asp	Asn	Leu	Lys	Thr
		1745					1750					1755			
	Gly	Val	Ser	Ile	Pro	Cys	Val	Cys	Gly	Arg	Asp	Ala	Thr	Gln	Tyr
		1760					1765					1770			
35	Leu	Val	Gln	Gln	Glu	Ser	Ser	Phe	Val	Met	Met	Ser	Ala	Pro	Pro
		1775					1780					1785			
40	Ala	Glu	Tyr	Lys	Leu	Gln	Gln	Gly	Thr	Phe	Leu	Cys	Ala	Asn	Glu
		1790					1795					1800			
	Tyr	Thr	Gly	Asn	Tyr	Gln	Cys	Gly	His	Tyr	Thr	His	Ile	Thr	Ala
		1805					1810					1815			
45	Lys	Glu	Thr	Leu	Tyr	Arg	Ile	Asp	Gly	Ala	His	Leu	Thr	Lys	Met
		1820					1825					1830			
50	Ser	Glu	Tyr	Lys	Gly	Pro	Val	Thr	Asp	Val	Phe	Tyr	Lys	Glu	Thr
		1835					1840					1845			
	Ser	Tyr	Thr	Thr	Thr	Ile	Lys	Pro	Val	Ser	Tyr	Lys	Leu	Asp	Gly
		1850					1855					1860			
55	Val	Thr	Tyr	Thr	Glu	Ile	Glu	Pro	Lys	Leu	Asp	Gly	Tyr	Tyr	Lys
		1865					1870					1875			

EP 1 694 829 B1

	Lys	Asp	Asn	Ala	Tyr	Tyr	Thr	Glu	Gln	Pro	Ile	Asp	Leu	Val	Pro
		1880					1885					1890			
5	Thr	Gln	Pro	Leu	Pro	Asn	Ala	Ser	Phe	Asp	Asn	Phe	Lys	Leu	Thr
		1895					1900					1905			
10	Cys	Ser	Asn	Thr	Lys	Phe	Ala	Asp	Asp	Leu	Asn	Gln	Met	Thr	Gly
		1910					1915					1920			
	Phe	Thr	Lys	Pro	Ala	Ser	Arg	Glu	Leu	Ser	Val	Thr	Phe	Phe	Pro
		1925					1930					1935			
15	Asp	Leu	Asn	Gly	Asp	Val	Val	Ala	Ile	Asp	Tyr	Arg	His	Tyr	Ser
		1940					1945					1950			
20	Ala	Ser	Phe	Lys	Lys	Gly	Ala	Lys	Leu	Leu	His	Lys	Pro	Ile	Val
		1955					1960					1965			
	Trp	His	Ile	Asn	Gln	Ala	Thr	Thr	Lys	Thr	Thr	Phe	Lys	Pro	Asn
		1970					1975					1980			
25	Thr	Trp	Cys	Leu	Arg	Cys	Leu	Trp	Ser	Thr	Lys	Pro	Val	Asp	Thr
		1985					1990					1995			
30	Ser	Asn	Ser	Phe	Glu	Val	Leu	Ala	Val	Glu	Asp	Thr	Gln	Gly	Met
		2000					2005					2010			
	Asp	Asn	Leu	Ala	Cys	Glu	Ser	Gln	Gln	Pro	Thr	Ser	Glu	Glu	Val
		2015					2020					2025			
35	Val	Glu	Asn	Pro	Thr	Ile	Gln	Lys	Glu	Val	Ile	Glu	Cys	Asp	Val
		2030					2035					2040			
40	Lys	Thr	Thr	Glu	Val	Val	Gly	Asn	Val	Ile	Leu	Lys	Pro	Ser	Asp
		2045					2050					2055			
	Glu	Gly	Val	Lys	Val	Thr	Gln	Glu	Leu	Gly	His	Glu	Asp	Leu	Met
		2060					2065					2070			
45	Ala	Ala	Tyr	Val	Glu	Asn	Thr	Ser	Ile	Thr	Ile	Lys	Lys	Pro	Asn
		2075					2080					2085			
50	Glu	Leu	Ser	Leu	Ala	Leu	Gly	Leu	Lys	Thr	Ile	Ala	Thr	His	Gly
		2090					2095					2100			
	Ile	Ala	Ala	Ile	Asn	Ser	Val	Pro	Trp	Ser	Lys	Ile	Leu	Ala	Tyr
		2105					2110					2115			
55	Val	Lys	Pro	Phe	Leu	Gly	Gln	Ala	Ala	Ile	Thr	Thr	Ser	Asn	Cys
		2120					2125					2130			

EP 1 694 829 B1

	Ala	Lys 2135	Arg	Leu	Ala	Gln	Arg 2140	Val	Phe	Asn	Asn	Tyr 2145	Met	Pro	Tyr
5	Val	Phe 2150	Thr	Leu	Leu	Phe	Gln 2155	Leu	Cys	Thr	Phe	Thr 2160	Lys	Ser	Thr
10	Asn	Ser 2165	Arg	Ile	Arg	Ala	Ser 2170	Leu	Pro	Thr	Thr	Ile 2175	Ala	Lys	Asn
	Ser	Val 2180	Lys	Ser	Val	Ala	Lys 2185	Leu	Cys	Leu	Asp	Ala 2190	Gly	Ile	Asn
15	Tyr	Val 2195	Lys	Ser	Pro	Lys	Phe 2200	Ser	Lys	Leu	Phe	Thr 2205	Ile	Ala	Met
20	Trp	Leu 2210	Leu	Leu	Leu	Ser	Ile 2215	Cys	Leu	Gly	Ser	Leu 2220	Ile	Cys	Val
	Thr	Ala 2225	Ala	Phe	Gly	Val	Leu 2230	Leu	Ser	Asn	Phe	Gly 2235	Ala	Pro	Ser
25	Tyr	Cys 2240	Asn	Gly	Val	Arg	Glu 2245	Leu	Tyr	Leu	Asn	Ser 2250	Ser	Asn	Val
30	Thr	Thr 2255	Met	Asp	Phe	Cys	Glu 2260	Gly	Ser	Phe	Pro	Cys 2265	Ser	Ile	Cys
	Leu	Ser 2270	Gly	Leu	Asp	Ser	Leu 2275	Asp	Ser	Tyr	Pro	Ala 2280	Leu	Glu	Thr
35	Ile	Gln 2285	Val	Thr	Ile	Ser	Ser 2290	Tyr	Lys	Leu	Asp	Leu 2295	Thr	Ile	Leu
40	Gly	Leu 2300	Ala	Ala	Glu	Trp	Val 2305	Leu	Ala	Tyr	Met	Leu 2310	Phe	Thr	Lys
	Phe	Phe 2315	Tyr	Leu	Leu	Gly	Leu 2320	Ser	Ala	Ile	Met	Gln 2325	Val	Phe	Phe
45	Gly	Tyr 2330	Phe	Ala	Ser	His	Phe 2335	Ile	Ser	Asn	Ser	Trp 2340	Leu	Met	Trp
50	Phe	Ile 2345	Ile	Ser	Ile	Val	Gln 2350	Met	Ala	Pro	Val	Ser 2355	Ala	Met	Val
	Arg	Met 2360	Tyr	Ile	Phe	Phe	Ala 2365	Ser	Phe	Tyr	Tyr	Ile 2370	Trp	Lys	Ser
55	Tyr	Val 2375	His	Ile	Met	Asp	Gly 2380	Cys	Thr	Ser	Ser	Thr 2385	Cys	Met	Met

EP 1 694 829 B1

	Cys	Tyr	Lys	Arg	Asn	Arg	Ala	Thr	Arg	Val	Glu	Cys	Thr	Thr	Ile
		2390					2395					2400			
5	Val	Asn	Gly	Met	Lys	Arg	Ser	Phe	Tyr	Val	Tyr	Ala	Asn	Gly	Gly
		2405					2410					2415			
10	Arg	Gly	Phe	Cys	Lys	Thr	His	Asn	Trp	Asn	Cys	Leu	Asn	Cys	Asp
		2420					2425					2430			
	Thr	Phe	Cys	Thr	Gly	Ser	Thr	Phe	Ile	Ser	Asp	Glu	Val	Ala	Arg
		2435					2440					2445			
15	Asp	Leu	Ser	Leu	Gln	Phe	Lys	Arg	Pro	Ile	Asn	Pro	Thr	Asp	Gln
		2450					2455					2460			
20	Ser	Ser	Tyr	Ile	Val	Asp	Ser	Val	Ala	Val	Lys	Asn	Gly	Ala	Leu
		2465					2470					2475			
	His	Leu	Tyr	Phe	Asp	Lys	Ala	Gly	Gln	Lys	Thr	Tyr	Glu	Arg	His
		2480					2485					2490			
25	Pro	Leu	Ser	His	Phe	Val	Asn	Leu	Asp	Asn	Leu	Arg	Ala	Asn	Asn
		2495					2500					2505			
30	Thr	Lys	Gly	Ser	Leu	Pro	Ile	Asn	Val	Ile	Val	Phe	Asp	Gly	Lys
		2510					2515					2520			
	Ser	Lys	Cys	Asp	Glu	Ser	Ala	Ser	Lys	Ser	Ala	Ser	Val	Tyr	Tyr
		2525					2530					2535			
35	Ser	Gln	Leu	Met	Cys	Gln	Pro	Ile	Leu	Leu	Leu	Asp	Gln	Ala	Leu
		2540					2545					2550			
40	Val	Ser	Asp	Val	Gly	Asp	Ser	Thr	Glu	Val	Ser	Val	Lys	Met	Phe
		2555					2560					2565			
	Asp	Ala	Tyr	Val	Asp	Thr	Phe	Ser	Ala	Thr	Phe	Ser	Val	Pro	Met
		2570					2575					2580			
45	Glu	Lys	Leu	Lys	Ala	Leu	Val	Ala	Thr	Ala	His	Ser	Glu	Leu	Ala
		2585					2590					2595			
50	Lys	Gly	Val	Ala	Leu	Asp	Gly	Val	Leu	Ser	Thr	Phe	Val	Ser	Ala
		2600					2605					2610			
	Ala	Arg	Gln	Gly	Val	Val	Asp	Thr	Asp	Val	Asp	Thr	Lys	Asp	Val
		2615					2620					2625			
55	Ile	Glu	Cys	Leu	Lys	Leu	Ser	His	His	Ser	Asp	Leu	Glu	Val	Thr
		2630					2635					2640			

EP 1 694 829 B1

	Gly	Asp 2645	Ser	Cys	Asn	Asn	Phe 2650	Met	Leu	Thr	Tyr	Asn 2655	Lys	Val	Glu
5	Asn	Met 2660	Thr	Pro	Arg	Asp	Leu 2665	Gly	Ala	Cys	Ile	Asp 2670	Cys	Asn	Ala
10	Arg	His 2675	Ile	Asn	Ala	Gln	Val 2680	Ala	Lys	Ser	His	Asn 2685	Val	Ser	Leu
	Ile	Trp 2690	Asn	Val	Lys	Asp	Tyr 2695	Met	Ser	Leu	Ser	Glu 2700	Gln	Leu	Arg
15	Lys	Gln 2705	Ile	Arg	Ser	Ala	Ala 2710	Lys	Lys	Asn	Asn	Ile 2715	Pro	Phe	Arg
20	Leu	Thr 2720	Cys	Ala	Thr	Thr	Arg 2725	Gln	Val	Val	Asn	Val 2730	Ile	Thr	Thr
	Lys	Ile 2735	Ser	Leu	Lys	Gly	Gly 2740	Lys	Ile	Val	Ser	Thr 2745	Cys	Phe	Lys
25	Leu	Met 2750	Leu	Lys	Ala	Thr	Leu 2755	Leu	Cys	Val	Leu	Ala 2760	Ala	Leu	Val
30	Cys	Tyr 2765	Ile	Val	Met	Pro	Val 2770	His	Thr	Leu	Ser	Ile 2775	His	Asp	Gly
	Tyr	Thr 2780	Asn	Glu	Ile	Ile	Gly 2785	Tyr	Lys	Ala	Ile	Gln 2790	Asp	Gly	Val
35	Thr	Arg 2795	Asp	Ile	Ile	Ser	Thr 2800	Asp	Asp	Cys	Phe	Ala 2805	Asn	Lys	His
40	Ala	Gly 2810	Phe	Asp	Ala	Trp	Phe 2815	Ser	Gln	Arg	Gly	Gly 2820	Ser	Tyr	Lys
	Asn	Asp 2825	Lys	Ser	Cys	Pro	Val 2830	Val	Ala	Ala	Ile	Ile 2835	Thr	Arg	Glu
45	Ile	Gly 2840	Phe	Ile	Val	Pro	Gly 2845	Leu	Pro	Gly	Thr	Val 2850	Leu	Arg	Ala
50	Ile	Asn 2855	Gly	Asp	Phe	Leu	His 2860	Phe	Leu	Pro	Arg	Val 2865	Phe	Ser	Ala
	Val	Gly 2870	Asn	Ile	Cys	Tyr	Thr 2875	Pro	Ser	Lys	Leu	Ile 2880	Glu	Tyr	Ser
55	Asp	Phe 2885	Ala	Thr	Ser	Ala	Cys 2890	Val	Leu	Ala	Ala	Glu 2895	Cys	Thr	Ile



EP 1 694 829 B1

	Phe	Lys 2900	Asp	Ala	Met	Gly	Lys 2905	Pro	Val	Pro	Tyr	Cys 2910	Tyr	Asp	Thr
5	Asn	Leu 2915	Leu	Glu	Gly	Ser	Ile 2920	Ser	Tyr	Ser	Glu	Leu 2925	Arg	Pro	Asp
10	Thr	Arg 2930	Tyr	Val	Leu	Met	Asp 2935	Gly	Ser	Ile	Ile	Gln 2940	Phe	Pro	Asn
	Thr	Tyr 2945	Leu	Glu	Gly	Ser	Val 2950	Arg	Val	Val	Thr	Thr 2955	Phe	Asp	Ala
15	Glu	Tyr 2960	Cys	Arg	His	Gly	Thr 2965	Cys	Glu	Arg	Ser	Glu 2970	Val	Gly	Ile
20	Cys	Leu 2975	Ser	Thr	Ser	Gly	Arg 2980	Trp	Val	Leu	Asn	Asn 2985	Glu	His	Tyr
	Arg	Ala 2990	Leu	Ser	Gly	Val	Phe 2995	Cys	Gly	Val	Asp	Ala 3000	Met	Asn	Leu
25	Ile	Ala 3005	Asn	Ile	Phe	Thr	Pro 3010	Leu	Val	Gln	Pro	Val 3015	Gly	Ala	Leu
30	Asp	Val 3020	Ser	Ala	Ser	Val	Val 3025	Ala	Gly	Gly	Ile	Ile 3030	Ala	Ile	Leu
	Val	Thr 3035	Cys	Ala	Ala	Tyr	Tyr 3040	Phe	Met	Lys	Phe	Arg 3045	Arg	Val	Phe
35	Gly	Glu 3050	Tyr	Asn	His	Val	Val 3055	Ala	Ala	Asn	Ala	Leu 3060	Leu	Phe	Leu
40	Met	Ser 3065	Phe	Thr	Ile	Leu	Cys 3070	Leu	Val	Pro	Ala	Tyr 3075	Ser	Phe	Leu
	Pro	Gly 3080	Val	Tyr	Ser	Val	Phe 3085	Tyr	Leu	Tyr	Leu	Thr 3090	Phe	Tyr	Phe
45	Thr	Asn 3095	Asp	Val	Ser	Phe	Leu 3100	Ala	His	Leu	Gln	Trp 3105	Phe	Ala	Met
50	Phe	Ser 3110	Pro	Ile	Val	Pro	Phe 3115	Trp	Ile	Thr	Ala	Ile 3120	Tyr	Val	Phe
	Cys	Ile 3125	Ser	Leu	Lys	His	Cys 3130	His	Trp	Phe	Phe	Asn 3135	Asn	Tyr	Leu
55	Arg	Lys 3140	Arg	Val	Met	Phe	Asn 3145	Gly	Val	Thr	Phe	Ser 3150	Thr	Phe	Glu

EP 1 694 829 B1

	Glu	Ala	Ala	Leu	Cys	Thr	Phe	Leu	Leu	Asn	Lys	Glu	Met	Tyr	Leu
		3155					3160					3165			
5	Lys	Leu	Arg	Ser	Glu	Thr	Leu	Leu	Pro	Leu	Thr	Gln	Tyr	Asn	Arg
		3170					3175					3180			
	Tyr	Leu	Ala	Leu	Tyr	Asn	Lys	Tyr	Lys	Tyr	Phe	Ser	Gly	Ala	Leu
10		3185					3190					3195			
	Asp	Thr	Thr	Ser	Tyr	Arg	Glu	Ala	Ala	Cys	Cys	His	Leu	Ala	Lys
		3200					3205					3210			
15	Ala	Leu	Asn	Asp	Phe	Ser	Asn	Ser	Gly	Ala	Asp	Val	Leu	Tyr	Gln
		3215					3220					3225			
	Pro	Pro	Gln	Thr	Ser	Ile	Thr	Ser	Ala	Val	Leu	Gln	Ser	Gly	Phe
20		3230					3235					3240			
	Arg	Lys	Met	Ala	Phe	Pro	Ser	Gly	Lys	Val	Glu	Gly	Cys	Met	Val
		3245					3250					3255			
25	Gln	Val	Thr	Cys	Gly	Thr	Thr	Thr	Leu	Asn	Gly	Leu	Trp	Leu	Asp
		3260					3265					3270			
	Asp	Thr	Val	Tyr	Cys	Pro	Arg	His	Val	Ile	Cys	Thr	Ala	Glu	Asp
30		3275					3280					3285			
	Met	Leu	Asn	Pro	Asn	Tyr	Glu	Asp	Leu	Leu	Ile	Arg	Lys	Ser	Asn
		3290					3295					3300			
35	His	Ser	Phe	Leu	Val	Gln	Ala	Gly	Asn	Val	Gln	Leu	Arg	Val	Ile
		3305					3310					3315			
	Gly	His	Ser	Met	Gln	Asn	Cys	Leu	Leu	Arg	Leu	Lys	Val	Asp	Thr
40		3320					3325					3330			
	Ser	Asn	Pro	Lys	Thr	Pro	Lys	Tyr	Lys	Phe	Val	Arg	Ile	Gln	Pro
		3335					3340					3345			
45	Gly	Gln	Thr	Phe	Ser	Val	Leu	Ala	Cys	Tyr	Asn	Gly	Ser	Pro	Ser
		3350					3355					3360			
	Gly	Val	Tyr	Gln	Cys	Ala	Met	Arg	Pro	Asn	His	Thr	Ile	Lys	Gly
50		3365					3370					3375			
	Ser	Phe	Leu	Asn	Gly	Ser	Cys	Gly	Ser	Val	Gly	Phe	Asn	Ile	Asp
		3380					3385					3390			
55	Tyr	Asp	Cys	Val	Ser	Phe	Cys	Tyr	Met	His	His	Met	Glu	Leu	Pro
		3395					3400					3405			

EP 1 694 829 B1

	Thr	Gly	Val	His	Ala	Gly	Thr	Asp	Leu	Glu	Gly	Lys	Phe	Tyr	Gly
		3410					3415					3420			
5	Pro	Phe	Val	Asp	Arg	Gln	Thr	Ala	Gln	Ala	Ala	Gly	Thr	Asp	Thr
		3425					3430					3435			
10	Thr	Ile	Thr	Leu	Asn	Val	Leu	Ala	Trp	Leu	Tyr	Ala	Ala	Val	Ile
		3440					3445					3450			
	Asn	Gly	Asp	Arg	Trp	Phe	Leu	Asn	Arg	Phe	Thr	Thr	Thr	Leu	Asn
		3455					3460					3465			
15	Asp	Phe	Asn	Leu	Val	Ala	Met	Lys	Tyr	Asn	Tyr	Glu	Pro	Leu	Thr
		3470					3475					3480			
20	Gln	Asp	His	Val	Asp	Ile	Leu	Gly	Pro	Leu	Ser	Ala	Gln	Thr	Gly
		3485					3490					3495			
	Ile	Ala	Val	Leu	Asp	Met	Cys	Ala	Ala	Leu	Lys	Glu	Leu	Leu	Gln
		3500					3505					3510			
25	Asn	Gly	Met	Asn	Gly	Arg	Thr	Ile	Leu	Gly	Ser	Thr	Ile	Leu	Glu
		3515					3520					3525			
30	Asp	Glu	Phe	Thr	Pro	Phe	Asp	Val	Val	Arg	Gln	Cys	Ser	Gly	Val
		3530					3535					3540			
	Thr	Phe	Gln	Gly	Lys	Phe	Lys	Lys	Ile	Val	Lys	Gly	Thr	His	His
		3545					3550					3555			
35	Trp	Met	Leu	Leu	Thr	Phe	Leu	Thr	Ser	Leu	Leu	Ile	Leu	Val	Gln
		3560					3565					3570			
40	Ser	Thr	Gln	Trp	Ser	Leu	Phe	Phe	Phe	Val	Tyr	Glu	Asn	Ala	Phe
		3575					3580					3585			
	Leu	Pro	Phe	Thr	Leu	Gly	Ile	Met	Ala	Ile	Ala	Ala	Cys	Ala	Met
		3590					3595					3600			
45	Leu	Leu	Val	Lys	His	Lys	His	Ala	Phe	Leu	Cys	Leu	Phe	Leu	Leu
		3605					3610					3615			
50	Pro	Ser	Leu	Ala	Thr	Val	Ala	Tyr	Phe	Asn	Met	Val	Tyr	Met	Pro
		3620					3625					3630			
	Ala	Ser	Trp	Val	Met	Arg	Ile	Met	Thr	Trp	Leu	Glu	Leu	Ala	Asp
		3635					3640					3645			
55	Thr	Ser	Leu	Ser	Gly	Tyr	Arg	Leu	Lys	Asp	Cys	Val	Met	Tyr	Ala
		3650					3655					3660			

EP 1 694 829 B1

	Ser	Ala	Leu	Val	Leu	Leu	Ile	Leu	Met	Thr	Ala	Arg	Thr	Val	Tyr
		3665					3670					3675			
5	Asp	Asp	Ala	Ala	Arg	Arg	Val	Trp	Thr	Leu	Met	Asn	Val	Ile	Thr
		3680					3685					3690			
10	Leu	Val	Tyr	Lys	Val	Tyr	Tyr	Gly	Asn	Ala	Leu	Asp	Gln	Ala	Ile
		3695					3700					3705			
	Ser	Met	Trp	Ala	Leu	Val	Ile	Ser	Val	Thr	Ser	Asn	Tyr	Ser	Gly
		3710					3715					3720			
15	Val	Val	Thr	Thr	Ile	Met	Phe	Leu	Ala	Arg	Ala	Ile	Val	Phe	Val
		3725					3730					3735			
20	Cys	Val	Glu	Tyr	Tyr	Pro	Leu	Leu	Phe	Ile	Thr	Gly	Asn	Thr	Leu
		3740					3745					3750			
	Gln	Cys	Ile	Met	Leu	Val	Tyr	Cys	Phe	Leu	Gly	Tyr	Cys	Cys	Cys
		3755					3760					3765			
25	Cys	Tyr	Phe	Gly	Leu	Phe	Cys	Leu	Leu	Asn	Arg	Tyr	Phe	Arg	Leu
		3770					3775					3780			
30	Thr	Leu	Gly	Val	Tyr	Asp	Tyr	Leu	Val	Ser	Thr	Gln	Glu	Phe	Arg
		3785					3790					3795			
	Tyr	Met	Asn	Ser	Gln	Gly	Leu	Leu	Pro	Pro	Lys	Ser	Ser	Ile	Asp
		3800					3805					3810			
35	Ala	Phe	Lys	Leu	Asn	Ile	Lys	Leu	Leu	Gly	Ile	Gly	Gly	Lys	Pro
		3815					3820					3825			
40	Cys	Ile	Lys	Val	Ala	Thr	Val	Gln	Ser	Lys	Met	Ser	Asp	Val	Lys
		3830					3835					3840			
	Cys	Thr	Ser	Val	Val	Leu	Leu	Ser	Val	Leu	Gln	Gln	Leu	Arg	Val
		3845					3850					3855			
45	Glu	Ser	Ser	Ser	Lys	Leu	Trp	Ala	Gln	Cys	Val	Gln	Leu	His	Asn
		3860					3865					3870			
50	Asp	Ile	Leu	Leu	Ala	Lys	Asp	Thr	Thr	Glu	Ala	Phe	Glu	Lys	Met
		3875					3880					3885			
	Val	Ser	Leu	Leu	Ser	Val	Leu	Leu	Ser	Met	Gln	Gly	Ala	Val	Asp
		3890					3895					3900			
55	Ile	Asn	Arg	Leu	Cys	Glu	Glu	Met	Leu	Asp	Asn	Arg	Ala	Thr	Leu
		3905					3910					3915			

EP 1 694 829 B1

	Gln	Ala	Ile	Ala	Ser	Glu	Phe	Ser	Ser	Leu	Pro	Ser	Tyr	Ala	Ala
		3920					3925					3930			
5	Tyr	Ala	Thr	Ala	Gln	Glu	Ala	Tyr	Glu	Gln	Ala	Val	Ala	Asn	Gly
		3935					3940					3945			
10	Asp	Ser	Glu	Val	Val	Leu	Lys	Lys	Leu	Lys	Lys	Ser	Leu	Asn	Val
		3950					3955					3960			
	Ala	Lys	Ser	Glu	Phe	Asp	Arg	Asp	Ala	Ala	Met	Gln	Arg	Lys	Leu
		3965					3970					3975			
15	Glu	Lys	Met	Ala	Asp	Gln	Ala	Met	Thr	Gln	Met	Tyr	Lys	Gln	Ala
		3980					3985					3990			
20	Arg	Ser	Glu	Asp	Lys	Arg	Ala	Lys	Val	Thr	Ser	Ala	Met	Gln	Thr
		3995					4000					4005			
	Met	Leu	Phe	Thr	Met	Leu	Arg	Lys	Leu	Asp	Asn	Asp	Ala	Leu	Asn
		4010					4015					4020			
25	Asn	Ile	Ile	Asn	Asn	Ala	Arg	Asp	Gly	Cys	Val	Pro	Leu	Asn	Ile
		4025					4030					4035			
30	Ile	Pro	Leu	Thr	Thr	Ala	Ala	Lys	Leu	Met	Val	Val	Val	Pro	Asp
		4040					4045					4050			
	Tyr	Gly	Thr	Tyr	Lys	Asn	Thr	Cys	Asp	Gly	Asn	Thr	Phe	Thr	Tyr
		4055					4060					4065			
35	Ala	Ser	Ala	Leu	Trp	Glu	Ile	Gln	Gln	Val	Val	Asp	Ala	Asp	Ser
		4070					4075					4080			
40	Lys	Ile	Val	Gln	Leu	Ser	Glu	Ile	Asn	Met	Asp	Asn	Ser	Pro	Asn
		4085					4090					4095			
	Leu	Ala	Trp	Pro	Leu	Ile	Val	Thr	Ala	Leu	Arg	Ala	Asn	Ser	Ala
		4100					4105					4110			
45	Val	Lys	Leu	Gln	Asn	Asn	Glu	Leu	Ser	Pro	Val	Ala	Leu	Arg	Gln
		4115					4120					4125			
50	Met	Ser	Cys	Ala	Ala	Gly	Thr	Thr	Gln	Thr	Ala	Cys	Thr	Asp	Asp
		4130					4135					4140			
	Asn	Ala	Leu	Ala	Tyr	Tyr	Asn	Asn	Ser	Lys	Gly	Gly	Arg	Phe	Val
		4145					4150					4155			
55	Leu	Ala	Leu	Leu	Ser	Asp	His	Gln	Asp	Leu	Lys	Trp	Ala	Arg	Phe
		4160					4165					4170			

EP 1 694 829 B1

5 Pro Lys Ser Asp Gly Thr Gly Thr Ile Tyr Thr Glu Leu Glu Pro  
 4175 4180 4185  
 10 Pro Cys Arg Phe Val Thr Asp Thr Pro Lys Gly Pro Lys Val Lys  
 4190 4195 4200  
 15 Tyr Leu Tyr Phe Ile Lys Gly Leu Asn Asn Leu Asn Arg Gly Met  
 4205 4210 4215  
 20 Val Leu Gly Ser Leu Ala Ala Thr Val Arg Leu Gln Ala Gly Asn  
 4220 4225 4230  
 25 Ala Thr Glu Val Pro Ala Asn Ser Thr Val Leu Ser Phe Cys Ala  
 4235 4240 4245  
 30 Phe Ala Val Asp Pro Ala Lys Ala Tyr Lys Asp Tyr Leu Ala Ser  
 4250 4255 4260  
 35 Gly Gly Gln Pro Ile Thr Asn Cys Val Lys Met Leu Cys Thr His  
 4265 4270 4275  
 40 Thr Gly Thr Gly Gln Ala Ile Thr Val Thr Pro Glu Ala Asn Met  
 4280 4285 4290  
 45 Asp Gln Glu Ser Phe Gly Gly Ala Ser Cys Cys Leu Tyr Cys Arg  
 4295 4300 4305  
 50 Cys His Ile Asp His Pro Asn Pro Lys Gly Phe Cys Asp Leu Lys  
 4310 4315 4320  
 55 Gly Lys Tyr Val Gln Ile Pro Thr Thr Cys Ala Asn Asp Pro Val  
 4325 4330 4335  
 60 Gly Phe Thr Leu Arg Asn Thr Val Cys Thr Val Cys Gly Met Trp  
 4340 4345 4350  
 65 Lys Gly Tyr Gly Cys Ser Cys Asp Gln Leu Arg Glu Pro Leu Met  
 4355 4360 4365  
 70 Gln Ser Ala Asp Ala Ser Thr Phe Leu Asn Gly Phe Ala Val  
 4370 4375 4380  
 75 <210> 75  
 80 <211> 2695  
 85 <212> PRT  
 90 <213> CORONAVIRUS  
 95 <400> 75

EP 1 694 829 B1

Arg Val Cys Gly Val Ser Ala Ala Arg Leu Thr Pro Cys Gly Thr Gly  
 1 5 10 15  
 5 Thr Ser Thr Asp Val Val Tyr Arg Ala Phe Asp Ile Tyr Asn Glu Lys  
 20 25 30  
 Val Ala Gly Phe Ala Lys Phe Leu Lys Thr Asn Cys Cys Arg Phe Gln  
 35 40 45  
 10 Glu Lys Asp Glu Glu Gly Asn Leu Leu Asp Ser Tyr Phe Val Val Lys  
 50 55 60  
 15 Arg His Thr Met Ser Asn Tyr Gln His Glu Glu Thr Ile Tyr Asn Leu  
 65 70 75 80  
 Val Lys Asp Cys Pro Ala Val Ala Val His Asp Phe Phe Lys Phe Arg  
 85 90 95  
 20 Val Asp Gly Asp Met Val Pro His Ile Ser Arg Gln Arg Leu Thr Lys  
 100 105 110  
 25 Tyr Thr Met Ala Asp Leu Val Tyr Ala Leu Arg His Phe Asp Glu Gly  
 115 120 125  
 Asn Cys Asp Thr Leu Lys Glu Ile Leu Val Thr Tyr Asn Cys Cys Asp  
 130 135 140  
 30 Asp Asp Tyr Phe Asn Lys Lys Asp Trp Tyr Asp Phe Val Glu Asn Pro  
 145 150 155 160  
 35 Asp Ile Leu Arg Val Tyr Ala Asn Leu Gly Glu Arg Val Arg Gln Ser  
 165 170 175  
 Leu Leu Lys Thr Val Gln Phe Cys Asp Ala Met Arg Asp Ala Gly Ile  
 180 185 190  
 40 Val Gly Val Leu Thr Leu Asp Asn Gln Asp Leu Asn Gly Asn Trp Tyr  
 195 200 205  
 45 Asp Phe Gly Asp Phe Val Gln Val Ala Pro Gly Cys Gly Val Pro Ile  
 210 215 220  
 Val Asp Ser Tyr Tyr Ser Leu Leu Met Pro Ile Leu Thr Leu Thr Arg  
 225 230 235 240  
 50 Ala Leu Ala Ala Glu Ser His Met Asp Ala Asp Leu Ala Lys Pro Leu  
 245 250 255  
 55 Ile Lys Trp Asp Leu Leu Lys Tyr Asp Phe Thr Glu Glu Arg Leu Cys  
 260 265 270

EP 1 694 829 B1

Leu Phe Asp Arg Tyr Phe Lys Tyr Trp Asp Gln Thr Tyr His Pro Asn  
 275 280 285  
 5  
 Cys Ile Asn Cys Leu Asp Asp Arg Cys Ile Leu His Cys Ala Asn Phe  
 290 295 300  
 Asn Val Leu Phe Ser Thr Val Phe Pro Pro Thr Ser Phe Gly Pro Leu  
 305 310 315  
 10  
 Val Arg Lys Ile Phe Val Asp Gly Val Pro Phe Val Val Ser Thr Gly  
 325 330 335  
 Tyr His Phe Arg Glu Leu Gly Val Val His Asn Gln Asp Val Asn Leu  
 340 345 350  
 15  
 His Ser Ser Arg Leu Ser Phe Lys Glu Leu Leu Val Tyr Ala Ala Asp  
 355 360 365  
 20  
 Pro Ala Met His Ala Ala Ser Gly Asn Leu Leu Leu Asp Lys Arg Thr  
 370 375 380  
 Thr Cys Phe Ser Val Ala Ala Leu Thr Asn Asn Val Ala Phe Gln Thr  
 385 390 395  
 25  
 Val Lys Pro Gly Asn Phe Asn Lys Asp Phe Tyr Asp Phe Ala Val Ser  
 405 410 415  
 30  
 Lys Gly Phe Phe Lys Glu Gly Ser Ser Val Glu Leu Lys His Phe Phe  
 420 425 430  
 Phe Ala Gln Asp Gly Asn Ala Ala Ile Ser Asp Tyr Asp Tyr Tyr Arg  
 435 440 445  
 35  
 Tyr Asn Leu Pro Thr Met Cys Asp Ile Arg Gln Leu Leu Phe Val Val  
 450 455 460  
 40  
 Glu Val Val Asp Lys Tyr Phe Asp Cys Tyr Asp Gly Gly Cys Ile Asn  
 465 470 475  
 Ala Asn Gln Val Ile Val Asn Asn Leu Asp Lys Ser Ala Gly Phe Pro  
 485 490 495  
 45  
 Phe Asn Lys Trp Gly Lys Ala Arg Leu Tyr Tyr Asp Ser Met Ser Tyr  
 500 505 510  
 50  
 Glu Asp Gln Asp Ala Leu Phe Ala Tyr Thr Lys Arg Asn Val Ile Pro  
 515 520 525  
 Thr Ile Thr Gln Met Asn Leu Lys Tyr Ala Ile Ser Ala Lys Asn Arg  
 530 535 540  
 55



EP 1 694 829 B1

Ala Arg Thr Val Ala Gly Val Ser Ile Cys Ser Thr Met Thr Asn Arg  
545 550 555 560

5 Gln Phe His Gln Lys Leu Leu Lys Ser Ile Ala Ala Thr Arg Gly Ala  
565 570 575

Thr Val Val Ile Gly Thr Ser Lys Phe Tyr Gly Gly Trp His Asn Met  
10 580 585 590

Leu Lys Thr Val Tyr Ser Asp Val Glu Thr Pro His Leu Met Gly Trp  
595 600 605

15 Asp Tyr Pro Lys Cys Asp Arg Ala Met Pro Asn Met Leu Arg Ile Met  
610 615 620

Ala Ser Leu Val Leu Ala Arg Lys His Asn Thr Cys Cys Asn Leu Ser  
20 625 630 635 640

His Arg Phe Tyr Arg Leu Ala Asn Glu Cys Ala Gln Val Leu Ser Glu  
645 650 655

25 Met Val Met Cys Gly Gly Ser Leu Tyr Val Lys Pro Gly Gly Thr Ser  
660 665 670

Ser Gly Asp Ala Thr Thr Ala Tyr Ala Asn Ser Val Phe Asn Ile Cys  
30 675 680 685

Gln Ala Val Thr Ala Asn Val Asn Ala Leu Leu Ser Thr Asp Gly Asn  
690 695 700

35 Lys Ile Ala Asp Lys Tyr Val Arg Asn Leu Gln His Arg Leu Tyr Glu  
705 710 715 720

Cys Leu Tyr Arg Asn Arg Asp Val Asp His Glu Phe Val Asp Glu Phe  
40 725 730 735

Tyr Ala Tyr Leu Arg Lys His Phe Ser Met Met Ile Leu Ser Asp Asp  
740 745 750

45 Ala Val Val Cys Tyr Asn Ser Asn Tyr Ala Ala Gln Gly Leu Val Ala  
755 760 765

Ser Ile Lys Asn Phe Lys Ala Val Leu Tyr Tyr Gln Asn Asn Val Phe  
50 770 775 780

Met Ser Glu Ala Lys Cys Trp Thr Glu Thr Asp Leu Thr Lys Gly Pro  
785 790 795 800

55 His Glu Phe Cys Ser Gln His Thr Met Leu Val Lys Gln Gly Asp Asp  
805 810 815

EP 1 694 829 B1

Tyr Val Tyr Leu Pro Tyr Pro Asp Pro Ser Arg Ile Leu Gly Ala Gly  
 820 825 830  
 5  
 Cys Phe Val Asp Asp Ile Val Lys Thr Asp Gly Thr Leu Met Ile Glu  
 835 840 845  
 Arg Phe Val Ser Leu Ala Ile Asp Ala Tyr Pro Leu Thr Lys His Pro  
 850 855 860  
 10  
 Asn Gln Glu Tyr Ala Asp Val Phe His Leu Tyr Leu Gln Tyr Ile Arg  
 865 870 875 880  
 15  
 Lys Leu His Asp Glu Leu Thr Gly His Met Leu Asp Met Tyr Ser Val  
 885 890 895  
 Met Leu Thr Asn Asp Asn Thr Ser Arg Tyr Trp Glu Pro Glu Phe Tyr  
 900 905 910  
 20  
 Glu Ala Met Tyr Thr Pro His Thr Val Leu Gln Ala Val Gly Ala Cys  
 915 920 925  
 25  
 Val Leu Cys Asn Ser Gln Thr Ser Leu Arg Cys Gly Ala Cys Ile Arg  
 930 935 940  
 Arg Pro Phe Leu Cys Cys Lys Cys Cys Tyr Asp His Val Ile Ser Thr  
 945 950 955 960  
 30  
 Ser His Lys Leu Val Leu Ser Val Asn Pro Tyr Val Cys Asn Ala Pro  
 965 970 975  
 35  
 Gly Cys Asp Val Thr Asp Val Thr Gln Leu Tyr Leu Gly Gly Met Ser  
 980 985 990  
 Tyr Tyr Cys Lys Ser His Lys Pro Pro Ile Ser Phe Pro Leu Cys Ala  
 995 1000 1005  
 40  
 Asn Gly Gln Val Phe Gly Leu Tyr Lys Asn Thr Cys Val Gly Ser  
 1010 1015 1020  
 45  
 Asp Asn Val Thr Asp Phe Asn Ala Ile Ala Thr Cys Asp Trp Thr  
 1025 1030 1035  
 Asn Ala Gly Asp Tyr Ile Leu Ala Asn Thr Cys Thr Glu Arg Leu  
 1040 1045 1050  
 50  
 Lys Leu Phe Ala Ala Glu Thr Leu Lys Ala Thr Glu Glu Thr Phe  
 1055 1060 1065  
 55  
 Lys Leu Ser Tyr Gly Ile Ala Thr Val Arg Glu Val Leu Ser Asp  
 1070 1075 1080

EP 1 694 829 B1

	Arg	Glu	Leu	His	Leu	Ser	Trp	Glu	Val	Gly	Lys	Pro	Arg	Pro	Pro
		1085					1090					1095			
5	Leu	Asn	Arg	Asn	Tyr	Val	Phe	Thr	Gly	Tyr	Arg	Val	Thr	Lys	Asn
		1100					1105					1110			
	Ser	Lys	Val	Gln	Ile	Gly	Glu	Tyr	Thr	Phe	Glu	Lys	Gly	Asp	Tyr
10		1115					1120					1125			
	Gly	Asp	Ala	Val	Val	Tyr	Arg	Gly	Thr	Thr	Thr	Tyr	Lys	Leu	Asn
		1130					1135					1140			
15	Val	Gly	Asp	Tyr	Phe	Val	Leu	Thr	Ser	His	Thr	Val	Met	Pro	Leu
		1145					1150					1155			
	Ser	Ala	Pro	Thr	Leu	Val	Pro	Gln	Glu	His	Tyr	Val	Arg	Ile	Thr
20		1160					1165					1170			
	Gly	Leu	Tyr	Pro	Thr	Leu	Asn	Ile	Ser	Asp	Glu	Phe	Ser	Ser	Asn
		1175					1180					1185			
25	Val	Ala	Asn	Tyr	Gln	Lys	Val	Gly	Met	Gln	Lys	Tyr	Ser	Thr	Leu
		1190					1195					1200			
	Gln	Gly	Pro	Pro	Gly	Thr	Gly	Lys	Ser	His	Phe	Ala	Ile	Gly	Leu
30		1205					1210					1215			
	Ala	Leu	Tyr	Tyr	Pro	Ser	Ala	Arg	Ile	Val	Tyr	Thr	Ala	Cys	Ser
		1220					1225					1230			
35	His	Ala	Ala	Val	Asp	Ala	Leu	Cys	Glu	Lys	Ala	Leu	Lys	Tyr	Leu
		1235					1240					1245			
	Pro	Ile	Asp	Lys	Cys	Ser	Arg	Ile	Ile	Pro	Ala	Arg	Ala	Arg	Val
40		1250					1255					1260			
	Glu	Cys	Phe	Asp	Lys	Phe	Lys	Val	Asn	Ser	Thr	Leu	Glu	Gln	Tyr
		1265					1270					1275			
45	Val	Phe	Cys	Thr	Val	Asn	Ala	Leu	Pro	Glu	Thr	Thr	Ala	Asp	Ile
		1280					1285					1290			
	Val	Val	Phe	Asp	Glu	Ile	Ser	Met	Ala	Thr	Asn	Tyr	Asp	Leu	Ser
50		1295					1300					1305			
	Val	Val	Asn	Ala	Arg	Leu	Arg	Ala	Lys	His	Tyr	Val	Tyr	Ile	Gly
		1310					1315					1320			
55	Asp	Pro	Ala	Gln	Leu	Pro	Ala	Pro	Arg	Thr	Leu	Leu	Thr	Lys	Gly
		1325					1330					1335			

EP 1 694 829 B1

	Thr	Leu 1340	Glu	Pro	Glu	Tyr	Phe 1345	Asn	Ser	Val	Cys	Arg 1350	Leu	Met	Lys
5	Thr	Ile 1355	Gly	Pro	Asp	Met	Phe 1360	Leu	Gly	Thr	Cys	Arg 1365	Arg	Cys	Pro
10	Ala	Glu 1370	Ile	Val	Asp	Thr	Val 1375	Ser	Ala	Leu	Val	Tyr 1380	Asp	Asn	Lys
	Leu	Lys 1385	Ala	His	Lys	Asp	Lys 1390	Ser	Ala	Gln	Cys	Phe 1395	Lys	Met	Phe
15	Tyr	Lys 1400	Gly	Val	Ile	Thr	His 1405	Asp	Val	Ser	Ser	Ala 1410	Ile	Asn	Arg
20	Pro	Gln 1415	Ile	Gly	Val	Val	Arg 1420	Glu	Phe	Leu	Thr	Arg 1425	Asn	Pro	Ala
	Trp	Arg 1430	Lys	Ala	Val	Phe	Ile 1435	Ser	Pro	Tyr	Asn	Ser 1440	Gln	Asn	Ala
25	Val	Ala 1445	Ser	Lys	Ile	Leu	Gly 1450	Leu	Pro	Thr	Gln	Thr 1455	Val	Asp	Ser
30	Ser	Gln 1460	Gly	Ser	Glu	Tyr	Asp 1465	Tyr	Val	Ile	Phe	Thr 1470	Gln	Thr	Thr
	Glu	Thr 1475	Ala	His	Ser	Cys	Asn 1480	Val	Asn	Arg	Phe	Asn 1485	Val	Ala	Ile
35	Thr	Arg 1490	Ala	Lys	Ile	Gly	Ile 1495	Leu	Cys	Ile	Met	Ser 1500	Asp	Arg	Asp
40	Leu	Tyr 1505	Asp	Lys	Leu	Gln	Phe 1510	Thr	Ser	Leu	Glu	Ile 1515	Pro	Arg	Arg
	Asn	Val 1520	Ala	Thr	Leu	Gln	Ala 1525	Glu	Asn	Val	Thr	Gly 1530	Leu	Phe	Lys
45	Asp	Cys 1535	Ser	Lys	Ile	Ile	Thr 1540	Gly	Leu	His	Pro	Thr 1545	Gln	Ala	Pro
50	Thr	His 1550	Leu	Ser	Val	Asp	Ile 1555	Lys	Phe	Lys	Thr	Glu 1560	Gly	Leu	Cys
	Val	Asp 1565	Ile	Pro	Gly	Ile	Pro 1570	Lys	Asp	Met	Thr	Tyr 1575	Arg	Arg	Leu
55	Ile	Ser 1580	Met	Met	Gly	Phe	Lys 1585	Met	Asn	Tyr	Gln	Val 1590	Asn	Gly	Tyr

EP 1 694 829 B1

	Pro	Asn	Met	Phe	Ile	Thr	Arg	Glu	Glu	Ala	Ile	Arg	His	Val	Arg
		1595					1600					1605			
5	Ala	Trp	Ile	Gly	Phe	Asp	Val	Glu	Gly	Cys	His	Ala	Thr	Arg	Asp
		1610					1615					1620			
10	Ala	Val	Gly	Thr	Asn	Leu	Pro	Leu	Gln	Leu	Gly	Phe	Ser	Thr	Gly
		1625					1630					1635			
	Val	Asn	Leu	Val	Ala	Val	Pro	Thr	Gly	Tyr	Val	Asp	Thr	Glu	Asn
		1640					1645					1650			
15	Asn	Thr	Glu	Phe	Thr	Arg	Val	Asn	Ala	Lys	Pro	Pro	Pro	Gly	Asp
		1655					1660					1665			
20	Gln	Phe	Lys	His	Leu	Ile	Pro	Leu	Met	Tyr	Lys	Gly	Leu	Pro	Trp
		1670					1675					1680			
	Asn	Val	Val	Arg	Ile	Lys	Ile	Val	Gln	Met	Leu	Ser	Asp	Thr	Leu
		1685					1690					1695			
25	Lys	Gly	Leu	Ser	Asp	Arg	Val	Val	Phe	Val	Leu	Trp	Ala	His	Gly
		1700					1705					1710			
30	Phe	Glu	Leu	Thr	Ser	Met	Lys	Tyr	Phe	Val	Lys	Ile	Gly	Pro	Glu
		1715					1720					1725			
	Arg	Thr	Cys	Cys	Leu	Cys	Asp	Lys	Arg	Ala	Thr	Cys	Phe	Ser	Thr
		1730					1735					1740			
35	Ser	Ser	Asp	Thr	Tyr	Ala	Cys	Trp	Asn	His	Ser	Val	Gly	Phe	Asp
		1745					1750					1755			
40	Tyr	Val	Tyr	Asn	Pro	Phe	Met	Ile	Asp	Val	Gln	Gln	Trp	Gly	Phe
		1760					1765					1770			
	Thr	Gly	Asn	Leu	Gln	Ser	Asn	His	Asp	Gln	His	Cys	Gln	Val	His
		1775					1780					1785			
45	Gly	Asn	Ala	His	Val	Ala	Ser	Cys	Asp	Ala	Ile	Met	Thr	Arg	Cys
		1790					1795					1800			
50	Leu	Ala	Val	His	Glu	Cys	Phe	Val	Lys	Arg	Val	Asp	Trp	Ser	Val
		1805					1810					1815			
	Glu	Tyr	Pro	Ile	Ile	Gly	Asp	Glu	Leu	Arg	Val	Asn	Ser	Ala	Cys
		1820					1825					1830			
55	Arg	Lys	Val	Gln	His	Met	Val	Val	Lys	Ser	Ala	Leu	Leu	Ala	Asp
		1835					1840					1845			

## EP 1 694 829 B1

	Lys	Phe	Pro	Val	Leu	His	Asp	Ile	Gly	Asn	Pro	Lys	Ala	Ile	Lys
		1850					1855					1860			
5	Cys	Val	Pro	Gln	Ala	Glu	Val	Glu	Trp	Lys	Phe	Tyr	Asp	Ala	Gln
		1865					1870					1875			
10	Pro	Cys	Ser	Asp	Lys	Ala	Tyr	Lys	Ile	Glu	Glu	Leu	Phe	Tyr	Ser
		1880					1885					1890			
	Tyr	Ala	Thr	His	His	Asp	Lys	Phe	Thr	Asp	Gly	Val	Cys	Leu	Phe
		1895					1900					1905			
15	Trp	Asn	Cys	Asn	Val	Asp	Arg	Tyr	Pro	Ala	Asn	Ala	Ile	Val	Cys
		1910					1915					1920			
20	Arg	Phe	Asp	Thr	Arg	Val	Leu	Ser	Asn	Leu	Asn	Leu	Pro	Gly	Cys
		1925					1930					1935			
	Asp	Gly	Gly	Ser	Leu	Tyr	Val	Asn	Lys	His	Ala	Phe	His	Thr	Pro
		1940					1945					1950			
25	Ala	Phe	Asp	Lys	Ser	Ala	Phe	Thr	Asn	Leu	Lys	Gln	Leu	Pro	Phe
		1955					1960					1965			
30	Phe	Tyr	Tyr	Ser	Asp	Ser	Pro	Cys	Glu	Ser	His	Gly	Lys	Gln	Val
		1970					1975					1980			
	Val	Ser	Asp	Ile	Asp	Tyr	Val	Pro	Leu	Lys	Ser	Ala	Thr	Cys	Ile
		1985					1990					1995			
35	Thr	Arg	Cys	Asn	Leu	Gly	Gly	Ala	Val	Cys	Arg	His	His	Ala	Asn
		2000					2005					2010			
40	Glu	Tyr	Arg	Gln	Tyr	Leu	Asp	Ala	Tyr	Asn	Met	Met	Ile	Ser	Ala
		2015					2020					2025			
	Gly	Phe	Ser	Leu	Trp	Ile	Tyr	Lys	Gln	Phe	Asp	Thr	Tyr	Asn	Leu
		2030					2035					2040			
45	Trp	Asn	Thr	Phe	Thr	Arg	Leu	Gln	Ser	Leu	Glu	Asn	Val	Ala	Tyr
		2045					2050					2055			
50	Asn	Val	Val	Asn	Lys	Gly	His	Phe	Asp	Gly	His	Ala	Gly	Glu	Ala
		2060					2065					2070			
	Pro	Val	Ser	Ile	Ile	Asn	Asn	Ala	Val	Tyr	Thr	Lys	Val	Asp	Gly
		2075					2080					2085			
55	Ile	Asp	Val	Glu	Ile	Phe	Glu	Asn	Lys	Thr	Thr	Leu	Pro	Val	Asn
		2090					2095					2100			

EP 1 694 829 B1

	Val	Ala	Phe	Glu	Leu	Trp	Ala	Lys	Arg	Asn	Ile	Lys	Pro	Val	Pro
		2105					2110					2115			
5	Glu	Ile	Lys	Ile	Leu	Asn	Asn	Leu	Gly	Val	Asp	Ile	Ala	Ala	Asn
		2120					2125					2130			
10	Thr	Val	Ile	Trp	Asp	Tyr	Lys	Arg	Glu	Ala	Pro	Ala	His	Val	Ser
		2135					2140					2145			
	Thr	Ile	Gly	Val	Cys	Thr	Met	Thr	Asp	Ile	Ala	Lys	Lys	Pro	Thr
		2150					2155					2160			
15	Glu	Ser	Ala	Cys	Ser	Ser	Leu	Thr	Val	Leu	Phe	Asp	Gly	Arg	Val
		2165					2170					2175			
20	Glu	Gly	Gln	Val	Asp	Leu	Phe	Arg	Asn	Ala	Arg	Asn	Gly	Val	Leu
		2180					2185					2190			
	Ile	Thr	Glu	Gly	Ser	Val	Lys	Gly	Leu	Thr	Pro	Ser	Lys	Gly	Pro
		2195					2200					2205			
25	Ala	Gln	Ala	Ser	Val	Asn	Gly	Val	Thr	Leu	Ile	Gly	Glu	Ser	Val
		2210					2215					2220			
30	Lys	Thr	Gln	Phe	Asn	Tyr	Phe	Lys	Lys	Val	Asp	Gly	Ile	Ile	Gln
		2225					2230					2235			
	Gln	Leu	Pro	Glu	Thr	Tyr	Phe	Thr	Gln	Ser	Arg	Asp	Leu	Glu	Asp
		2240					2245					2250			
35	Phe	Lys	Pro	Arg	Ser	Gln	Met	Glu	Thr	Asp	Phe	Leu	Glu	Leu	Ala
		2255					2260					2265			
40	Met	Asp	Glu	Phe	Ile	Gln	Arg	Tyr	Lys	Leu	Glu	Gly	Tyr	Ala	Phe
		2270					2275					2280			
	Glu	His	Ile	Val	Tyr	Gly	Asp	Phe	Ser	His	Gly	Gln	Leu	Gly	Gly
		2285					2290					2295			
45	Leu	His	Leu	Met	Ile	Gly	Leu	Ala	Lys	Arg	Ser	Gln	Asp	Ser	Pro
		2300					2305					2310			
50	Leu	Lys	Leu	Glu	Asp	Phe	Ile	Pro	Met	Asp	Ser	Thr	Val	Lys	Asn
		2315					2320					2325			
	Tyr	Phe	Ile	Thr	Asp	Ala	Gln	Thr	Gly	Ser	Ser	Lys	Cys	Val	Cys
		2330					2335					2340			
55	Ser	Val	Ile	Asp	Leu	Leu	Leu	Asp	Asp	Phe	Val	Glu	Ile	Ile	Lys
		2345					2350					2355			

EP 1 694 829 B1

	Ser	Gln	Asp	Leu	Ser	Val	Ile	Ser	Lys	Val	Val	Lys	Val	Thr	Ile
		2360					2365					2370			
5	Asp	Tyr	Ala	Glu	Ile	Ser	Phe	Met	Leu	Trp	Cys	Lys	Asp	Gly	His
		2375					2380					2385			
10	Val	Glu	Thr	Phe	Tyr	Pro	Lys	Leu	Gln	Ala	Ser	Gln	Ala	Trp	Gln
		2390					2395					2400			
	Pro	Gly	Val	Ala	Met	Pro	Asn	Leu	Tyr	Lys	Met	Gln	Arg	Met	Leu
		2405					2410					2415			
15	Leu	Glu	Lys	Cys	Asp	Leu	Gln	Asn	Tyr	Gly	Glu	Asn	Ala	Val	Ile
		2420					2425					2430			
20	Pro	Lys	Gly	Ile	Met	Met	Asn	Val	Ala	Lys	Tyr	Thr	Gln	Leu	Cys
		2435					2440					2445			
	Gln	Tyr	Leu	Asn	Thr	Leu	Thr	Leu	Ala	Val	Pro	Tyr	Asn	Met	Arg
		2450					2455					2460			
25	Val	Ile	His	Phe	Gly	Ala	Gly	Ser	Asp	Lys	Gly	Val	Ala	Pro	Gly
		2465					2470					2475			
30	Thr	Ala	Val	Leu	Arg	Gln	Trp	Leu	Pro	Thr	Gly	Thr	Leu	Leu	Val
		2480					2485					2490			
	Asp	Ser	Asp	Leu	Asn	Asp	Phe	Val	Ser	Asp	Ala	Asp	Ser	Thr	Leu
		2495					2500					2505			
35	Ile	Gly	Asp	Cys	Ala	Thr	Val	His	Thr	Ala	Asn	Lys	Trp	Asp	Leu
		2510					2515					2520			
40	Ile	Ile	Ser	Asp	Met	Tyr	Asp	Pro	Arg	Thr	Lys	His	Val	Thr	Lys
		2525					2530					2535			
	Glu	Asn	Asp	Ser	Lys	Glu	Gly	Phe	Phe	Thr	Tyr	Leu	Cys	Gly	Phe
		2540					2545					2550			
45	Ile	Lys	Gln	Lys	Leu	Ala	Leu	Gly	Gly	Ser	Ile	Ala	Val	Lys	Ile
		2555					2560					2565			
50	Thr	Glu	His	Ser	Trp	Asn	Ala	Asp	Leu	Tyr	Lys	Leu	Met	Gly	His
		2570					2575					2580			
	Phe	Ser	Trp	Trp	Thr	Ala	Phe	Val	Thr	Asn	Val	Asn	Ala	Ser	Ser
		2585					2590					2595			
55	Ser	Glu	Ala	Phe	Leu	Ile	Gly	Ala	Asn	Tyr	Leu	Gly	Lys	Pro	Lys
		2600					2605					2610			



EP 1 694 829 B1

5  
 Glu Gln Ile Asp Gly Tyr Thr Met His Ala Asn Tyr Ile Phe Trp  
 2615 2620 2625  
 Arg Asn Thr Asn Pro Ile Gln Leu Ser Ser Tyr Ser Leu Phe Asp  
 2630 2635 2640  
 10 Met Ser Lys Phe Pro Leu Lys Leu Arg Gly Thr Ala Val Met Ser  
 2645 2650 2655  
 Leu Lys Glu Asn Gln Ile Asn Asp Met Ile Tyr Ser Leu Leu Glu  
 2660 2665 2670  
 15 Lys Gly Arg Leu Ile Ile Arg Glu Asn Asn Arg Val Val Val Ser  
 2675 2680 2685  
 20 Ser Asp Ile Leu Val Asn Asn  
 2690 2695

25  
 <210> 76  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <223> amorce S/L3/+4932  
 30  
 <400> 76 20  
 ccacacacag cttgtggata 20

35  
 <210> 77  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 40 <223> amorce S/L4/+6401  
 <400> 77 20  
 ccgaagttgt aggcaatgtc 20

45  
 <210> 78  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 50 <220>  
 <223> amorce S/L4/+6964  
 <400> 78  
 ttgtgtgctc cttctattg 20

55  
 <210> 79  
 <211> 20  
 <212> DNA

<213> Artificial sequence  
 <220>  
 <223> amorce S/L4/-/6817  
 5  
 <400> 79  
 ccggcatcca aacataattt 20  
 <210> 80  
 10 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 15 <223> amorce S/L5/-/7633  
 <400> 80  
 tggtcagtag ggttgattgg 20  
 20 <210> 81  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 25 <220>  
 <223> S/L5/-/8127  
 <400> 81  
 catcctttgt gtcaacatcg 20  
 30 <210> 82  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 35 <220>  
 <223> amorce S/L5/-/8633  
 <400> 82  
 40 gtcacgagtg acaccatcct 20  
 <210> 83  
 <211> 20  
 <212> DNA  
 45 <213> Artificial sequence  
 <220>  
 <223> amorce S/L5/+ /7839  
 50 <400> 83  
 atgcgacgag tctgcttcta 20  
 <210> 84  
 <211> 20  
 55 <212> DNA  
 <213> Artificial sequence  
 <220>

<223> amorce S/L5/+8785  
  
 <400> 84  
 5      ttcatagtgc ctggcttacc      20  
  
 <210> 85  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 10  
  
 <220>  
 <223> amorce S/L5/+8255  
  
 <400> 85  
 15      atcttggcgc atgtattgac      20  
  
 <210> 86  
 <211> 20  
 <212> DNA  
 20      <213> Artificial sequence  
  
 <220>  
 <223> amorce S/L6/-9422  
  
 25      <400> 86  
          tgcattagca gcaacaacat      20  
  
 <210> 87  
 <211> 20  
 30      <212> DNA  
          <213> Artificial sequence  
  
 <220>  
 <223> amorce S/L6/-9966  
 35  
  
 <400> 87  
          tctgcagaac agcagaagtg      20  
  
 <210> 88  
 40      <211> 20  
          <212> DNA  
          <213> Artificial sequence  
  
 <220>  
 45      <223> amorce S/L6/-10542  
  
 <400> 88  
          cctgtgcagt ttgtctgtca      20  
  
 50      <210> 89  
          <211> 20  
          <212> DNA  
          <213> Artificial sequence  
  
 55      <220>  
          <223> amorce S/L6/+10677  
  
 <400> 89

ccttgtggca atgaagtaca 20  
 <210> 90  
 <211> 20  
 5 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <223> amorce S/L6/+ /10106  
 10  
 <400> 90  
 atgtcatttg cacagcagaa 20  
 <210> 91  
 15 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 20 <223> amorce S/L6/+ /9571  
 <400> 91  
 cttcaatggt ttgccatgt 20  
 25 <210> 92  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 30 <220>  
 <223> amorce S/L7/- /11271  
 <400> 92  
 35 tgcgagctgt catgagaata 20  
 <210> 93  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 40 <220>  
 <223> amorce S/L7/- /11801  
 <400> 93  
 45 aaccgagagc agtaccacag 20  
 <210> 94  
 <211> 20  
 <212> DNA  
 50 <213> Artificial sequence  
 <220>  
 <223> amorce S/L7/- /12383  
 55 <400> 94  
 ttggctgct gtagtcaatg 20  
 <210> 95

<211> 20  
 <212> DNA  
 <213> Artificial sequence

5      <220>  
       <223> amorce S/L7/+12640

10      <400> 95  
       ctacgacaga tgtcctgtgc      20

15      <210> 96  
       <211> 20  
       <212> DNA  
       <213> Artificial sequence

20      <220>  
       <223> amorce S/L7/+12088

25      <400> 96  
       gagcaggctg tagctaattg      20

30      <210> 97  
       <211> 20  
       <212> DNA  
       <213> Artificial sequence

35      <220>  
       <223> amorce S/L7/+11551

40      <400> 97  
       ttaggctatt gttgctgctg      20

45      <210> 98  
       <211> 20  
       <212> DNA  
       <213> Artificial sequence

50      <220>  
       <223> amorce S/L8/-13160

55      <400> 98  
       cagacaacat gaagcaccac      20

60      <210> 99  
       <211> 20  
       <212> DNA  
       <213> Artificial sequence

65      <220>  
       <223> amorce S/L8/-13704

70      <400> 99  
       cgctgacgtg atatatgtgg      20

75      <210> 100  
       <211> 20  
       <212> DNA  
       <213> Artificial sequence

<220>  
 <223> amorce S/L8/-/14284  
  
 <400> 100  
 5 tgcacaatga aggatacacc 20  
  
 <210> 101  
 <211> 20  
 <212> DNA  
 10 <213> Artificial sequence  
  
 <220>  
 <223> amorce S/L8/+14453  
  
 15 <400> 101  
 acatagctcg cgtctcagtt 20  
  
 <210> 102  
 <211> 20  
 20 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> amorce S/L8/+13968  
 25  
 <400> 102  
 ggcattgtag gcgtactgac 20  
  
 <210> 103  
 <211> 19  
 30 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 35 <223> amorce S/L8/+13401  
  
 <400> 103  
 gtttgcggtg taagtgcag 19  
  
 40 <210> 104  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
  
 45 <220>  
 <223> amorce S/L9/-/15098  
  
 <400> 104  
 50 tagtggcggc tattgacttc 20  
  
 <210> 105  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 55  
 <220>  
 <223> amorce S/L9/-/15677

<400> 105  
 ctaaaccttg agccgcatag 20  
  
 5 <210> 106  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
  
 10 <220>  
 <223> amorce S/L9/-/16247  
  
 <400> 106  
 catggtcata gcagcacttg 20  
  
 15 <210> 107  
 <211> 21  
 <212> DNA  
 <213> Artificial sequence  
  
 20 <220>  
 <223> amorce S/L9/+ /16323  
  
 <400> 107  
 ccaggtgtg atgtcactga t 21  
 25  
 <210> 108  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 30  
 <220>  
 <223> amorce S/L9/+ /15858  
  
 <400> 108  
 ccttaccag atccatcaag 20  
 35  
 <210> 109  
 <211> 20  
 <212> DNA  
 40 <213> Artificial sequence  
  
 <220>  
 <223> amorce S/L9/+ /15288  
  
 45 <400> 109  
 cgcaaacata acacttgctg 20  
  
 <210> 110  
 <211> 20  
 50 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> amorce S/L10/-/16914  
 55  
 <400> 110  
 agtgttggt acaagccagt 20

<210> 111  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 5  
 <220>  
 <223> amorce S/L10/-/17466  
 <400> 111  
 10 gttccaagga acatgtctgg 20  
 <210> 112  
 <211> 20  
 <212> DNA  
 15 <213> Artificial sequence  
 <220>  
 <223> amorce S/L10/-/18022  
 20 <400> 112  
 aggtgcctgt gtaggatgaa 20  
 <210> 113  
 <211> 20  
 25 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <223> amorce S/L10/+18245  
 30 <400> 113  
 gggctgtcat gcaactagag 20  
 <210> 114  
 35 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 40 <223> amorce S/L10/+17663  
 <400> 114  
 tcttacacgc aatcctgctt 20  
 45 <210> 115  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 50 <220>  
 <223> amorce S/L10/+17061  
 <400> 115  
 55 taccatctg ctgcatagt 20  
 <210> 116  
 <211> 20  
 <212> DNA



<213> Artificial sequence  
 <220>  
 <223> amorce S/L11/-/18877  
 5  
 <400> 116  
 gcaagcagaa ttaaccctca 20  
 <210> 117  
 10 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 15 <223> amorce S/L11/-/19396  
 <400> 117  
 agcaccacct aaattgcatc 20  
 20 <210> 118  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 25 <220>  
 <223> amorce S/L11/-/20002  
 <400> 118  
 30 tgggcccttt gaagtggtta 20  
 <210> 119  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 35 <220>  
 <223> amorce S/L11/+ /20245  
 <400> 119  
 40 tcgaacacat cgtttatgga 20  
 <210> 120  
 <211> 20  
 <212> DNA  
 45 <213> Artificial sequence  
 <220>  
 <223> amorce S/L11/+ /19611  
 50 <400> 120  
 gaagcacctg ttccatcat 20  
 <210> 121  
 <211> 20  
 55 <212> DNA  
 <213> Artificial sequence  
 <220>

<223> amorce S/L11/+ /19021  
  
 <400> 121  
 acgatgctca gccatgtagt           20  
 5  
  
 <210> 122  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 10  
  
 <220>  
 <223> amorce SARS/L1/F3/+ /800  
  
 <400> 122  
 gaggtgcagt cactcgctat           20  
 15  
  
 <210> 123  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 20  
  
 <220>  
 <223> amorce SARS/L1/F4/+ /1391  
  
 <400> 123  
 cagagattgg acctgagcat           20  
 25  
  
 <210> 124  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 30  
  
 <220>  
 <223> amorce SARS/L1/F5/+ /1925  
 35  
  
 <400> 124  
 cagcaaacca ctcaattcct           20  
  
 <210> 125  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 40  
  
 <220>  
 <223> amorce SARS/L1/R3/- /1674  
 45  
  
 <400> 125  
 aaatgatggc aacctctca           20  
  
 <210> 126  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 50  
  
 <220>  
 <223> amorce SARS/L1/R4/- /1107  
 55  
  
 <400> 126

cacgtggttg aatgactttg 20  
 <210> 127  
 <211> 20  
 5 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <223> amorce SARS/L1/R5/-/520  
 10  
 <400> 127  
 atttctgcaa ccagctcaac 20  
 <210> 128  
 15 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 20 <223> amorce SARS/L2/F3/+ /2664  
 <400> 128  
 cgcattgtct cctggtttac 20  
 25 <210> 129  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 30 <220>  
 <223> amorce SARS/L2/F4/+ /3232  
 <400> 129  
 gagattgagc cagaaccaga 20  
 35 <210> 130  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
 40 <220>  
 <223> amorce SARS/L2/F5/+ /3746  
 <400> 130  
 45 atgagcaggt tgtcatggat 20  
 <210> 131  
 <211> 20  
 <212> DNA  
 50 <213> Artificial sequence  
 <220>  
 <223> amorce SARS/L2/R3/-/3579  
 55 <400> 131  
 ctgccttaag aagctggatg 20  
 <210> 132

<211> 20  
 <212> DNA  
 <213> Artificial sequence

5      <220>  
       <223> amorce SARS/L2/R4/-/2991

<400> 132  
 tttcttcacc agcatcatca        20

10      <210> 133  
       <211> 20  
       <212> DNA  
       <213> Artificial sequence

15      <220>  
       <223> amorce SARS/L2/R5/-/2529

<400> 133  
 20      caccgttctt gagaacaacc        20

<210> 134  
 <211> 20  
 <212> DNA  
 25      <213> Artificial sequence

<220>  
 <223> amorce SARS/L3/F3/+ /4708

30      <400> 134  
       tctttggctg gctcttacag        20

<210> 135  
 <211> 20  
 35      <212> DNA  
       <213> Artificial sequence

<220>  
 <223> amorce SRAS/L3/F4/+ /5305

40      <400> 135  
       gctggtgatg ctgctaact        20

<210> 136  
 <211> 20  
 45      <212> DNA  
       <213> Artificial sequence

<220>  
 50      <223> amorce SARS/L3/F5/+ /5822

<400> 136  
 ccatcaagcc tgtgtcgat        20

55      <210> 137  
       <211> 20  
       <212> DNA  
       <213> Artificial sequence

<220>  
 <223> amorce SARS/L3/R3/-/5610  
  
 <400> 137  
 5 cagggtgtgc agacatcata 20  
  
 <210> 138  
 <211> 20  
 <212> DNA  
 10 <213> Artificial sequence  
  
 <220>  
 <223> amorce SARS/L3/R4/-/4988  
  
 15 <400> 138  
 aacatcagca ccatccaagt 20  
  
 <210> 139  
 <211> 20  
 20 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> amorce SARS/L3/R5/-/4437  
 25  
 <400> 139  
 atcggacacc atagtcaacg 20  
  
 <210> 140  
 30 <211> 7788  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 35 <223> gène S synthétique  
  
 <400> 140  
  
 40  
  
 45  
  
 50  
  
 55

EP 1 694 829 B1

	tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta	60
	ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc	120
5	aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg	180
	gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc	240
	gcctggctga ccgcccacg acccccgccc attgacgtca ataatgacgt atgttcccat	300
10	agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc	360
	ccacttggca gtacatcaag tgtatcatat gccaagtccg cccctattg acgtcaatga	420
	cggtaaattg cccgcctggc attatgcca gtacatgacc ttacgggact ttcctacttg	480
15	gcagtacatc tacgtattag tcatcgctat taccatgggt atgcggtttt ggcagtacac	540
	caatgggctg ggatagcggc ttgactcacg gggatttcca agtctccacc ccattgacgt	600
	caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaataaccc	660
20	cgccccgttg acgcaaattg gcggtaggcg tgtacgggtg gaggtctata taagcagagc	720
	tcgttttagtg aaccgtcaga tctactagaag ctttattgcg gtagtttatc acagttaa	780
	tgctaacgca gtcagtgtct ctgacacaac agtctcgaac ttaagctgca gaagttggc	840
25	gtgaggcact gggcaggtaa gtatcaagg tacaagacag gtttaaggag accaatagaa	900
	actgggcttg tcgagacaga gaagactctt gcgtttctga taggcaccta ttggtcttac	960
30		
35		
40		
45		
50		
55		

	tgacatccac	tttgcctttc	tctccacagg	tgtccactcc	cagttcaatt	acagctctta	1020
	aggctagagt	acttaatacg	actcactata	ggctagcgga	tccaccatgt	tcattcttct	1080
5	gctgttcctg	accctgacca	gcggcagcga	cctggaccgg	tgcaccacct	tcgacgacgt	1140
	gcaggccccc	aactacaccc	agcacaccag	cagcatgcgg	ggcgtgtact	accccgacga	1200
	gatctttcgg	agcgacaccc	tgtacctgac	ccaggacctg	ttcctgcccct	tctacagcaa	1260
10	cgtgaccggc	ttccacacca	tcaaccacac	cttcggcaac	cccgtgatcc	ccttcaagga	1320
	cggcatctac	ttcgccgcca	ccgagaagag	caacgtggtg	cggggctggg	tgttcggcag	1380
	caccatgaac	aacaagagcc	agagcgtgat	catcatcaac	aacagcacca	acgtggtgat	1440
15	ccgggcctgc	aacttcgagc	tgtgcgacaa	ccccttcttc	gccgtgtcca	aacctatggg	1500
	caccagacc	cacaccatga	tcttcgacaa	cgccttcaac	tgcaccttcg	agtacatcag	1560
	cgacgccttc	agcctggacg	tgagcgagaa	gagcggcaac	ttcaagcacc	tgcgggagtt	1620
20	cgtgttcaag	aacaaggacg	gcttcctgta	cgtgtacaag	ggctaccagc	ccatcgacgt	1680
	ggtgagagac	ctgcccagcg	gcttcaacac	cctgaagccc	atcttcaagc	tgcccctggg	1740
	catcaacatc	accaacttcc	gggccatcct	gaccgccttt	agccctgccc	aggacatctg	1800
25	gggcaccagc	gccgccgcct	acttcgtggg	ctacctgaag	cctaccacct	tcattgtgaa	1860
	gtacgacgag	aacggcacca	tcaccgacgc	cgtggactgc	agccagaacc	ccctggccga	1920
	gctgaagtgc	agcgtgaaga	gcttcgagat	cgacaagggc	atctaccaga	ccagcaactt	1980
30	cagagtgggtg	cctagcggcg	atgtgggtgcg	gttccccaat	atcaccaacc	tgtgcccctt	2040
	cggcgaagtg	ttcaacgcca	ccaagtctcc	cagcgtgtac	gcctgggagc	ggaagaagat	2100
	cagcaactgc	gtggccgact	acagcgtgct	gtacaactcc	accttcttca	gcaccttcaa	2160
35	gtgctacggc	gtgagcgcca	ccaagctgaa	cgacctgtgc	ttcagcaacg	tgtacgccga	2220
	cagcttcgtg	gtgaagggcg	acgacgtgag	acagatcgcc	cctggccaga	ccggcgtgat	2280
	cgccgactac	aactacaagc	tgcccgacga	cttcatgggc	tgcggtgctgg	cctggaacac	2340
40	ccggaacatc	gacgccacaa	gcaccggcaa	ctacaattac	aagtaccgct	acctgcgga	2400
	cggcaagctg	cggcccttcg	agcgggacat	ctccaacgtg	cccttcagcc	ccgacggcaa	2460
	gccctgcacc	ccccctgccc	tgaactgcta	ctggcccctg	aacgactacg	gcttctacac	2520
45	caccaccggc	atcggctatc	agccctacag	agtgggtggg	ctgagcttcg	agctgctgaa	2580
	cgcccctgcc	accgtgtgcg	gcccgaagct	gagcaccgac	ctgatcaaga	accagtgcgt	2640
	gaacttcaac	ttcaacggcc	tgaccggcac	cggcgtgctg	acccccagca	gcaagcgctt	2700
50	ccagcccttc	cagcagttcg	gccgggatgt	gagcgacttc	accgacagcg	tgcgggaccc	2760
	caagaccagc	gagatcctgg	acatcagccc	ctgcagcttc	ggcggcgtgt	ccgtgatcac	2820
	ccccggcacc	aacgccagca	gcgaagtggc	cgtgctgtac	caggacgtga	actgcaccga	2880
55	cgtgagcacc	gccatccacg	ccgaccagct	gacccccgcc	tggcggatct	acagaccggg	2940
	gaacaacgtg	ttccagaccc	aggccggctg	cctgatcggc	gccgagcacg	tggacaccag	3000

EP 1 694 829 B1

	ctacgagtgc	gacatcccca	ttggcgccgg	aatctgcgcc	agctaccaca	ccgtgagcct	3060
	gctgcggagc	accagccaga	agtccatcgt	ggcctacacc	atgagcctgg	gcgccgacag	3120
5	cagcatcgcc	tacagcaaca	acaccatcgc	catccccacc	aacttcagca	tctccatcac	3180
	caccgaagtg	atgcccgtga	gcatggccaa	gacaagcgtg	gattgcaaca	tgtacatctg	3240
	cggcgacagc	accgagtgcg	ccaacctgct	gctgcagtac	ggcagcttct	gcacccagct	3300
10	gaaccggggc	ctgagcggca	tcgccgccga	gcaggaccgg	aacaccagag	aagtgttcgc	3360
	ccaagtgaag	cagatgtata	agacccccac	cctgaagtac	ttcggggggt	tcaacttctc	3420
	tcagatcctg	cccgaccctc	tgaagcccac	caagcgctcc	ttcatcgagg	acctgctggt	3480
15	caacaaagtg	accctggccg	acgccggctt	tatgaagcag	tacggcgagt	gcctgggcga	3540
	catcaacgcc	cgggacctga	tctgcgcccc	gaagtttaac	gggctgaccg	tgctgcccc	3600
	cctgctgacc	gacgacatga	tcgccgccta	tacagccgcc	ctggtgagcg	gcaccgccac	3660
20	cgccggctgg	accttcggag	ccggagccgc	cctgcagatc	cccttcgcc	tgcatgtggc	3720
	ctaccggttc	aacggcatcg	gcgtgaccca	gaacgtgctg	tacgagaacc	agaagcagat	3780
	cgccaaccag	ttcaacaagg	ccatcagcca	gatccaggag	agcctgacca	caaccagcac	3840
25	cgccctgggc	aagctgcagg	acgtggtgaa	ccagaacgcc	caggccctga	acaccctggt	3900
	gaagcagctg	agcagcaact	tcggcgccat	cagctctgtg	ctgaacgaca	tcctgagcag	3960
	gctggacaaa	gtggaggccg	aagtgcagat	cgaccggctg	atcaccggac	gcctgcagtc	4020
30	cctgcagacc	tacgtgaccc	agcagctgat	cagagccgcc	gagatccggg	ccagcgccaa	4080
	tctggccgcc	accaagatga	gcgagtgcgt	gctgggccag	agcaagagag	tggaacttctg	4140
	cggcaagggc	tatcacctga	tgagcttccc	ccaggccgcc	ccccacggcg	tggtgttcct	4200
35	gcacgtgacc	tacgtgccta	gccaggagcg	gaacttcacc	accgccccag	ccatctgcc	4260
	cgagggcaag	gcctacttcc	cccgggaggg	cgtgttcgtg	tttaacggca	ccagctgggt	4320
	catcacccag	cgcaacttct	tcagccccc	gatcatcacc	acagacaaca	ccttcgtgtc	4380
40	cggcaactgt	gatgtggtga	tcggcatcat	caataacacc	gtgtacgacc	ccctgcagcc	4440
	cgagctggac	agcttcaagg	aggagctgga	caaatacttc	aagaaccaca	cctccccga	4500
	cgtggacctg	ggcgatatca	gcggcatcaa	cgcctccgtg	gtgaacatcc	agaaggagat	4560
45	cgacagactg	aacgaagtgg	ccaagaacct	gaacgagagc	ctgatcgacc	tgaggagct	4620
	gggcaagtac	gagcagtaca	tcaagtggcc	ctggtacgtg	tggtggggt	tcacgcggg	4680
	cctgatcgcc	atcgtgatgg	tgaccatcct	gctgtgctgc	atgaccagct	gctgtagctg	4740
50	cctgaaaggc	gcctgcagct	gtggcagctg	ctgcaagttc	gacgaggacg	acagcgagcc	4800
	cgtgctgaag	ggcgtgaagc	tgactacac	ctgataactc	gagaattcac	gcgtggtacc	4860
	tctagagtcg	acccggggcg	ccgcttcgag	cagacatgat	aagatacatt	gatgagtttg	4920
55	gacaaaccac	aactagaatg	cagtgaaaaa	aatgctttat	ttgtgaaatt	tgtgatgcta	4980
	ttgctttatt	tgtaaccatt	ataagctgca	ataaacaagt	taacaacaac	aattgcattc	5040



EP 1 694 829 B1

	attttatgtt tcaggttcag ggggagatgt gggaggtttt ttaaagcaag taaaacctct	5100
	acaaatgttg taaaatcgat aaggatccgg gctggcgtaa tagcgaagag gcccgcaccg	5160
5	atcgcccttc ccaacagttg cgcagcctga atggcgaatg gacgcgccct gtagcggcgc	5220
	attaagcgcg gcggggtgtg tggttacgcg cagcgtgacc gctacacttg ccagcgcctt	5280
	agcgcgccgt cctttcgctt tcttcccttc ctttctcgcc acgttcgccg gctttccccg	5340
10	tcaagctcta aatcgggggc tccctttagg gttccgattt agagctttac ggcacctcga	5400
	ccgcaaaaaa cttgatttgg gtgatggttc acgtagtggg ccatcgccct gatagacggt	5460
	ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt tccaaactgg	5520
15	aacaacactc aaccctatct cggctctattc ttttgattta taagggattt tgccgatttc	5580
	ggcctatttg ttaaaaaatg agctgattta acaaataatt aacgcgaatt ttaacaaaat	5640
	attaacgttt acaatttcgc ctgatgcggt attttctcct tacgcatctg tgcggtattt	5700
20	cacaccgcat atggtgcact ctgagtacaa tctgctctga tgccgcatag ttaagccagc	5760
	cccgacaccc gccaacaccc gctgacgcgc cctgacgggc ttgtctgctc ccggcatccg	5820
	cttacagaca agctgtgacc gtctccggga gctgcatgtg tcagaggttt tcaccgtcat	5880
25	caccgaaacg cgcgagacga aagggcctcg tgatacgcct atttttatag gttaatgtca	5940
	tgataataat ggtttcttag acgtcaggtg gcacttttcg gggaaatgtg cgcggaaccc	6000
	ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct	6060
30	gataaatgct tcaataatat tgaaaaagga agagtatgag tattcaacat ttccgtgtcg	6120
	cccttatttc cttttttgcg gcattttgcc ttcctgtttt tgctcaccca gaaacgctgg	6180
	tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc gaactggatc	6240
35	tcaacagcgg taagatcctt gagagttttc gccccgaaga acgttttcca atgatgagca	6300
	cttttaaaagt tctgctatgt ggcgcggtat tatcccgtat tgacgccggg caagagcaac	6360
	tcggtcgccg catacactat tctcagaatg acttggttga gtactacca gtcacagaaa	6420
40	agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata accatgagtg	6480
	ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag ctaaccgctt	6540
	ttttgcacaa catgggggat catgtaactc gccttgatcg ttgggaaccg gagctgaatg	6600
45	aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaacgttgc	6660
	gcaaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta atagactgga	6720
	tggaggcgga taaagttgca ggaccacttc tgcgctcggc ccttccggct ggctggttta	6780
50	ttgctgataa atctggagcc ggtgagcgtg ggtctcgcgg tatcattgca gcactggggc	6840
	cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag gcaactatgg	6900
	atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat tggttaactgt	6960
55	cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa	7020
	ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt	7080

5 cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga gatccttttt 7140  
 ttctgcgcggt aatctgctgc ttgcaaaca aaaaaccacc gctaccagcg gtggtttgtt 7200  
 tgccggatca agagctacca actctttttt cgaaggtaac tggcttcagc agagcgcaga 7260  
 taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag aactctgtag 7320  
 10 caccgcctac atacctcgct ctgctaatac tgttaccagt ggctgctgcc agtggcgata 7380  
 agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg cagcgggtcgg 7440  
 gctgaacggg ggggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga 7500  
 15 gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca 7560  
 ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt ccagggggaa 7620  
 acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt 7680  
 20 tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg gcctttttac 7740  
 ggttcctggc cttttgctgg ctttttgctc acatggctcg acagatct 7788

25 <210> 141  
 <211> 23  
 <212> DNA  
 <213> Artificial sequence

30 <220>  
 <223> amorce SNE-S1

<400> 141 23  
 ggttgggatt atccaaaatg tga 23

35 <210> 142  
 <211> 24  
 <212> DNA  
 <213> Artificial sequence

40 <220>  
 <223> amorce SNE-AS1

<400> 142 24  
 gcatcatcag aaagaatcat catg 24

45 <210> 143  
 <211> 21  
 <212> DNA  
 <213> Artificial sequence

50 <220>  
 <223> amorce SAR1-S

55 <400> 143  
 cctctcttgt tcttgctgc a 21

<210> 144  
 <211> 21

<212> DNA  
 <213> Artificial sequence

5 <220>  
 <223> amorce SAR1-AS

<400> 144  
 tatagtgagc cgccacacat g 21

10 <210> 145  
 <211> 45  
 <212> DNA  
 <213> Artificial sequence

15 <220>  
 <223> amorce PCR

20 <400> 145  
 ataggatcca ccatgtttat ttcttatta ttcttactc tcact 45

25 <210> 146  
 <211> 37  
 <212> DNA  
 <213> Artificial sequence

30 <220>  
 <223> amorce PCR

35 <400> 146  
 atactcgagt tatgtgtaat gtaatttgac acccttg 37

40 <210> 147  
 <211> 45  
 <212> DNA  
 <213> Artificial sequence

45 <220>  
 <223> amorce PCR

50 <400> 147  
 ataggatcca ccatgtttat ttcttatta ttcttactc tcact 45

55 <210> 148  
 <211> 36  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> amorce PCR

<400> 148  
 acctccggat ttaatatatt gctcatattt tcccaa 36

<210> 149  
 <211> 13  
 <212> PRT  
 <213> Artificial sequence

# EP 1 694 829 B1

<220>

<223> extrémité N-terminale de la protéine S du SRAS-CoV (acides amines 1 à 13)

<400> 149

5

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly  
1 5 10

10

<210> 150

<211> 10

<212> PRT

<213> Artificial sequence

15

<220>

<223> oligopeptide

<400> 150

20

Ser Gly Asp Tyr Lys Asp Asp Asp Asp Lys  
1 5 10

25

<210> 151

<211> 34

<212> DNA

<213> Artificial sequence

30

<220>

<223> amorce PCR

<400> 151

actagctagc ggatccacca tgtcatctt cctg 34

35

<210> 152

<211> 33

<212> DNA

<213> Artificial sequence

40

<220>

<223> amorce PCR

<400> 152

agtatccgga ctgatgtac tgctcgact tgc 33

45

<210> 153

<211> 59

<212> DNA

<213> Artificial sequence

50

<220>

<223> oligonucléotide

<400> 153

tatgagcttt tttttttt ttttttggc atataaatag actcggcgcg ccatctgca 59

55

<210> 154

<211> 53  
 <212> DNA  
 <213> Artificial sequence  
 5 <220>  
 <223> oligonucléotide  
 <400> 154  
 gatggcgcgc cgagtctatt tatatgccaa aaaaaaaaaa aaaaaaaagc tca 53  
 10  
 <210> 155  
 <211> 45  
 <212> DNA  
 <213> Artificial sequence  
 15 <220>  
 <223> amorce PCR  
 <400> 155  
 20 atacgtacga ccatgtttat ttcttatta ttcttactc tcact 45  
 <210> 156  
 <211> 40  
 <212> DNA  
 25 <213> Artificial sequence  
 <220>  
 <223> amorce PCR  
 30 <400> 156  
 atagcgcgcgt cattatgtgt aatgtaattt gacacccttg 40  
 <210> 157  
 <211> 20  
 35 <212> DNA  
 <213> Artificial sequence  
 <220>  
 <223> amorce PCR  
 40 <400> 157  
 ccatttcaac aattggccg 20  
 <210> 158  
 45 <211> 45  
 <212> DNA  
 <213> Artificial sequence  
 <220>  
 50 <223> amorce PCR  
 <400> 158  
 ataggatccg cgcgctcatt atttatcgtc gtcatcttta taatc 45  
 55

## Revendications

1. Souche isolée ou purifiée de coronavirus humain associé au syndrome respiratoire aigu sévère, **caractérisée en**

**ce que** son génome présente sous la forme d'ADN complémentaire un codon sérine en position 23220-23222 du gène de la protéine S ou un codon glycine en position 25298-25300 du gène de l'ORF3, et un codon alanine en position 7918-7920 de l'ORF1a ou un codon sérine en position 26857-26859 du gène de la protéine M, lesdites positions étant indiquées en référence à la séquence Genbank AY274119.3.

2. Souche isolée ou purifiée de coronavirus selon la revendication 1, **caractérisée en ce que** l'équivalent ADN de son génome présente une séquence correspondant à la séquence SEQ ID NO : 1.
3. Polynucléotide isolé ou purifié, **caractérisé en ce que** sa séquence est celle du génome de la souche isolée de coronavirus selon la revendication 1 ou la revendication 2.
4. Polynucléotide isolé ou purifié selon la revendication 3, **caractérisé en ce que** sa séquence est la SEQ ID NO : 1.
5. Fragment du polynucléotide selon la revendication 3, **caractérisé en ce qu'il** comprend les ou est constitué des bases ou paires de bases consécutives des positions 7919 à 23220, 7919 à 25298, 16622 à 23220, 19064 à 23220, 16622 à 25298, 19064 à 25298, 23220 à 24872, 23220 à 26857, 24872 à 25298, ou 25298 à 26857.
6. Réactif de détection d'un coronavirus associé au SRAS, **caractérisé en ce qu'il** est sélectionné dans le groupe constitué par une souche isolée de coronavirus selon la revendication 1 ou la revendication 2 et un polynucléotide selon l'une quelconque des revendications 3, 4 et 5.
7. Utilisation d'un produit sélectionné dans le groupe constitué par :
  - (a) une souche isolée de coronavirus selon la revendication 1 ou la revendication 2,
  - (b) un polynucléotide selon la revendication 3, la revendication 4 ou la revendication 5,
 pour la préparation d'un réactif de détection et éventuellement de génotypage d'un coronavirus associé au SRAS.
8. Kit ou coffret de détection d'un coronavirus associé au SRAS, **caractérisé en ce qu'il** comprend au moins un réactif sélectionné dans le groupe constitué par :
  - une souche isolée de coronavirus selon la revendication 1 ou la revendication 2 et,
  - un polynucléotide selon la revendication 3, la revendication 4 ou la revendication 5.

## Claims

1. Isolated or purified strain of human coronavirus associated with severe acute respiratory syndrome, **characterised in that** its genome exhibits, in the form of complementary DNA, a serine codon at position 23220-23222 of the gene for the S protein or a glycine codon at position 25298-25300 of the gene for ORF3, and an alanine codon at position 7918-7920 for ORF1a or a serine codon at position 26857-26859 of the gene for the M protein, said positions being indicated with reference to GenBank sequence AY274119.3.
2. Isolated or purified coronavirus strain according to Claim 1, **characterised in that** the DNA equivalent of its genome exhibits a sequence corresponding to sequence SEQ ID No. 1.
3. Isolated or purified polynucleotide, **characterised in that** its sequence is that of the genome of the isolated coronavirus strain according to Claim 1 or Claim 2.
4. Isolated or purified polynucleotide according to Claim 3, **characterised in that** its sequence is SEQ ID No. 1.
5. Fragment of the polynucleotide according to Claim 3, **characterised in that** it includes or is constituted by the consecutive bases or base pairs of positions 7919 to 23220, 7919 to 25298, 16622 to 23220, 19064 to 23220, 16622 to 25298, 19064 to 25298, 23220 to 24872, 23220 to 26857, 24872 to 25298, or 25298 to 26857.
6. Reagent for detection of a coronavirus associated with SARS, **characterised in that** it is selected from the group constituted by an isolated coronavirus strain according to Claim 1 or Claim 2 and a polynucleotide according to any one of Claims 3, 4 and 5.

7. Use of a product selected from the group constituted by

- (a) an isolated coronavirus strain according to Claim 1 or Claim 2,
- (b) a polynucleotide according to Claim 3, Claim 4 or Claim 5,

for the preparation of a detection reagent and possibly for genotyping of a coronavirus associated with SARS.

8. Kit or box for detecting a coronavirus associated with SARS, **characterised in that** it includes at least one reagent selected from the group constituted by:

- an isolated coronavirus strain according to Claim 1 or Claim 2 and,
- a polynucleotide according to Claim 3, Claim 4 or Claim 5.

**Patentansprüche**

1. Isolierter oder gereinigter humaner Coronavirusstamm, der mit dem Schweren Akuten Respiratorischen Syndrom assoziiert ist, **dadurch gekennzeichnet, dass** sein Genom in Form von komplementärer DNA an Position 23220-23222 des Gens für das S-Protein ein Serincodon oder an Position 25298-25300 des Gens für den ORF3 ein Glycincodon, und ein Alanincodon an Position 7918-7920 von ORF1a oder ein Serincodon an Position 26857-26859 des Gens für das M-Protein hat, wobei die Positionen in Bezug auf die Genbanksequenz AY274119.3 angegeben sind.

2. Isolierter oder gereinigter Coronavirusstamm nach Anspruch 1, **dadurch gekennzeichnet, dass** die DNA, die seinem Genom äquivalent ist, eine Sequenz hat, die der Sequenz SEQ ID NO:1 entspricht.

3. Isoliertes oder gereinigtes Polynukleotid, **dadurch gekennzeichnet, dass** seine Sequenz jene des Genoms des isolierten Coronavirusstamms nach Anspruch 1 oder Anspruch 2 ist.

4. Isoliertes oder gereinigtes Polynukleotid nach Anspruch 3, **dadurch gekennzeichnet, dass** seine Sequenz SEQ ID NO: 1 ist.

5. Fragment des Polynukleotids nach Anspruch 3, **dadurch gekennzeichnet, dass** es die aufeinanderfolgenden Basen oder Basenpaare der Positionen 7919 bis 23220, 7919 bis 25298, 16622 bis 23220, 19064 bis 23220, 16622 bis 25298, 19064 bis 25298, 23220 bis 24872, 23220 bis 26857, 24872 bis 25298 oder 25298 bis 26857 umfasst oder aus diesen besteht.

6. Reagenz zum Nachweis eines mit SARS assoziierten Coronavirus, **dadurch gekennzeichnet, dass** es ausgewählt ist aus der Gruppe bestehend aus einem isolierten Coronavirusstamm nach Anspruch 1 oder Anspruch 2 und einem Polynukleotid nach einem der Ansprüche 3, 4 und 5.

7. Verwendung eines Produktes ausgewählt aus der Gruppe bestehend aus:

- (a) einem isolierten Coronavirusstamm nach Anspruch 1 oder Anspruch 2,
- (b) einem Polynukleotid nach Anspruch 3, Anspruch 4 oder Anspruch 5,

zur Herstellung eines Reagenzes zum Nachweis und gegebenenfalls zur Genotypisierung eines mit SARS assoziierten Coronavirus.

8. Kit oder Set zum Nachweis eines mit SARS assoziierten Coronavirus, **dadurch gekennzeichnet, dass** es wenigstens ein Reagenz umfasst, ausgewählt aus der Gruppe bestehend aus:

- einem isolierten Coronavirusstamm nach Anspruch 1 oder Anspruch 2, und
- einem Polynukleotid nach Anspruch 3, Anspruch 4 oder Anspruch 5.

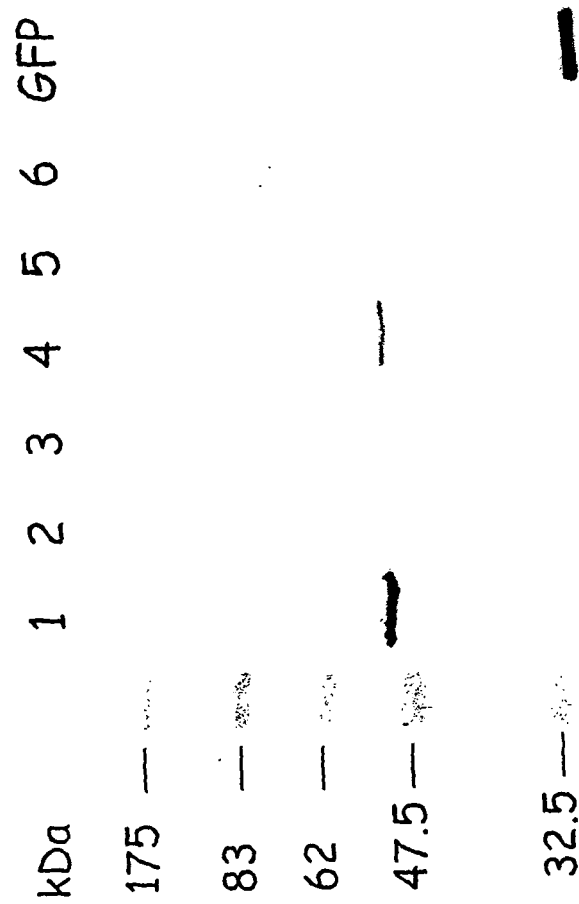


FIGURE 1



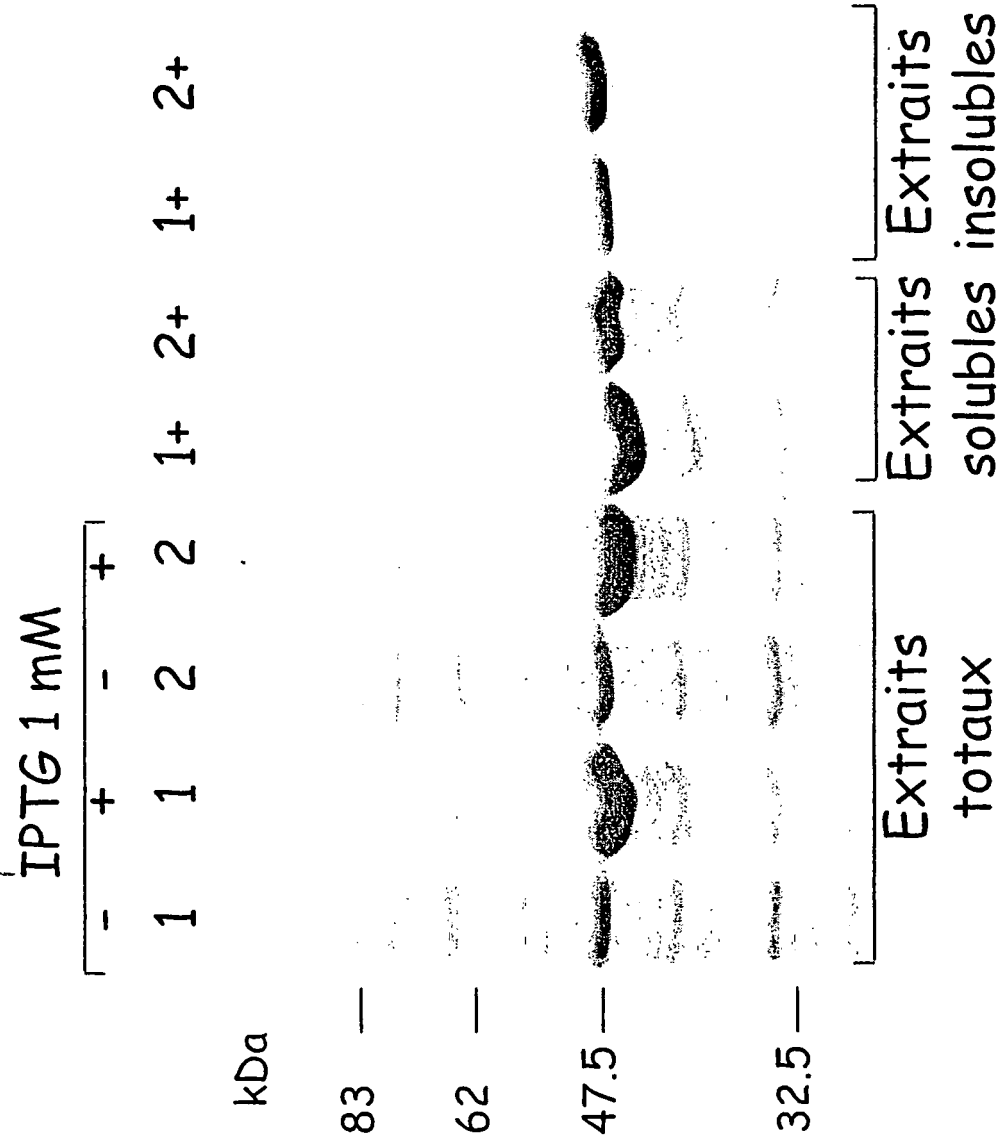


FIGURE 2

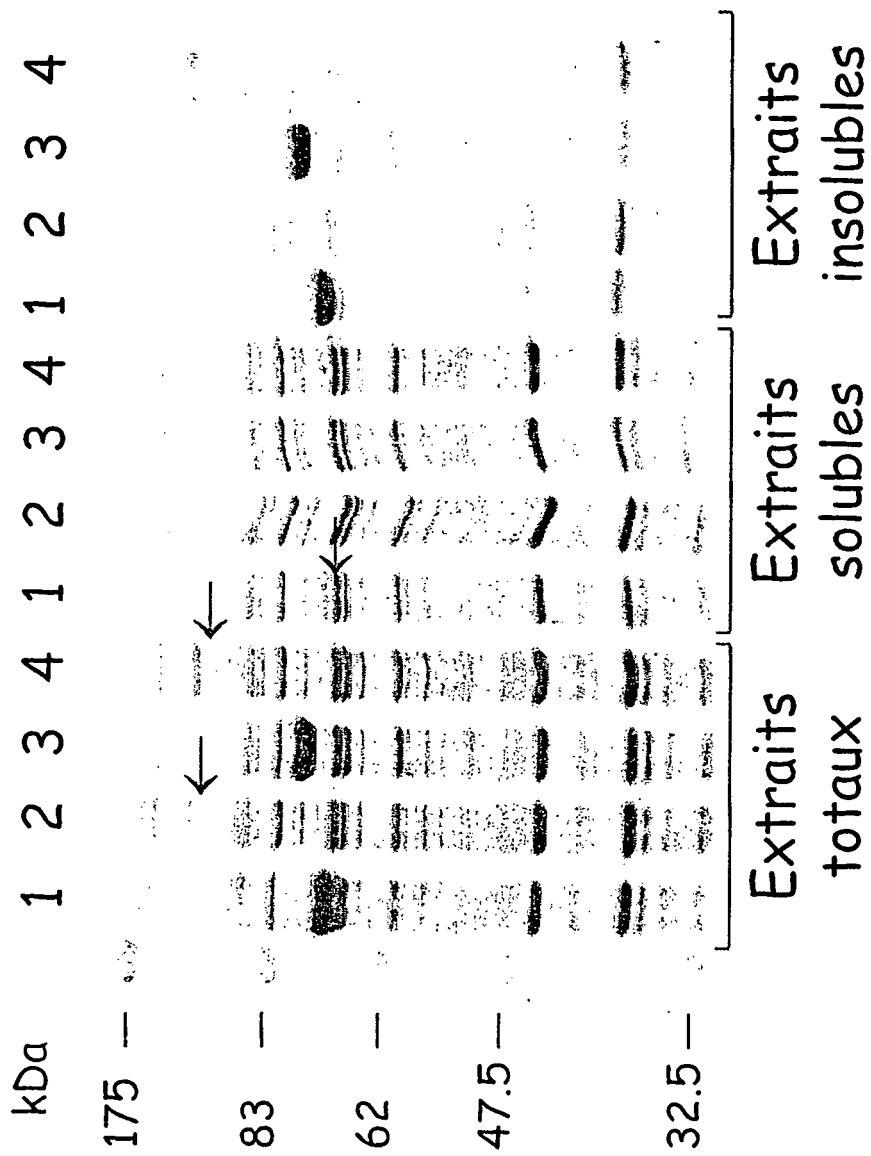


FIGURE 3

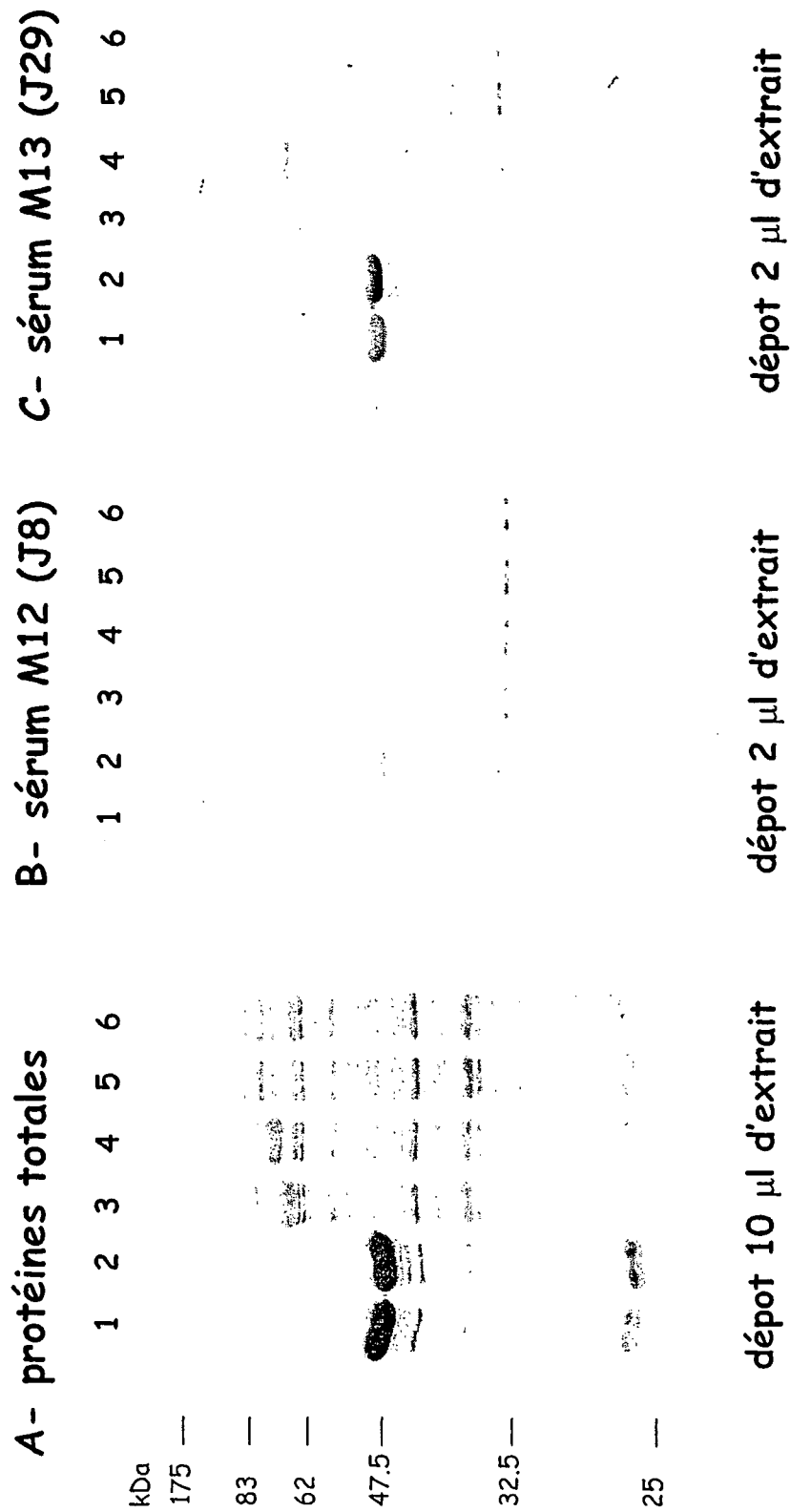


FIGURE 4

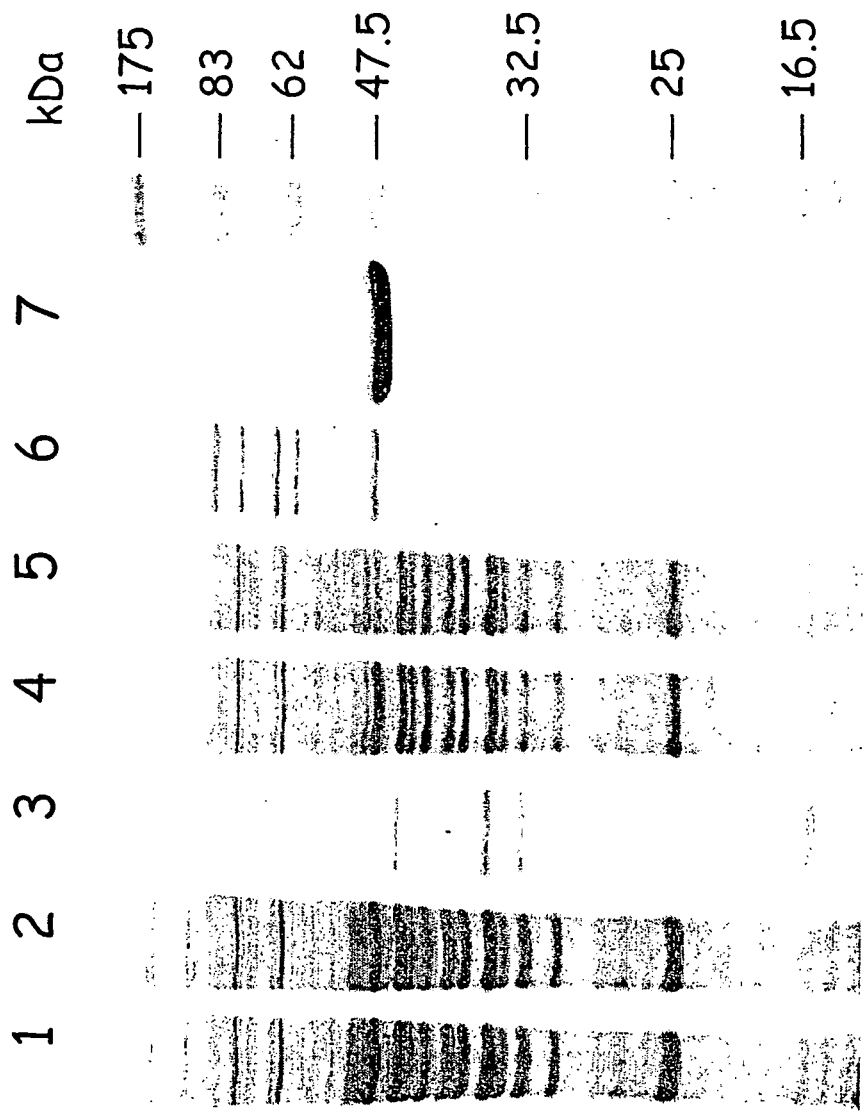


FIGURE 5

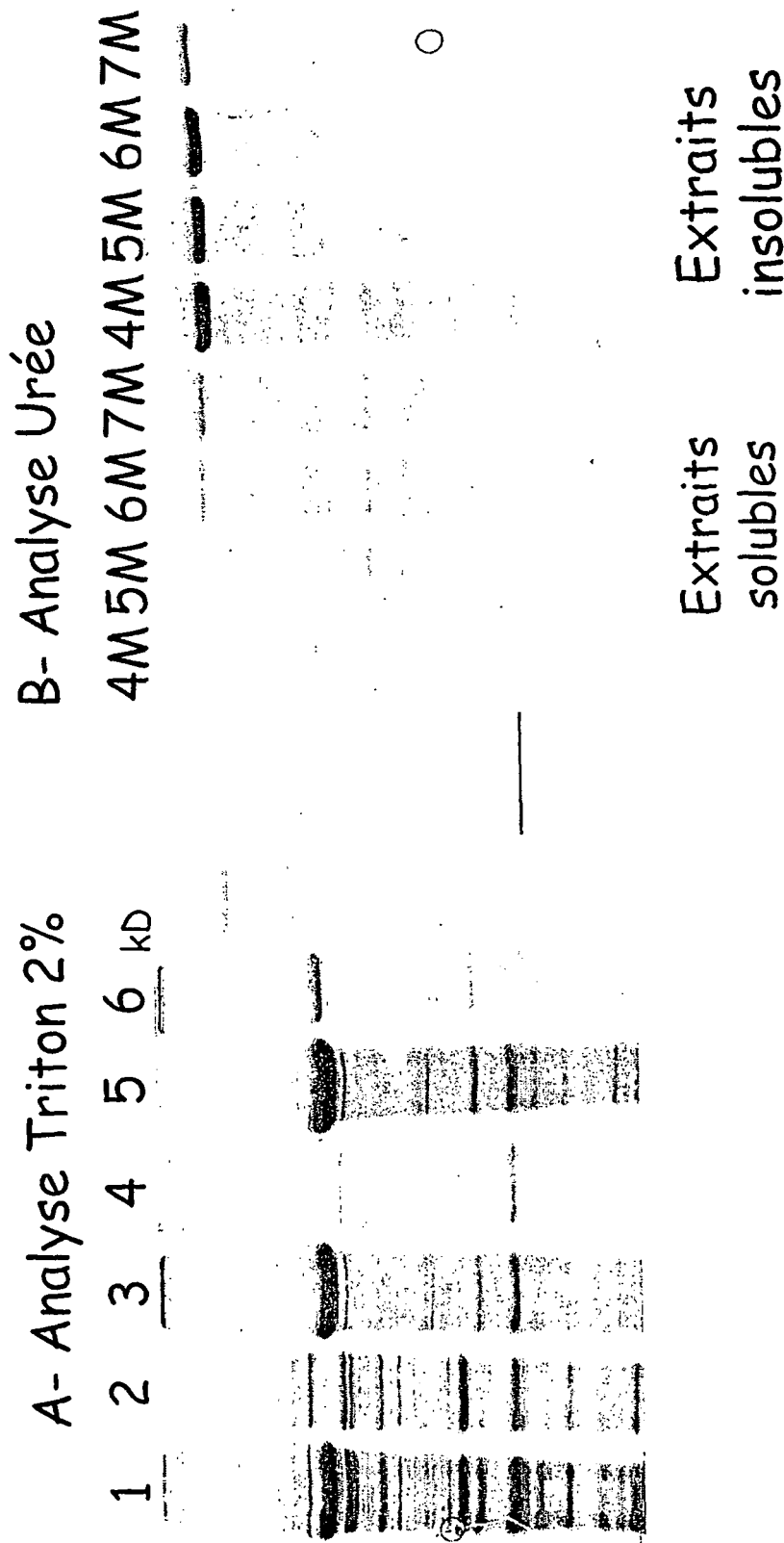


FIGURE 6

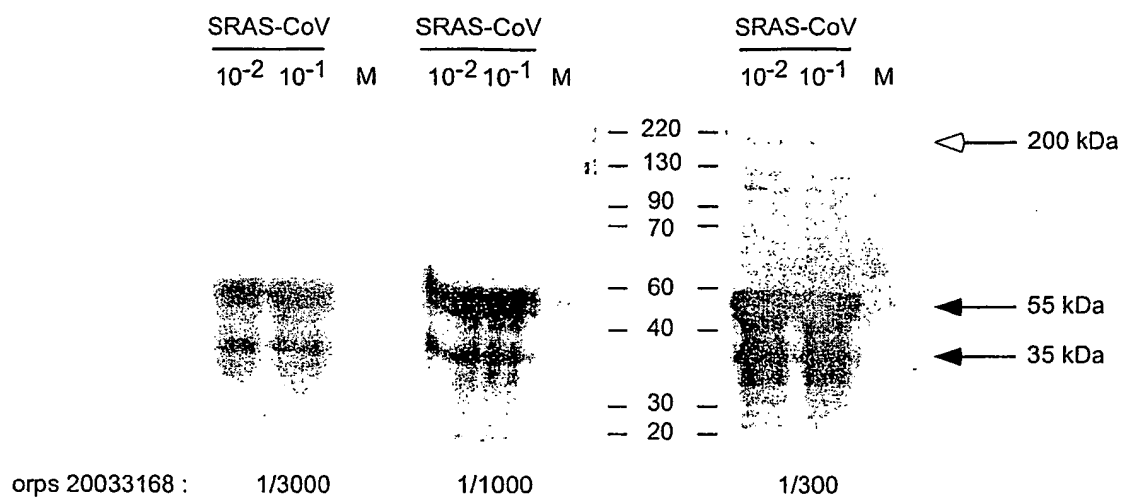
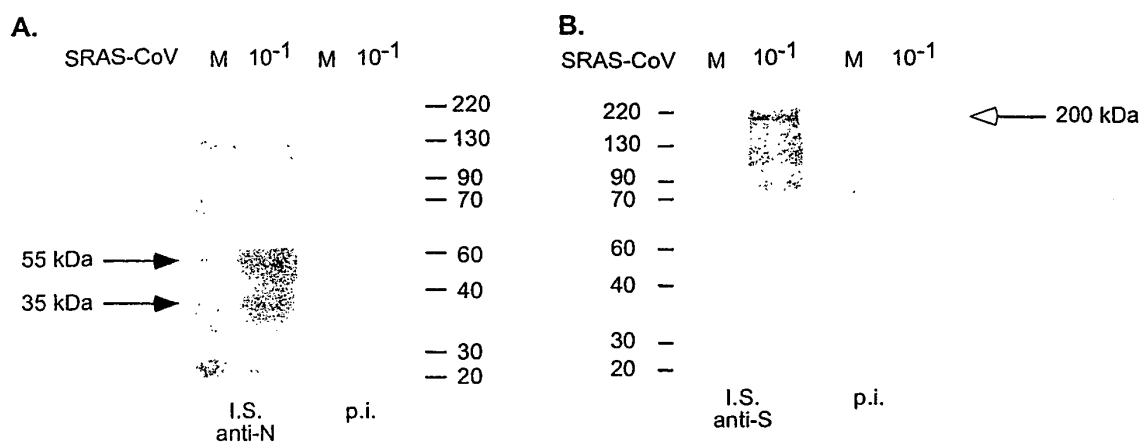
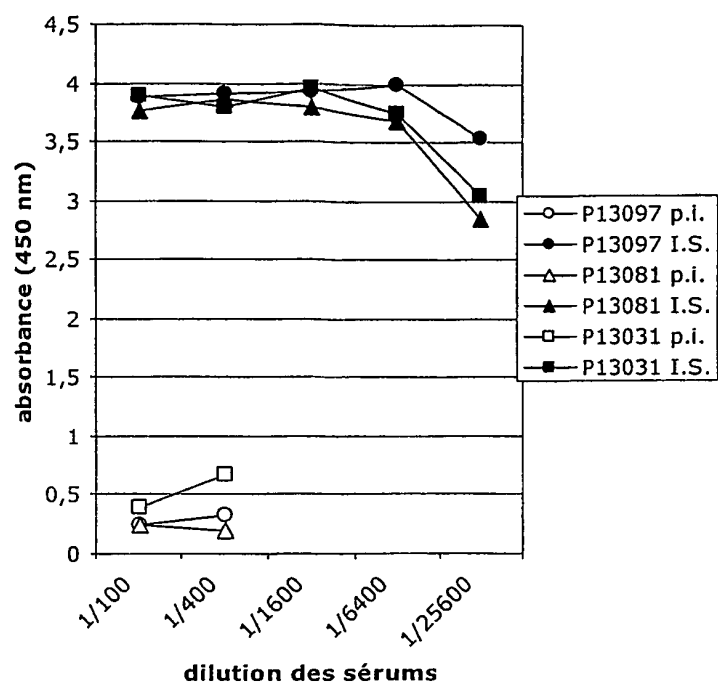


FIGURE 7



**FIGURE 8**

A



B

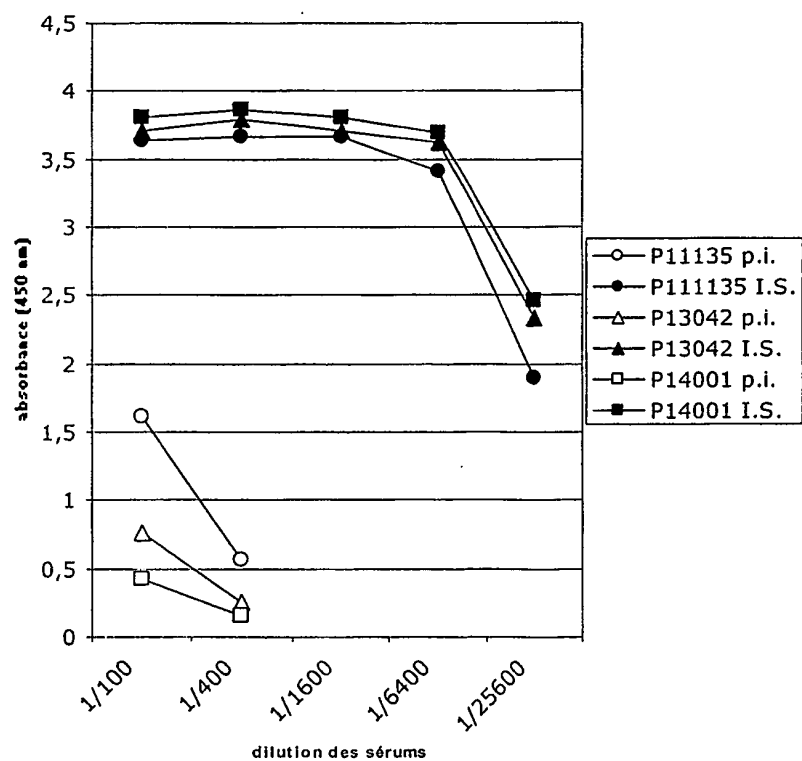


FIGURE 9



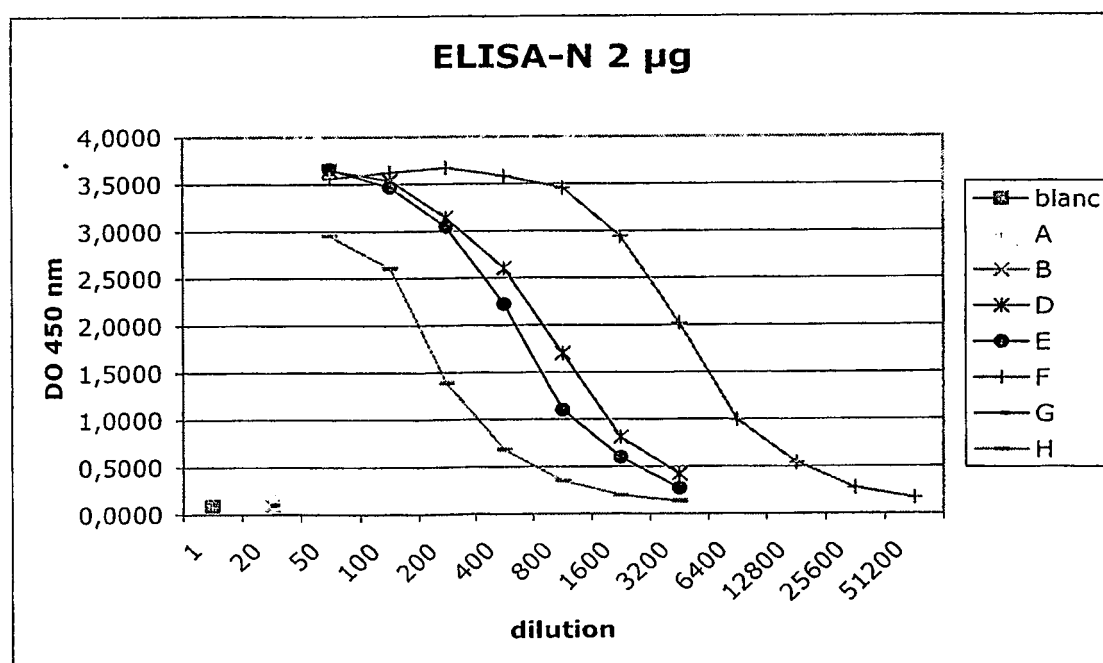
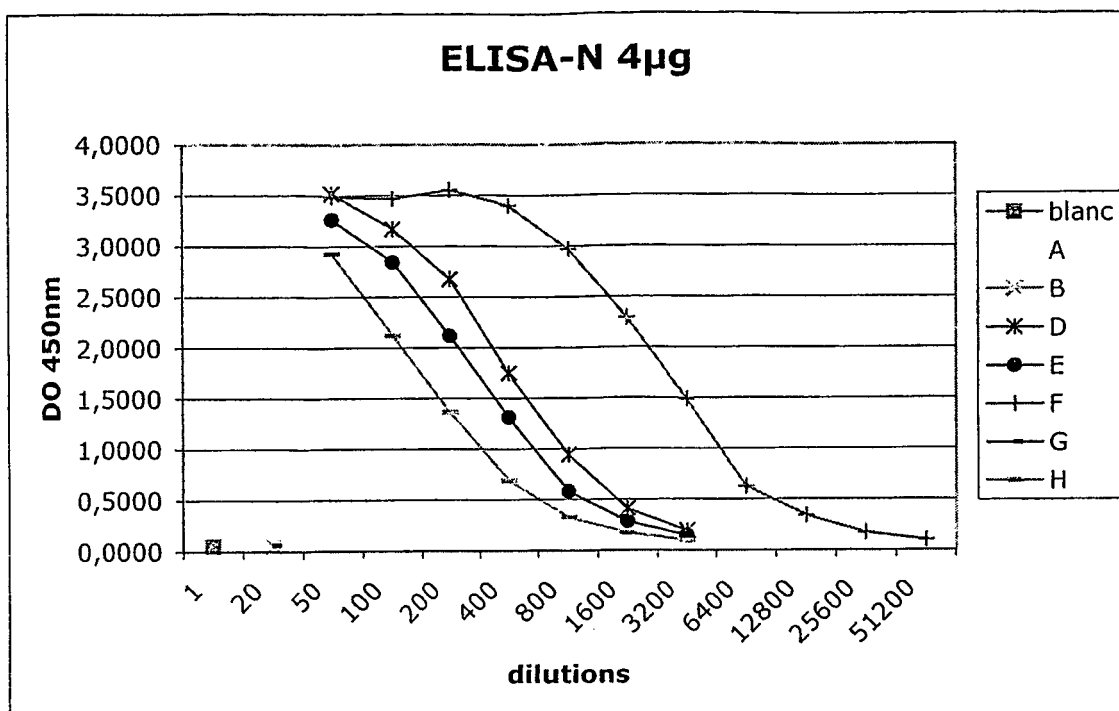


FIGURE 10a

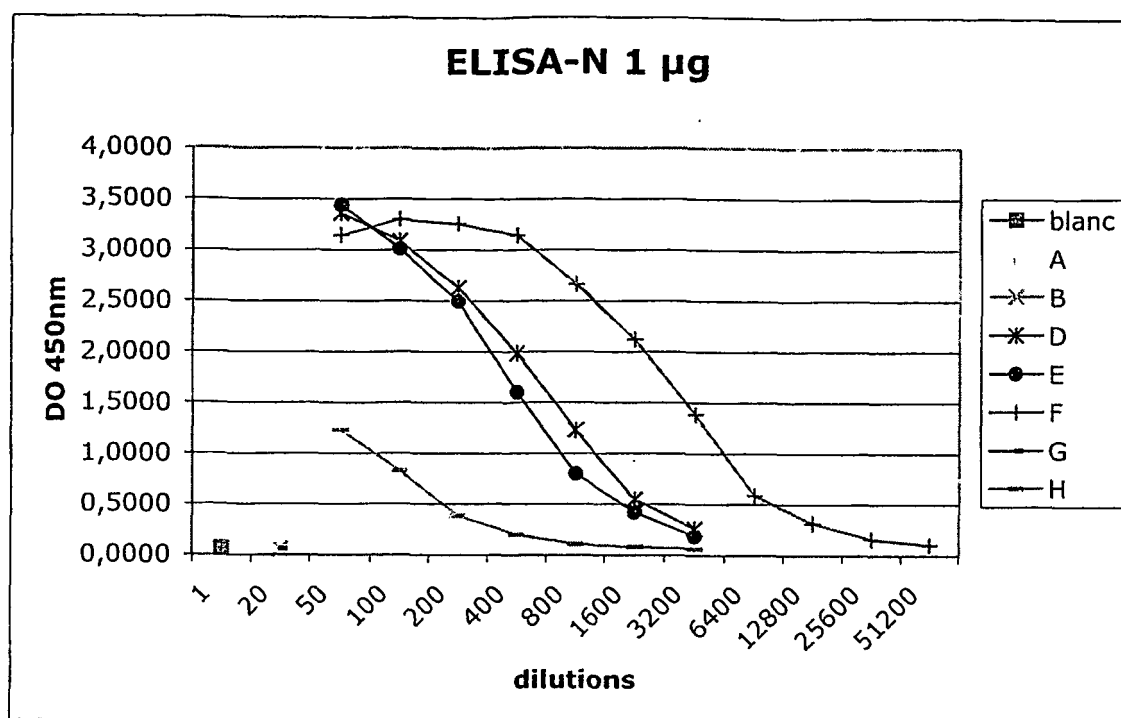


FIGURE 10b

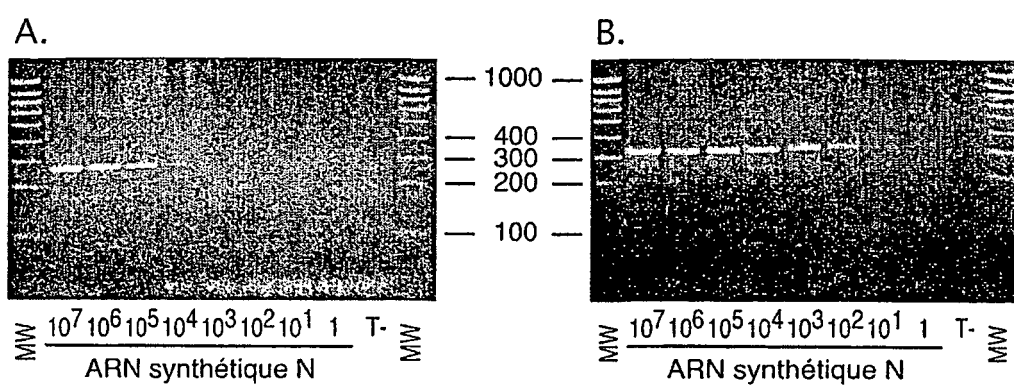


FIGURE 11

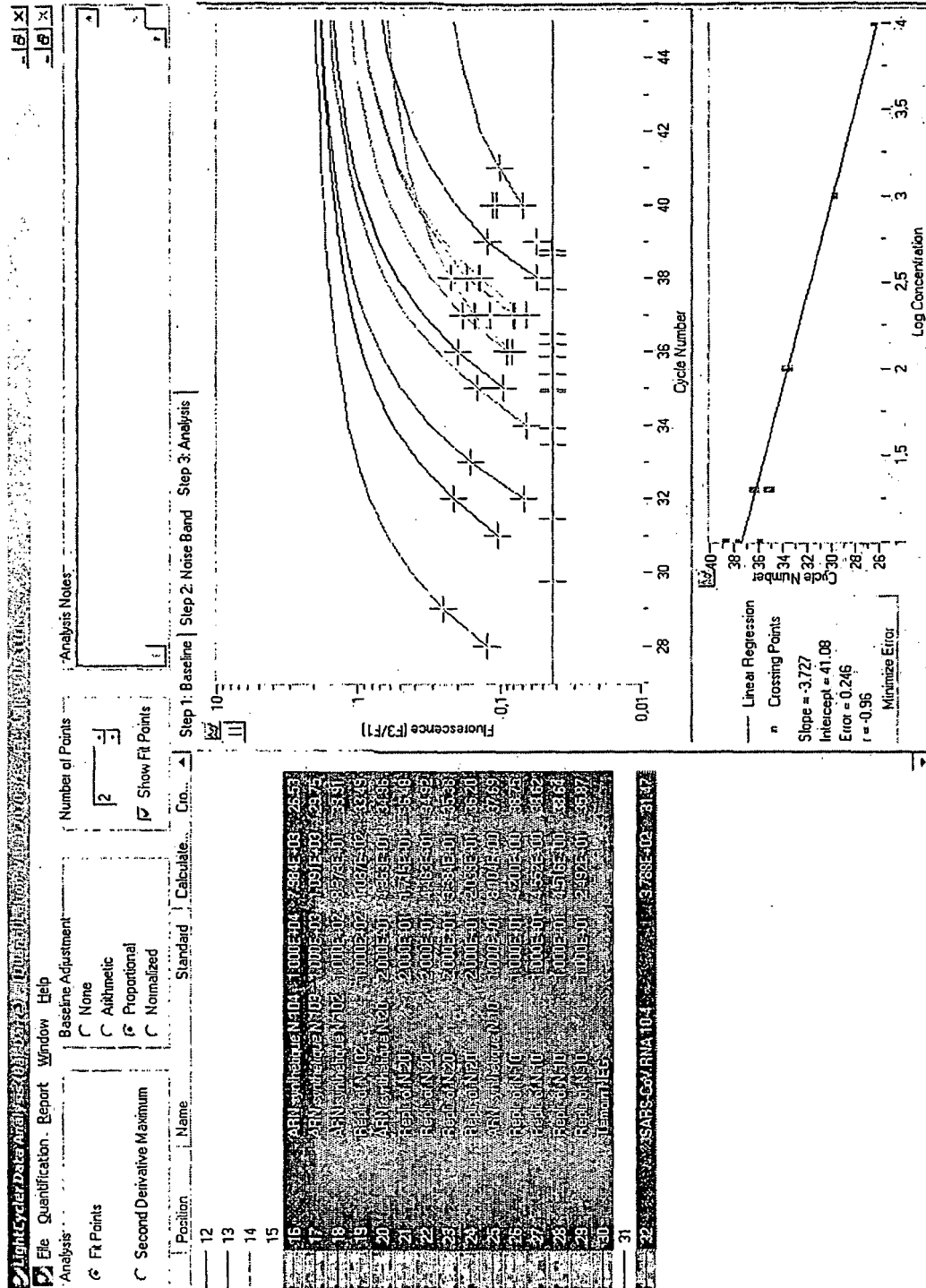


FIGURE 12

```

                                >< XhoII
                                >< Sau3AI
                                >< NdeII
>< MvaI                        > < TthHB8I
>< EcoRII                      > < TaqI      >< MflI
                                >< Sau3AI    >< MboI
>< Ecl136I                    >< NdeII    >< DpnII
>< DsaV                        >< MboI>< MnlI>< DpnI
                                >< DpnII    >< BstYI
                                >< DpnI     >< BspAI
>< BsaJI                      >< BspAI     >< Bsp143I
                                >< Bsp143I>< BglII
ATATTAGGTT TTTACCTACC CAGGAAAAGC CAACCAACCT CGATCTCTTG TAGATCTGTT CTCTAAACGA
   10         20         30         40         50         60         70

                                >< VneI
                                >< SphI
                                >< SnoI
                                >< RmaI
                                >< PaeI    >< SduI
                                >< NspI    >< NspII
                                >< NspHI   >< HgiAI
                                >< NlaIII  >< Bsp1286I
                                >< MaeI    >< BmyI
                                >< ApaLI
                                >< Alw44I
>< Tru9I
>< MseI        >< BbvI
>< DraI        >< AluI          > < Fnu4HI       >< Alw21I
ACTTTAAAT CTGTGTAGCT GTCGCTCGGC TGCATGCCTA GTGCACCTAC GCAGTATAAA CAATAATAAA
   80         90         100        110        120        130        140

                                >< SfcI
                                >< PstI
                                >< MnlI
                                >< Ksp632I
                                >< EarI
                                >< Eam1104I
TTTTACTGTC GTTGACAAGA AACGAGTAAC TCGTCCCTCT TCTGCAGACT GCTTACGGTT TCGTCCGTGT
   150        160        170        180        190        200        210

                                >< TthHB8I
                                >< StyI
>< TaqI                >< RmaI            >< ScrFI
>< Sau3AI              >< MaeI            >< NciI
>< NdeII               >< EcoT14I       >< MspI
>< MboI                >< Eco130I       >< MaeIII
>< DpnII               >< BssT1I        >< HpaII
>< DpnI                >< BsaJI          >< HapII
>< BspAI               >< BlnI           >< DsaV
>< Bsp143I             >< AvrII          >< BcnI
TGCAGTCGAT CATCAGCAT A CCTAGGTTTC GTCCGGGTGT GACCGAAAGG TAAGATGGAG AGCCTTGTTT
   220        230        240        250        260        270        280

                                >< RmaI
                                >< Esp3I  >< MaeII
                                >< BsmAI    >< MaeI
                                >< Alw26I  >< BsmBI
TTGGTGTCAA CGAGAAAACA CACGTCCAAC TCAGTTTGCC TGTCCTTCAG GTTAGAGACG TGCTAGTGCG
   290        300        310        320        330        340        350

```

FIGURE 13.1

# EP 1 694 829 B1

```

    >< Sau96I
      >< PssI
    >< Pali
    >< NspIV
      >< MnlI
      >< HaeIII
    >< EcoO109I
      >< DraII>< MboII >< PmlI
    >< MnlI >< Cfr13I >< PmaCI
    >< Ksp632I >< BsuRI >< MaeII
    >< HinfI >< BsiZI>< EcoNI >< Eco72I
      >< EarI >< BshI >< BslI >< BsaAI
    >< PleI >< Eam1104I>< AsuI >< BsiYI>< BbrPI >< MnlI
TGGCTTCGGG GACTCTGTGG AAGAGGCCCT ATCGGAGGCA CGTGAACACC TCAAAAATGG CACTTGTGGT
    360      370      380      390      400      410      420

                                >< Tru9I
                                >< SfaNI
    >< RmaI >< RsaI >< Csp6I >< BspWI >< MseI
    >< MaeI >< AluI >< AfaI >< AluI >< MaeII
CTAGTAGAGC TGGAAAAAGG CGTACTGCCC CAGCTTGAAC AGCCCTATGT GTTCATTAAA CGTTCTGATG
    430      440      450      460      470      480      490

    >< Pali
    >< HaeIII
    >< Tru9I >< GdiII >< RsaI
    >< MseI >< EaeI McrI ><
    >< Esp4I >< BsuRI >< BsmI BsiEI ><
    >< AflII >< BshI >< AluI >< BscCI >< AfaI
CCTTAAGCAC CAATCACGGC CACAAGGTCG TTGAGCTGGT TGCAGAAATG GACGGCATTC AGTACGGTCG
    500      510      520      530      540      550      560

                                >< NspI
                                >< NspHI
    >< ScaI >< RsaI >< NlaIII
    >< Csp6I >< BslI >< MboII
    >< BsrI >< BsiYI >< MboII
    >< AciI >< AfaI >< AflIII >< MunI >< AciI
TAGCGGTATA ACGGTGAG TACTCGTGCC ACATGTGGGC GAAACCCCAA TTGCATACCG CAATGTTCTT
    570      580      590      600      610      620      630

                                >< TthHB8I
                                >< TagI
                                >< Sau3AI
                                >< NdeII
                                >< MboI
                                >< DpnII
                                >< DpnI
                                >< ClaI
                                >< Bsu15I
                                >< BspDI
                                >< BspAI
                                >< Bsp143I
                                >< Bsp106I
                                >< BsiXI MaeIII >
                                >< BscI>< SfaNI DdeI ><
                                >< BanIII BfrI ><
CTTCGTAAGA ACGTAATAA GGGAGCCGGT GGTCATAGCT ATGGCATCGA TCTAAAGTCT TATGACTTAG
    640      650      660      670      680      690      700

```

FIGURE 13.2

```
>< Sau3AI  
>< NdeII  
>< MboI  
    >< HphI                                     VneI ><  
        >< DpnII                               SnoI ><  
            >< BspAI                           > < NlaIII  
                >< AlwI>< DpnI                 >< DdeI   ApaLI ><  
                    >< AluI       >< Bsp143I      >< MboII >< BsrI     Alw44I ><  
GTGACGAGCT TGGCACTGAT CCCATTGAAG ATTATGAACA AACTTGAAC ACTAAGCATG GCAGTGGTGCG  
710         720         730         740         750         760         770  
  
                >< SstI  
                >< SduI  
                >< SacI  
                >< NspII  
                >< MnlI                                       Sau96I ><  
                    >< HgiAI                                   >< TthHB8I   PalI ><  
>< SduI               >< Eco24I                             >< TaqI     NspIV ><  
>< NspII              >< Ecl136II                         > < SalI     HaeIII ><  
>< HgiAI             >< Bsp1286I                       > < RtrI     Cfr13I ><  
    >< DraIII          >< BmyI                            >< HindII   BsuRI ><  
>< Bsp1286I          >< BanII                          >< HincII   BsiZI ><  
>< BmyI              >< Alw21I                        >< BsgI     BshI ><  
>< Alw21I           >< AluI              >< MaeIII   >< AccI     AsuI ><  
ACTCCGTGAA CTCACCTCGTG AGCTCAATGG AGGTGCAGTC ACTCGCTATG TCGACAACAA TTTCTGTGGC  
780         790         800         810         820         830         840  
  
                                >< ThaI  
                                >< ThaI  
                                >< MvnI  
                                >< MvnI  
                                >< HinPII  
                                >< Hin6I           > < VneI  
                                >< HhaI           > < SnoI  
                                >< CfoI                   >< SduI  
                                >< BstUI                     NspII ><  
                                >< BstUI                     HgiAI ><  
                                >< Bsp50I   Bsp1286I ><  
                                >< Bsp50I                   >< BmyI  
                                >< AciI           > < ApaLI  
                                >< AccII           > < Alw44I  
>< Acc65I           >< MnII   >< SfaNI   >< AccII   Alw21I ><  
CCAGATGGGT ACCCTCTTGA TTGCATCAAA GATTTTCTCG CACGCGCGGG CAAGTCAATG TGCACTCTTT  
850         860         870         880         890         900         910  
  
                >< TthHB8I  
                >< TthHB8I  
                >< TaqI  
                >< TaqI  
                >< MnII  
                >< Ksp632I                                       NlaIII ><  
                >< HinfI>< PleI                                 >< NlaIII  
                >< Eam1104I   >< MboII   >< MaeIII   EcoRII ><  
                >< EarI   > < BbvI>< AccI >< Fnu4HI   DsaV ><  
CCGAACAAC TgATTACATc GAgTcGAAGA GAAGTGTCTA CTGCTGCCGT GACCAtGAGc ATGAAttGC  
920         930         940         950         960         970         980  
  
                                >< TthHB8I  
                                >< TaqI  
                                >< SfuI  
                                >< NspV>< Tru9I  
>< ScrFI           >< HinPII           >< LspI>< MseI
```

FIGURE 13.3

# EP 1 694 829 B1

```

>< MvaI      >< Hin6I      >< SduI      >< Csp45I
>< Ecl136I   >< HhaI      >< NspII     >< BstBI
>< BstOI     >< HaeII     >< HgiAI     >< Bsp119I
>< BstNI     >< Eco47III   >< Bsp1286I  >< BsiCI
>< BsiLI     >< CfoI      >< BmyI      >< Bpu14I
>< ApyI >< DdeI >< Bsp143II >< AluI    >< Alw21I    >< AsuII
CTGGTTCACCT GAGCGCTCTG ATAAGAGCTA CGAGCACCAG ACACCCTTCG AAATTAAGAG TGCCAAGAAA
    990         1000         1010         1020         1030         1040         1050

                                >< Tru9I
                                >< MseI
                                >< BsmI
                                >< BscCI
                                >< MnlI
TTTGACACTT TCAAAGGGGA ATGCCCAAAG TTTGTGTTTC CTCTTAACTC AAAAGTCAAA GTCATTCAAC
    1060         1070         1080         1090         1100         1110         1120

>< PmlI
>< PmaCI
>< MaeII
>< Eco72I
>< BsaAI
>< BbrPI
>< AflIII    >< MnlI>< DdeI
CACGTGTTGA AAAGAAAAAG ACTGAGGGTT TCATGGGGCG TATACGCTCT GTGTACCCTG TTGCATCTCC
    1130         1140         1150         1160         1170         1180         1190

>< SfaNI
>< MaeIII    >< AccI
ACAGGAGTGT AACAATATGC ACTTGCTCTAC CTTGATGAAA TGTAATCATT GCGATGAAGT TTCATGGCAG
    1200         1210         1220         1230         1240         1250         1260

                                >< SinI
                                >< Sau96I
                                PssI ><
                                >< Psp5II
                                >< PpuMI
                                >< NspIV
                                >< NspHII
                                >< Eco47I
                                >< DraII
                                >< Cfr13I
                                >< BsiZI
                                >< Bme18I
                                >< AvaII
                                >< AsuI
>< MaeII     EcoO109I ><AflIII >
ACGTGCGACT TTCTGAAAGC CACTTGTGAA CATTGTGGCA CTGAAAATTT AGTTATTGAA GGACCTACTA
    1270         1280         1290         1300         1310         1320         1330

                                Van91I ><
                                SinI ><
                                Sau96I ><
                                PflMI ><
                                NspIV ><
                                NspHII >
                                Eco47I ><
                                Cfr13I ><
                                BslI ><
                                BsiZI ><
                                BsiYI ><
                                Bme18I ><
                                AvaII ><
                                AsuI ><

>< RsaI
>< NspI
>< NlaIV
>< NlaIII
>< NspHI>< KpnI
>< Eco64I
>< Csp6I
>< BscBI
>< Bani
>< Asp718
>< AfaI
>< AccB1I

```

FIGURE 13. 4



```

    >< Acc65I          >< SfcI          >< NlaIII          AccB7I ><
CATGTGGGTA CCTACCTACT AATGCTGTAG TGAAAATGCC ATGTCTGCC TGTCAAGACC CAGAGATTGG
    1340          1350          1360          1370          1380          1390          1400

                                >< TthHB8I
                                >< TaqI>< MnlI
                                >< HinfI

    >< DdeI          >< PleI          >< AciI
ACCTGAGCAT AGTGTTCAG ATTATCACAA CCACTCAAAC ATTGAAACTC GACTCCGCAA GGGAGGTAGG
    1410          1420          1430          1440          1450          1460          1470

    >< RmaI          NlaIV ><
    >< MnlI          >< BsrI
    >< MaeI          >< BbvI          >< Fnu4HI          BscBI ><
ACTAGATGTT TTGGAGGCTG TGTGTTTGGC TATGTTGGCT GCTATAATAA GCGTGCCTAC TGGGTTCCTC
    1480          1490          1500          1510          1520          1530          1540

                                XhoII ><
                                Sau3AI ><
                                NdeII ><
                                MflI ><
                                MboI ><
                                >< MaeIII
                                >< Eco31I          DpnII ><
                                >< HaeIII          >< BsrI          >< MnlI DpnI >
    >< RmaI          >< BsuRI          >< BsrI          >< BsmAI          BstYI ><
    >< MnlI          >< DdeI          >< BspWI          >< BsaI>< HphI          BspAI ><
    >< MaeI          >< BshI>< BglI          >< Alw26I          Bsp143I >
GTGCTAGTGC TGATATTGGC TCAGGCCATA CTGGCATTAC TGGTGACAAT GTGGAGACCT TGAATGAGGA
    1550          1560          1570          1580          1590          1600          1610

                                >< Tru9I
                                >< MseI
                                >< MaeII          >< Tru9I
                                >< HpaI          >< MnlI
                                >< HindII          >< Ksp632I
                                >< HinfI >< PleI >< HincII          >< EarI
    >< AlwI          >< DdeI          >< AflIII          >< MseI          >< Eam1104I
TCTCCTTGAG ATACTGAGTC GTGAACGTGT TAACATTAAC ATTGTTGGCG ATTTTCATTT GAATGAAGAG
    1620          1630          1640          1650          1660          1670          1680

    >< MboII          PleI ><
    >< BstXI          >< SfaNI          >< HinfI
GTTGCCATCA TTTTGGCATC TTTCTCTGCT TCTACAAGTG CCTTTATTGA CACTATAAAG AGTCTTGATT
    1690          1700          1710          1720          1730          1740          1750

                                >< StyI
                                >< MaeIII
                                >< EcoT14I
                                >< PleI          >< Eco130I
                                >< MaeIII          >< BssT1I          BslI ><
                                >< HinfI>< AciI          >< BsaJI          BsiYI ><
ACAAGTCTTT CAAAACCATT GTTGAGTCCT GCGGTAAC TAAGATTACC AAGGGAAAGC CCGTAAAGG
    1760          1770          1780          1790          1800          1810          1820

    >< Sau3AI          >< Van91I
    >< NdeII          >< PflMI
    >< MboI          >< DraIII
    >< DpnII          >< BslI
    >< DpnI >< Tru9I          >< BsiYI
    >< BspAI >< MseI          >< BbvI          >< MnlI
    >< Bsp143I          >< AccB7I          Fnu4HI ><

```

FIGURE 135

EP 1 694 829 B1

```

TGCTTGAAC ATTGGACAAC AGAGATCAGT TTTAACACCA CTGTGTGGTT TTCCCTCACA GGCTGCTGGT
1830      1840      1850      1860      1870      1880      1890

      >< ThaI
      >< SfaNI
      >< MvnI
      >< HinPII
>< HinPII
      >< Hin6I
>< Hin6I
      >< HhaI
>< Sau3AI      >< HhaI
>< NdeII      >< CfoI
>< MboI      >< CfoI
>< DpnII      >< BstUI
      >< DpnI      >< BssHII
>< BspAI      >< Bsp50I
      >< Bsp143I      >< AccII
GTTATCAGAT CAATTTTTCG GCGCACACTT GATGCAGCAA ACCACTCAAT TCCTGATTTC CAAAGAGCAG
1900      1910      1920      1930      1940      1950      1960

      >< TthHB8I
      >< StyI
      >< NcoI
      >< HindII
      >< HincII
      >< HinfI
      >< EcoT14I
      >< Eco57I
      >< TaqI>< Eco130I
>< SalI >< DsaI
>< RtrI >< BssT1I
      >< BsaHI
      >< BbiII>< NlaIII
      >< AcyI >< HgaI
>< MaeIII
      >< BbvI
      >< MaeII >< AccI>< BsaJI
CTGTCACCAT ACTTGATGGT ATTTCTGAAC AGTCATTACG TCTTGTCGAC GCCATGGTTT ATACTTCAGA
1970      1980      1990      2000      2010      2020      2030

      >< RsaI
      >< NdeI      >< Csp6I
      >< BspMI
      >< MaeIII >< BsrI >< AfaI      >< DdeI
CCTGCTCACC AACAGTGTCA TTATTATGGC ATATGTAAC TGGTCTTTG TACAACAGAC TTCTCAGTGG
2040      2050      2060      2070      2080      2090      2100

      >< StuI
      >< Pali
      >< HaeIII
      >< Eco147I
      >< SduI      >< DdeI
      >< NspII      >< BsuRI
      >< Bsp1286I      >< BshI
      >< BmyI      >< AatI      >< MnlI
TTGTCTAATC TTTTGGGCAC TACTGTTGAA AAATCAGGC CTATCTTTGA ATGGATTGAG GCGAAACTTA
2110      2120      2130      2140      2150      2160      2170

      >< TfiI
      >< HinfI
      >< FokI
      >< SfaNI >< BsgI
      >< Tth111I ><
GTGCAGGAGT TGAATTTCTC AAGGATGCTT GGGAGATTCT CAAATTTCTC ATTACAGGTG TTTTTCACAT
2180      2190      2200      2210      2220      2230      2240

```

FIGURE 13.6

```

Tru9I ><
MseI ><
HpaI >
HindII >
HincII >
>< Eco57I
CGTCAAGGGT CAAATACAGG TTGCTTCAGA TAACATCAAG GATTGTGTAA AATGCTTCAT TGATGTTGTT
2250      2260      2270      2280      2290      2300      2310

>< Sau3AI
>< NdeII
>< MboI
> < MaeIII
>< FbaI
>< DpnII
>< DpnI
>< BspAI
>< Bsp143I
>< TthHB8I
>< TaqI
AACAAGGCAC TCGAAATGTG CATTGATCAA GTCACATATCG CTGGCGCAAA GTTGCGATCA CTCAACTTAG
2320      2330      2340      2350      2360      2370      2380

>< Sau3AI
>< NdeII
>< DpnII
>< DpnIMboII ><
>< DdeI ><
>< Bsp143I
>< HhaI
>< MboIBfrI ><
>< CfoI
>< BspAI
>< BbsI ><
>< PvuII
>< MaeII
>< Psp5I
>< Bst1107I
>< BsaAI
>< BbvI
>< Fnu4HI
>< Fnu4HI
>< HphI
>< DrdI
>< AccI
>< AluI
GTGAAGTCTT CATCGCTCAA AGCAAGGGAC TTTACCGTCA GTGTATACGT GGCAAGGAGC AGCTGCAACT
2390      2400      2410      2420      2430      2440      2450

>< Tru9I
>< NlaIV
>< MseI
>< MnlI
>< Esp4I
>< Eco64I
>< BscBI
>< NlaIII >< BanI
>< AflIII
>< BbvI
>< AccBII
>< MaeIII
>< TfiI
>< HinfI
>< HphI
>< AfaI
ACTCATGCCT CTTAAGGCAC CAAAAGAAGT AACCTTTCTT GAAGGTGATT CACATGACAC AGTACTTACC
2460      2470      2480      2490      2500      2510      2520

> < XhoI
>< TthHB8I
>< TthHB8I>< TaqI
> < SfaI
> < PaeR7I
> < NspIII
>< HphI >< HinfI
> < Eco88I
> < CcrI
>< Esp3I >< BsaHI
> < BcoI
>< BsmAI >< BbiII
> < AvaI >< HgaI
>< TaqI > < Ama87I>< BsmBI
>< Alw26I >< AcyI
>< AluI
TCTGAGGAGG TTGTTCTCAA GAACGGTGAA CTCGAAGCAC TCGAGACGCC CGTTGATAGC TTCACAAATG
2530      2540      2550      2560      2570      2580      2590

```

FIGURE 13.7

# EP 1 694 829 B1

```

                                >< Pali >< NlaIII
                                >< HaeIII >< MnlI
                                >< BsuRI >< DdeI >< Tru9I
                                >< BshI >< BfrI >< MseI
>< AluI >< BsrI >< BshI >< BfrI >< MseI
GAGCTATCGT TGGCACACCA GTCTGTGTAA ATGGCCTCAT GCTCTTAGAG ATTAAGGACA AAGAACAATA
2600 2610 2620 2630 2640 2650 2660

                                >< VneI
                                Tru9I ><
                                >< SnoI
                                >< SduI
                                >< NspII
                                MseI ><
                                >< HgiAI
                                Bsp1286I >< BslI ><
                                BsiYI ><
                                >< BmyI
                                >< ApaLI
                                >< Tru9I >< Alw44I
                                >< MseI >< Alw21I
CTGCGCATTG TCTCCTGGTT TACTGGCTAC AAACAATGTC TTTCGCTTAA AAGGGGGTGC ACCAATTAAA
2670 2680 2690 2700 2710 2720 2730

                                >< TfiI
>< MaeIII >< MboII >< MaeIII >< HinfI AluI ><
GGTGTAACCT TTGGAGAAGA TACTGTTTGG GAAGTTCAAG GTTACAAGAA TGTGAGAATC ACATTTGAGC
2740 2750 2760 2770 2780 2790 2800

                                >< RsaI
                                >< NlaIV
                                MaeIII ><
                                >< MspI>< KpnI
                                >< HpaII
                                >< HapII
                                >< Eco64I
                                >< SduI >< Csp6I
                                >< NspII >< TfiI >< BscBI
                                >< HgiAI >< BanI
                                >< Bsp1286I >< Asp718
                                >< BmyI >< HinfI >< AfaI
                                >< Alw21I >< AccB1I
                                >< AflIII >< MseI >< AccI >< Acc65I
TTGATGAACG TGTGACAAA GTGCTTAATG AAAAGTGCTC TGTCTACACT GTTGAATCCG GTACCGAAGT
2810 2820 2830 2840 2850 2860 2870

                                >< Sau3AI
                                >< NdeII
                                >< MboI
                                >< DpnII
                                >< DpnI
                                >< NspI
                                >< NspHI
                                >< MboII >< BspAI
                                >< NlaIII >< BsrI >< Bsp143I
>< DdeI >< MnlI >< AlwNI >< BbsI >< AlwNI
TACTGAGTTT GCATGTGTTG TAGCAGAGGC TGTGTGAAG ACTTTACAAC CAGTTTCTGA TCTCCTTACC
2880 2890 2900 2910 2920 2930 2940

                                >< Sau3AI
                                >< NdeII
                                >< MboI
                                >< DpnII
                                >< DpnI
                                >< BspAI

```

FIGURE 13.8

EP 1 694 829 B1

```

    >< NlaIII>< Bsp143I          >< AluI          >< SfaNI
AACATGGGTA TTGATCTTGA TGAGTGGAGT GTAGCTACAT TCTACTTATT TGATGATGCT GGTGAAGAAA
    2950          2960          2970          2980          2990          3000          3010

                                >< SfaNI
                                >< MnlI
                                >< Ksp632I          >< MnlI
                                >< EarI          >< MboII
    >< MboII          >< GsuI          >< MnlI          >< Eam1104I          >< MboII
    >< BsaAI          >< HphI          >< MaeII>< BpmI          >< MnlI          >< Eam1104I          >< MboII
ACTTTTCATC ACGTATGTAT TGTTCCTTTT ACCCTCCAGA TGAGGAAGAA GAGGACGATG CAGAGTGTGA
    3020          3030          3040          3050          3060          3070          3080

                                >< RsaI
                                >< RsaI
                                >< NlaIII
                                >< MnlI          >< FokI
                                >< Csp6I          Eco31I ><
                                >< Csp6I          >< MamI BsmAI ><
    >< MboII          >< MboII          >< AfaI          >< BsiBI BsaI ><
    >< MboII          >< AfaI          >< BsaB1Alw26I ><
GGAAGAAGAA ATTGATGAAA CCTGTGAACA TGAGTACGGT ACAGAGGATG ATTATCAAGG TCTCCCTCTG
    3090          3100          3110          3120          3130          3140          3150

    >< NlaIV>< PvuII>< XmnI
    >< Eco64I >< Psp5I          >< TthHB8I
    >< MnlI >< DdeI          >< TaqI          >< MnlI          >< MboII
    >< BscBI>< NspBII >< MnlI          >< Ksp632I          >< MboII >< MboII
    >< BanI          >< MnlI          >< EarI          >< BsrI
    >< AccBII >< AluI >< Asp700I          >< Eam1104I >< MboII>< BbsI
GAATTTGGTG CCTCAGCTGA AACAGTTCGA GTTGAGGAAG AAGAAGAGGA AGACTGGCTG GATGATACTA
    3160          3170          3180          3190          3200          3210          3220

                                >< Tru9I
    >< FokI          >< MseI          >< Eco57I
    >< DdeI          >< BsrI>< MboII BsrI ><
CTGAGCAATC AGAGATTGAG CCAGAACCAG AACCTACACC TGAAGAACCA GTTAATCAGT TTACTGGTTA
    3230          3240          3250          3260          3270          3280          3290

    >< Tru9I          >< MnlI
    >< MseI          >< Tru9I >< HindII>< Tru9I          >< DraIII
    >< DraI          >< MseI          >< HincII>< MseI          >< BspWI
TTTAAAACTT ACTGACAATG TTGCCATTAA ATGTGTTGAC ATCGTTAAGG AGGCACAAAG TGCTAATCCT
    3300          3310          3320          3330          3340          3350          3360

                                >< VneI
                                >< SnoI
                                >< SduI
                                >< NspII
                                >< HgiAI
                                >< Bsp1286I
                                >< BmyI
                                >< ApaLI
                                >< Alw44I
    >< BbvI          >< HphI          >< NlaIII          >< BspMI          >< Alw21I
    >< Fnu4HI          >< BspMI          >< Alw21I
ATGGTGATTG TAAATGCTGC TAACATACAC CTGAAACATG GTGGTGGTGT AGCAGGTGCA CTCAACAAGG
    3370          3380          3390          3400          3410          3420          3430

                                >< Sau96I
                                >< Pali
                                >< NspIV
                                >< HaeIII
    >< NlaIV          >< Cfr13I

```

FIGURE 13.9

EP 1 694 829 B1

```

>< Eco64I                                >< BsuRI
>< BscBI                                > < Tru9I    >< BsiZI
>< BanI                                > < MseI      >< BshI      >< MnlI
>< AccB1I>< NlaIII                      >< AluI    >< AsuI    >< MnlI
CAACCAATGG TGCCATGCAA AAGGAGAGTG ATGATTACAT TAAGCTAAAT GGCCCTCTTA CAGTAGGAGG
3440          3450          3460          3470          3480          3490          3500

                                >< SinI
                                >< Sau96I
                                >< NspIV
                                >< NspHI>< NspHII
                                >< Eco47I
                                >< Cfr13I
                                >< NlaIII    >< BspMI
                                >< BsiZI
                                >< Bme18I
                                >< AvaII MnlI ><
                                >< NspI>< AsuI FokI ><
GTCTTGTTTG CTTTCTGGAC ATAATCTTGC TAAGAAGTGT CTGCATGTTG TTGGACCTAA CCTAAATGCA
3510          3520          3530          3540          3550          3560          3570

                                > < Tru9I
                                >< HphI> < MseI
                                >< Esp4I
                                >< AluI      > < NdeI
                                >< AflIII>< Fnu4HI    >< BbvI
GGTGAGGACA TCCAGCTTCT TAAGGCAGCA TATGAAAATT TCAATTCACA GGACATCTTA CTTGCACCAT
3580          3590          3600          3610          3620          3630          3640

                                RsaI ><
                                Csp6I ><
                                AfaI ><
                                >< Eco57I                      >< BcgI
TGTGTCAGC AGGCATATTT GGTGCTAAAC CACTTCAGTC TTTACAAGTG TGCGTGCAGA CGGTTCGTAC
3650          3660          3670          3680          3690          3700          3710

>< BsgI                                >< BspMI
>< BcgI/a                                >< AluI      >< NlaIII
ACAGGTTTAT ATTGCAGTCA ATGACAAAGC TCTTTATGAG CAGGTTGTCA TGGATTATCT TGATAACCTG
3720          3730          3740          3750          3760          3770          3780

                                >< MnlI
>< RmaI                                > < MnlI    >< NlaIV    >< TfiI    >< MboII
>< MaeI                                >< Eco57I    >< BscBI    >< HinfI    >< DdeI
AAGCCTAGAG TGGAAGCACC TAAACAAGAG GAGCCACCAA ACACAGAAGA TTCCAAAAC T GAGGAGAAAT
3790          3800          3810          3820          3830          3840          3850

                                >< Tru9I
                                >< StuI
                                >< Pali
                                >< MseI    >< MnlI    >< MaeIII
                                >< HaeIII    >< EcoO651
                                >< Eco147I    >< Eco91I
                                >< BsuRI      BstXI ><
                                >< BshI      >< BstPI
                                >< AatI      >< BstEII
CTGTCGTACA GAAGCCTGTC GATGTGAAGC CAAAAATTAA GGCCTGCATT GATGAGGTTA CCACAACACT
3860          3870          3880          3890          3900          3910          3920

                                TfiI ><
                                NlaIII ><
                                HinfI ><
>< DdeI                                >< EcoRV    >< HindIII

```

FIGURE 13.10

# EP 1 694 829 B1

```

>< BsrI      >< MboII      >< MaeIII      >< Eco32I      >< AluI
GGAAGAACT  AAGTTTCTTA  CCAATAAGTT  ACTCTTGTTT  GCTGATATCA  ATGGTAAGCT  TTACCATGAT
3930      3940      3950      3960      3970      3980      3990

      >< NspI
      >< NspHI
      >< NlaIII
      >< MnlI      >< SfaNI
      >< DdeI      >< EcoNI
      >< MboII >< BslI      >< NlaIII
>< DdeI      >< BfrI      >< HphI      >< BsiYI      >< FokI
TCTCAGAACA  TGCTTAGAGG  TGAAGATATG  TCTTTCCTTG  AGAAGGATGC  ACCTTACATG  GTAGGTGATG
4000      4010      4020      4030      4040      4050      4060

      >< SpeI
      >< RmaI
      >< MaeI      >< EcoRV>< HphI      >< SfaNI
      >< HphI      >< Eco32I      >< MnlI      >< DdeI
TTATCACTAG  TGGTGATATC  ACTTGTTGTTG  TAATACCCTC  CAAAAGGCT  GGTGGCACTA  CTGAGATGCT
4070      4080      4090      4100      4110      4120      4130

      >< ScrFI
      >< RsaI
      >< MvaI
      >< EcoRII
      >< Ecl136I
      >< DsaV
      >< Csp6I >< EcoNI
      >< BstOI
      >< BstNI
      >< BsiLI
      >< BsaJI
      >< BsaAI      >< BslI
      >< MaeII>< ApyI
      >< AfaI      >< BsiYI
      >< AluI      >< BsrI      >< MboII      >< MaeII>< ApyI
CTCAAGAGCT  TTGAAGAAAG  TGCCAGTTGA  TGAGTATATA  ACCACGTACC  CTGGACAAGG  ATGTGCTGGT
4140      4150      4160      4170      4180      4190      4200

      >< Tru9I
      >< MseI
      >< DdeI      >< Esp4I
      >< MnlI      >< BspWI
      >< FokI      >< AluI      >< AflII      >< Eco57I >< AfaI
TATACACTTG  AGGAAGCTAA  GACTGCTCTT  AAGAAATGCA  AATCTGCATT  TTATGTACTA  CCTTCAGAAG
4210      4220      4230      4240      4250      4260      4270

      >< ScrFI
      >< MvaI
      >< EcoRII
      >< XmnI      >< Ecl136I      NlaIII ><
      >< Ksp632I      >< RmaI      >< DsaV      Ksp632I ><
      >< EarI      >< TfiI>< MboII      >< BstOI      >< EarI
      >< Eam1104I      >< MaeI      >< BstNI      Eam1104I ><
      >< DdeI      >< HinfI      >< BsiLI      BsmAI ><
      >< BspWI      >< Asp700I      >< ApyI      Alw26I ><
CACCTAATGC  TAAGGAAGAG  ATTCTAGGAA  CTGTATCCTG  GAATTTGAGA  GAAATGCTTG  CTCATGCTGA
4280      4290      4300      4310      4320      4330      4340

      >< VspI      >< Zsp2I
      >< Tru9I      >< Ppu10I
      >< MseI      >< NsiI
      >< MboII      >< NlaIII      >< FokI
      >< Eco57I      >< Mph1103I      >< FokI

```

FIGURE 13. 11

# EP 1 694 829 B1

```

                >< AsnI          >< EcoT22I          >< BspWI
                >< AseI          >< AvaIII          >< BglI          >< MaeII
AGAGACAAGA AAATTAATGC CTATATGCAT GGATGTTAGA GCCATAATGG CAACCATCCA ACGTAAGTAT
4350      4360      4370      4380      4390      4400      4410

                >< SfaNI
                >< Tru9I          > < HindII          >< TfiI          >< SpeI
                >< MseI          > < HincII>< MboII          >< RmaI
                >< MnlI          >< DrdI >< HinfI          >< MaeI
AAAGGAATTA AAATTCAAGA GGGCATCGTT GACTATGGTG TCCGATTCTT CTTTATACT AGTAAAGAGC
4420      4430      4440      4450      4460      4470      4480

                >< MaeIII
>< SfcI          >< Fnu4HI          >< MunI
                >< AluI          >< AluI          >< AciI          MaeIII ><
CTGTAGCTTC TATTATTACG AAGCTGAACT CTCTAAATGA GCCGCTTGTC ACAATGCCAA TTGGTTATGT
4490      4500      4510      4520      4530      4540      4550

                >< ThaI
                >< MvnI
                >< MboII
                >< HinfII
                >< HinfII
                >< HinfI
                >< HinfI
                >< HhaI
                >< HhaI
                >< Tru9I          >< HhaI
>< NlaIII          >< Fnu4HI
                >< MseI          >< CfoI
                >< MnlI          >< CfoI
                >< Ksp632I          >< BstUI
                >< EarI          >< BssHII>< BspWI          >< Tru9I
                >< Eam1104I          >< Bsp50I          >< MseI
                >< BbvI          >< AccII          >< AluI          HphI ><
GACACATGGT TTTAATCTTG AAGAGGCTGC GCGCTGTATG CGTTCTCTTA AAGCTCCTGC CGTAGTGTCA
4560      4570      4580      4590      4600      4610      4620

                >< MaeIII
>< SfaNI          >< AlwNI          >< MnlI >< MnlI>< DdeI
GTATCATCAC CAGATGCTGT TACTACATAT AATGGATACC TCACTTCGTC ATCAAAGACA TCTGAGGAGC
4630      4640      4650      4660      4670      4680      4690

                >< SinI
                >< Sau96I
                >< NspIV
                >< NspHII
>< SduI          >< Eco47I
>< NspII          >< Cfr13I
>< HgiAI          >< BsiZI
>< Bsp1286I          >< Bmel8I          >< RsaI
>< BmyI          >< AvaII          >< Csp6I
>< Alw21I          >< AsuI          >< AfaI
ACTTTGTAGA AACAGTTTCT TTGGCTGGCT CTTACAGAGA TTGGTCCTAT TCAGGACAGC GTACAGAGTT
4700      4710      4720      4730      4740      4750      4760

                > < TthHB8I
                > < TaqI
                >< SduI
                >< Van91I >< NspII
                >< Tru9I          >< RsaI >< PflMI >< Eco24I
                >< MseI          >< HphI >< BslI >< Bsp1286I
                >< Esp4I          >< Csp6I >< BsiYI >< BmyI GsuI ><

```

FIGURE 13.12



# EP 1 694 829 B1

```

    >< AflIII >< MaeIII >< AfaI >< AccB7I >< BanIIBpmI ><
AGGTGTTGAA TTTCTTAAGC GTGGTGACAA AATTGTGTAC CACACTCTGG AGAGCCCCGT CGAGTTTCAT
    4770      4780      4790      4800      4810      4820      4830

                                >< Tru9I
                                >< PleI >< EcoNI
                                >< MnlI >< BslI
                                >< BsmAI >< BsiYI
    >< MnlI >< HphI >< HinfI>< Alw26I>< AciI >< MseI
CTTGACGGTG AGGTCTTTTC ACTTGACAAA CTAAAGAGTC TCTTATCCCT GCGGGAGGTT AAGACTATAA
    4840      4850      4860      4870      4880      4890      4900

                                >< AluI >< NdeI
AAGTGTTTAC AACTGTGGAC AACACTAATC TCCACACACA GCTTGTGGAT ATGTCTATGA CATATGGACA
    4910      4920      4930      4940      4950      4960      4970

    >< SinI
    >< Sau96I
    >< NspIV
    >< NspHII
    >< Eco47I
    >< Cfr13I
    >< BsiZI
    >< Bmel8I
    >< AvaII
    >< AsuI
                                >< MaeIII >< Tru9I >< MnlI
                                >< FokI >< MseI >< BspHI
GCAGTTTGGT CCAACATACT TGGATGGTGC TGATGTTACA AAAATTAAAC CTCATGTAAA TCATGAGGGT
    4980      4990      5000      5010      5020      5030      5040

                                >< TthHB8I
                                >< TaqI
    >< RsaI
    >< RmaI
    >< MaeI
    >< Csp6I
    >< AfaI
                                >< SnaBI
                                >< MaeII >< HindIII >< RsaI
                                >< Eco105I >< Csp6I
                                >< BsaAI >< AluI >< AfaI
AAGACTTTCT TTGTACTACC TAGTGATGAC ACACTACGTA GTGAAGCTTT CGAGTACTAC CATACTCTTG
    5050      5060      5070      5080      5090      5100      5110

    >< RsaI
    >< NspI
    >< NspHI
    >< NlaIII
    >< Csp6I >< Tru9I
    >< AflIII >< MseI
    >< AfaI >< DraI
                                MnlI >
                                BslI ><
                                BsiYI ><
ATGAGAGTTT TCTTGGTAGG TACATGTCTG CTTTAAACCA CACAAAGAAA TGGAAATTTT CTCAAGTTGG
    5120      5130      5140      5150      5160      5170      5180

    >< Tru9I >< Tru9I
    >< MseI >< MseI
                                >< MunI >< MaeI
                                AluI >
TGTTTAACT TCAATTAAAT GGGCTGATAA CAATTGTTAT TTGTCTAGTG TTTTATTAGC ACTTCAACAG
    5190      5200      5210      5220      5230      5240      5250

                                >< SfaNI
                                >< SduI
                                >< NspII
                                >< Eco24I
                                >< Bsp1286I
                                >< BmyI >< HphI >
                                >< BbvI Fnu4HI ><
                                >< BanII >< BspWI
    >< MnlI

```

FIGURE 13.13

# EP 1 694 829 B1

```

CTTGAAGTCA AATTCAATGC ACCAGCACTT CAAGAGGCTT ATTATAGAGC CCGTGCTGGT GATGCTGCTA
5260          5270          5280          5290          5300          5310          5320

>< VneI
>< SnaI
    >< SduI
    >< NspII
    >< HgiAI
    >< Bsp1286I
    >< BmyI
>< ApaLI
>< Alw44I
    >< Alw21I
    >< AluI
    MboII ><
    >< HphI
ACTTTTGTGC ACTCATACTC GCTTACAGTA ATAAACTGT TGGCGAGCTT GGTGATGTCA GAGAACTAT
5330          5340          5350          5360          5370          5380          5390

    > < SphI
    > < PaeI
    > < NspI
    > < NspHI >< TfiI
    >< Tru9I
    >< SfcI > < NlaIII>< HinfI
    >< MseI
GACCCATCTT CTACAGCATG CTAATTTGGA ATCTGCAAAG CGAGTTCTTA ATGTGGTGTG TAAACATTGT
5400          5410          5420          5430          5440          5450          5460

    >< Tru9I
    >< MseI
    >< AluI
    > < RsaI
    > < Csp6I
    Esp4I >
    >< AfaI
    AflII >
GGTCAGAAAA CTACTACCTT AACGGGTGTA GAAGCTGTGA TGTATATGGG TACTCTATCT TATGATAATC
5470          5480          5490          5500          5510          5520          5530

    >< RsaI
    >< MboII
    >< RmaI HinfI ><
    >< Csp6I
    >< MaeI >< BbsI
    >< AfaI
>< Tru9I
    >< SfaNI
>< MseI
    >< NlaIII
    >< AfaI
TTAAGACAGG TGTTTCCATT CCATGTGTGT GTGGTCGTGA TGCTACACAA TATCTAGTAC AACAAGAGTC
5540          5550          5560          5570          5580          5590          5600

    >< RsaI
    >< P1eI
    >< DdeI
    >< Csp6I
    >< BsgI
    >< BspWI >< BspMI
    >< AfaI
TTCTTTTGTG ATGATGTCTG CACCACCTGC TGAGTATAAA TTACAGCAAG GTACATTCTT ATGTGCGAAT
5610          5620          5630          5640          5650          5660          5670

    >< RsaI
    >< Eco31I
    >< DdeI
    > < MaeIII
    >< BsmAI
    >< Csp6I
    >< BsaI
    MnlI ><
    >< AfaI >< BsrI
    >< Alw26I
    HphI >
GAGTACACTG GTAACATATCA GTGTGGTCAT TACACTCATA TAACTGCTAA GGAGACCCTC TATCGTATTG
5680          5690          5700          5710          5720          5730          5740

    >< SstI
    >< SinI
    >< SduI
    >< Sau96I
    >< SacI
    >< NspIV
    >< NspII
    >< NspHII
    >< HgiAI
    > < RsaI
    >< MaeIII
    >< Eco24I
    >< Eco47I
    >< Ecl136II
    >< Cfr13I
    >< Bsp1286I
    >< BsiZI
    >< BmyI
    >< Bmel8I

```

FIGURE 13. 14

# EP 1 694 829 B1

```

    >< BanII
    >< Alw21I
    >< AluI
ACGGAGCTCA CCTTACAAAG ATGTCAGAGT ACAAAGGACC AGTGACTGAT GTTTTCTACA AGGAAACATC
    5750      5760      5770      5780      5790      5800      5810

    >< TthHB8I
    >< TaqI >< MaeIII
TTACTACTACA ACCATCAAGC CTGTGTCGTA TAAACTCGAT GGAGTTACTT ACACAGAGAT TGAACCAAAA
    5820      5830      5840      5850      5860      5870      5880

    >< RsaI
    >< Csp6I
    >< SfcI >< BbvI
    >< Fnu4HI
    >< AfaI
TTGGATGGGT ATTATAAAAA GGATAATGCT TACTATACAG AGCAGCCTAT AGACCTTGTA CCAACTCAAC
    5890      5900      5910      5920      5930      5940      5950

    Tru9I ><
    SwaI ><
    MseI ><
    > < NspI
    > < NspHI
    > < NlaIII
    >< AflIII
CATTACCAAA TGCGAGTTT GATAATTTCA AACTCACATG TTCTAACACA AAATTTGCTG ATGATTTAAA
    5960      5970      5980      5990      6000      6010      6020

    >< MboII
    >< AluI >< AluI >< MaeIII
TCAAATGACA GGCTTCACAA AGCCAGCTTC ACGAGAGCTA TCTGTCACAT TCTTCCCAGA CTTGAATGGC
    6030      6040      6050      6060      6070      6080      6090

    >< SfcI
GATGTAGTGG CTATTGACTA TAGACACTAT TCAGCGAGTT TCAAGAAAGG TGCTAAATTA CTGCATAAGC
    6100      6110      6120      6130      6140      6150      6160

    >< Tru9I
    >< ScrFI
    >< MvaI
    >< MseI
    >< EcoRII
    >< Ecl136I
    >< DsaV
    >< BstOI
    >< BstNI
    >< BsiLI
    >< MunI
    >< BstXI
    >< ApyI
    >< MaeII
    >< MaeII ><
    >< DraIII
    >< BstXI
CAATTGTTTG GCACATTAAC CAGGCTACAA CCAAGACAAC GTTCAAACCA AACACTTGGT GTTTACGTTG
    6170      6180      6190      6200      6210      6220      6230

    > < RsaI
    >< Csp6I
    > < AfaI >< BsrI
    >< MboII ><
    >< BbsI
TCTTTGGAGT ACAAAGCCAG TAGATACTTC AAATTCATTT GAAGTTCTGG CAGTAGAAGA CACACAAGGA
    6240      6250      6260      6270      6280      6290      6300

    >< HindII
    >< HincII
    >< MnlI
    >< MboII
    >< Eco57I
ATGGACAATC TTGCTTGTGA AAGTCAACAA CCCACCTCTG AAGAAGTAGT GGAAAATCCT ACCATACAGA
    6310      6320      6330      6340      6350      6360      6370

```

FIGURE 13.15

# EP 1 694 829 B1

```

                >< MaeIII                >< Tru9I
                >< MaeII                >< MseI
AGGAAGTCAT AGAGTGTGAC GTGAAACTA CCGAAGTTGT AGGCAATGTC ATACTTAAAC CATCAGATGA
    6380          6390          6400          6410          6420          6430          6440

                >< XhoII
                >< Sau3AI
                >< NlaIII
                >< NdeII
                >< MflI
                >< MboI
                >< DpnII
                >< DpnI
                >< BstYI
                >< BspAI
    >< Tru9I
    >< MseI                >< BspHI >< Bsp143I>< Fnu4HI
    > < MaeIII            >< MnlI >< BbvI            >< AlwI
AGGTGTTAAA GTAACACAAG AGTTAGGTCA TGAGGATCTT ATGGCTGCTT ATGTGGAAAA CACAAGCATT
    6450          6460          6470          6480          6490          6500          6510

                >< SauI
                >< RmaI
                >< MstII
                >< MaeI
                >< Eco81I
                >< DdeI
                >< CvnI
                >< Bsu36I
                >< Bse21I
                >< BfrI> < Tru9I
    >< Tru9I                >< AxyI> < MseI>< MunI                >< NlaIII
    >< MseI                >< AluI                >< AocI >< DraI                >< BbvI Fnu4HI ><
ACCATTAAGA AACCTAATGA GCTTTCAC TA GCTTAGGTT TAAAAACAAT TGCCACTCAT GGTATTGCTG
    6520          6530          6540          6550          6560          6570          6580

                >< VspI                >< StyI
                >< Tru9I                >< EcoT14I                > < DdeI
                >< MseI                >< Eco130I                >< BslI
                >< AsnI                >< BssT1I                >< BsiYI
                >< AseI                >< BsaJI                > < BfrI                >< Fnu4HI
CAATTAATAG TGTTCCTTGG AGTAAAATTT TGGCTTATGT CAAACCATT TTAGGACAAG CAGCAATTAC
    6590          6600          6610          6620          6630          6640          6650

                >< HinPII
                >< Hin6I                >< Tru9I
                >< HhaI                >< MaeII>< MseI
                >< DdeI                >< DraIII
    >< BbvI                >< CfoI                >< AflIII
AACATCAAAT TGCCTAAGA GATTAGCACA ACGTGTGTTT AACAATTATA TGCCTTATGT GTTTACATTA
    6660          6670          6680          6690          6700          6710          6720

                >< RsaI                > < RsaI>< XbaI
                >< Csp6I                >< Csp6I >< RmaI
    >< MunI >< AfaI                > < AfaI >< MaeI                >< AluI
TTGTTCCAAT TGTGTACTTT TACTAAAAGT ACCAATTCTA GAATTAGAGC TTCCTACCT ACAACTATTG
    6730          6740          6750          6760          6770          6780          6790

                >< VspI
                >< Tru9I
                >< NaeI
                >< MspI
                >< MseI

```

FIGURE 13. 16

```

                                >< HpaII
                                >< HapII
                                >< CfrI0I >< FokI
                                >< AsnI
                                >< AseI>< HphI>< MaeIII
                                >< Tru9I
                                >< DdeI
                                MaeIII >
                                >< MseI
                                >< BfrI
                                >< BbvI
CTAAAAATAG TGTTAAGAGT GTTGCTAAAT TATGTTTGA TGCCGGCATT AATTATGTGA AGTCACCCAA
6800 6810 6820 6830 6840 6850 6860

                                >< Tru9I
                                >< DdeI
                                MaeIII >
                                >< MseI
                                >< BfrI
                                >< BbvI
ATTTTCTAAA TTGTTACAA TCGCTATGTG GCTATTGTTG TTAAGTATTT GCTTAGGTTT TCTAATCTGT
6870 6880 6890 6900 6910 6920 6930

                                >< SduI
                                >< NspII
                                >< HgiAI
                                >< Bsp1286I
                                >< BmyI
                                >< Alw21I
                                >< RsaI
                                >< Csp6I
                                >< Fnu4HI
                                >< AfaI
                                >< SduI
                                >< NspII
                                >< HgiAI
                                >< Bsp1286I
                                >< BmyI
                                >< Alw21I
GTAAGTCTG CTTTGGTGT ACTCTTATCT AATTTTGGTG CTCCTTCTTA TTGTAATGGC GTTAGAGAAT
6940 6950 6960 6970 6980 6990 7000

                                >< Tru9I ><
                                >< MseI ><
                                >< Fnu4HI
                                >< BbvI >
                                >< Tru9I
                                >< MaeIII
                                >< Fnu4HI
                                >< MseI
                                >< MaeII
                                >< BbvI >
TGTATCTTAA TTCGTCTAAC GTTACTACTA TGGATTTCTG TGAAGGTTCT TTTCTTGCA GCATTTGTTT
7010 7020 7030 7040 7050 7060 7070

                                >< TfiI
                                >< MamI
                                >< HinfI
                                >< BsiBI
                                >< XmnI>< MaeIII
                                >< AluI >
                                >< PleI>< HinfI
                                >< BsaBI >< AluI
                                >< Asp700I
                                >< AfaI ><
AAGTGGATTA GACTCCCTTG ATTCTTATCC AGCTCTTGAA ACCATTGAGG TGACGATTTC ATCGTACAAG
7080 7090 7100 7110 7120 7130 7140

                                >< Pali
                                >< NspBII
                                >< HaeIII
                                >< GdiII
                                >< Fnu4HI
                                >< EaeI
                                >< DdeI
                                >< BsuRI
                                >< BshI >< BslI
                                >< RmaI
                                >< MaeI
                                >< AciI>< BsiYI
CTAGACTTGA CAATTTTAGG TCTGGCCGCT GAGTGGGTTT TGGCATATAT GTTGTTTACA AAATTCTTTT
7150 7160 7170 7180 7190 7200 7210

                                >< BspMI
                                >< RmaI
                                >< AluI
                                >< MaeI
ATTTATTAGG TCTTTCAGCT ATAATGCAGG TGTTCTTTGG CTATTTTGCT AGTCATTTC TCAGCAATTC
7220 7230 7240 7250 7260 7270 7280

                                >< RsaI ><
                                >< MboII
                                >< NlaIV
                                >< Eco64I
                                >< RsaI >< BscBI
                                >< Csp6I >< BanI
                                >< BsaBI ><
                                >< NlaIII
                                >< AfaI>< AccBI
                                >< AfaI ><

```

FIGURE 13.17

EP 1 694 829 B1

```

TTGGCTCATG TGGTTTATCA TTAGTATTGT ACAAATGGCA CCCGTTTCTG CAATGGTTAG GATGTACATC
  7290      7300      7310      7320      7330      7340      7350

                                TthHB8I ><
                                >< TaqI
                                MnlI ><
                                Ksp632I ><
                                >< FokI
                                >< EarI
                                >< MboII EarI ><
>< FokI
>< Eam1104I>< AluI>< MboII >< NlaIII Eam1104I ><
TTCTTTGCTT CTTTCTACTA CATATGGAAG AGCTATGTTT ATATCATGGA TGGTTGCACC TCTTCGACTT
  7360      7370      7380      7390      7400      7410      7420

                                XhoII ><
                                Sau3AI ><
                                NlaIII ><
                                NdeII ><
                                MflI ><
                                MboI ><
                                >< ThaI
                                >< MvnI
                                >< Ksp632I
                                >< EarI
                                >< Eam1104I
>< HinPII >< MluI
>< Hin6I >< BstUI
>< HhaI >< Bsp50I >< RsaI
>< NlaIII >< CfoI >< AflIII >< Csp6I >< Tru9I BspAI ><
>< BspWI >< BspWI >< AccII >< AfaI >< MseI BglII ><
GCATGATGTG CTATAAGCGC AATCGTGCCA CACGCGTTGA GTGTACAAC TTTGTTAATG GCATGAAGAG
  7430      7440      7450      7460      7470      7480      7490

                                >< Pali
                                >< HaeIII
                                >< DsaI
                                >< MunI
>< MboII >< BsuRI
>< DpnI >< BshI >< MunI BsmAI ><
>< Bsp143I >< MnlI >< BsaJI >< PleI>< HinfI Alw26I ><
ATCTTTCTAT GTCTATGCAA ATGGAGGCCG TGGCTTCTGC AAGACTCACA ATTGGAATTG TCTCAATTGT
  7500      7510      7520      7530      7540      7550      7560

                                >< RsaI
                                >< Csp6I
                                >< BsrI
                                >< AfaI
                                >< GsuI
                                >< BpmI
                                >< MaeIIIDraI ><
                                >< BsrI
GACACATTTT GCACTGGTAG TACATTCATT AGTGATGAAG TTGCTCGTGA TTTGTCACCTC CAGTTTAAAA
  7570      7580      7590      7600      7610      7620      7630

                                >< ThaI
                                >< MvnI
                                >< HphI
                                >< HinPII
                                >< HinPII
                                >< Hin6I
                                >< Hin6I
                                >< HhaI ><
                                >< HhaI
                                >< CfoI ><
                                >< CfoI
                                >< BstUI
                                >< BssHII
                                Bsp50I ><
                                >< BsrI
                                >< AccII
GACCAATCAA CCCTACTGAC CAGTCATCGT ATATTGTTGA TAGTGTTGCT GTGAAAAATG GCGCGCTTCA
  7640      7650      7660      7670      7680      7690      7700

```

FIGURE 13. 18

# EP 1 694 829 B1

```

                                >< FokI
                                >< BsmAI
                                >< MnlI
                                >< Alw26I
                                >< AciI
CCTCTACTTT GACAAGGCTG GTCAAAGAC CTATGAGAGA CATCCGCTCT CCCATTTTGT CAATTTAGAC
7710      7720      7730      7740      7750      7760      7770

                                >< VspI
                                >< Tru9I
                                >< MseI
                                >< AsnI
                                >< AseI
                                >< BcgI/a
> < AluI
AATTTGAGAG CTAACAACAC TAAAGGTTCA CTGCCTATTA ATGTCATAGT TTTTGATGGC AAGTCCAAAT
7780      7790      7800      7810      7820      7830      7840

                                >< SfcI
                                >< PvuII
                                >< RsaI
                                >< Psp5I
                                >< PleI
                                >< Csp6I
                                >< NspBII
>< HinfI >< DdeI >< BcgI >< AfaI >< AluI
GCGACGAGTC TGCTTCTAAG TCTGCTTCTG TGTACTACAG TCAGCTGATG TGCCAACCTA TTCTGTTGCT
7850      7860      7870      7880      7890      7900      7910

                                TthHB8I ><
                                TaqI ><
                                SalI ><
                                RtrI ><
                                HindII >
                                HincII >
                                >< ScaI
                                >< RsaI
                                >< Tru9I
                                >< Csp6I
                                >< SfaNI >< Eco57I
                                >< MaeII
                                >< AfaI
                                >< MseI
                                AccI ><
>< AluI
TGACCAAGCT CTTGTATCAG AC GTTGGAGA TAGTACTGAA GTTTCGGTTA AGATGTTTGA TGCTTATGTC
7920      7930      7940      7950      7960      7970      7980

                                >< Tru9I
                                >< MseI
                                > < Esp4I
                                > < AflIII
                                >< SfcI
                                >< BspWI
                                >< AluI
GACACCTTTT CAGCAACTTT TAGTGTTCCT ATGGA AAAAC TTAAGGCACT TGTTGCTACA GCTCACAGCG
7990      8000      8010      8020      8030      8040      8050

                                >< PvuII
                                >< Psp5I
                                >< NspBII
                                >< Fnu4HI
                                >< AluI
                                >< BbvI
AGTTAGCAAA GGGTGTAGCT TTAGATGGTG TCCTTTCTAC ATTCTGTGTC GCTGCCCGAC AAGGTGTTGT
8060      8070      8080      8090      8100      8110      8120

                                >< AluI
                                >< HindII
                                >< HincII
                                >< BsmAI
                                >< DdeI
                                >< FokI>< Alw26I
                                >< BfrI
                                >< MaeIII ><
TGATACCGAT GTTGACACAA AGGATGTTAT TGAATGTCTC AAAC TTTCAC ATCACTCTGA CTTAGAAGTG
8130      8140      8150      8160      8170      8180      8190

                                >< XhoII
                                Sau3AI ><
                                >< NdeII
                                >< MflI
                                >< MboI
                                >< NlaIII >< HgaI
                                >< HinfI >< DpnII
                                DpnI ><

```

FIGURE 13.19

```

                                Bsp143I ><
                                >< BsaHI >< BstYI
                                >< BbiIII >< BspAI
                                >< AcyI >< BglIII
                                >< MaeIII>< HphI
                                >< MaeIII >< HphI >< NlaIII
ACAGGTGACA GTTGTAAACA TTTCATGCTC ACCTATAATA AGGTTGAAAA CATGACGCCC AGAGATCTTG
      8200      8210      8220      8230      8240      8250      8260

      >< NspI
      >< NspHI
      >< NlaIII
>< HinPII
>< Hin6I
  >< HhaI
  >< CfoI
                                >< BspWI >< MaeIII
GCGCATGTAT TGACTGTAAT GCAAGGCATA TCAATGCCCA AGTAGCAAAA AGTCACAATG TTTCACTCAT
      8270      8280      8290      8300      8310      8320      8330

                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< PvuII
                                >< Psp5I
                                >< Eam1105I
                                >< NspBII
                                >< BbvI
                                >< Fnu4HI
                                >< AflIII
                                >< AluI >< BbvI >< Fnu4HI
CTGGAATGTA AAAGACTACA TGTCTTTATC TGAACAGCTG CGTAAACAAA TTCGTAGTGC TGCCAAGAAG
      8340      8350      8360      8370      8380      8390      8400

                                >< RmaI
                                >< MaeI >< Eam1105I
                                >< MboII
AACACATAC CTTTTAGACT AACTTGTGCT ACAACTAGAC AGGTTGTCAA TGTCACTAACT ACTAAATCT
      8410      8420      8430      8440      8450      8460      8470

                                >< Tru9I
                                >< Pali
                                >< MseI
                                >< HaeIII
                                >< ScaI
                                >< Esp4I
                                >< RsaI >< Tru9I
                                >< BsuRI
                                >< Csp6I >< MseI
                                >< BshI
                                >< AfaI >< DraI >< AflIII >< BbvI
CACTCAAGGG TGGTAAGATT GTTAGTACTT GTTTTAACT TATGCTTAAG GCCACATTAT TGTGCGTTCT
      8480      8490      8500      8510      8520      8530      8540

                                >< RsaI
                                >< Csp6I
                                >< BsrI
                                >< NlaIII
                                >< Fnu4HI
                                >< AfaI
                                >< MaeIII
TGCTGCATTG GTTTGTTATA TCGTTATGCC AGTACATACA TTGTCAATCC ATGATGGTTA CACAAATGAA
      8550      8560      8570      8580      8590      8600      8610

                                >< MaeIII
                                >< MaeIII
                                >< FokI
                                >< MaeIII
ATCATTGGTT ACAAAGCCAT TCAGGATGGT GTCACCTCGTG ACATCATTTT TACTGATGAT TGTTTTGCAA
      8620      8630      8640      8650      8660      8670      8680

                                SfcI >
                                Fnu4HI ><
                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< HgaI
                                >< BstXI
                                >< BbvI
                                >< AluI
ATAACATGC TGGTTTTGAC GCATGGTTTA GCCAGCGTGG TGGTTCATAC AAAAATGACA AAAGCTGCCC
      8690      8700      8710      8720      8730      8740      8750

```

FIGURE 13. 20



```

                                >< ScrFI
                                >< RsaI
                                >< MvaI >< MspI
                                >< EcoRII >< HpaII
                                >< Ecl136I>< NciI
                                >< DsaV >< HapII
                                >< BstOI>< DsaV
                                >< BstNI >< Csp6I
                                >< BsiLI >< BcnIDdeI ><
                                >< ApyI >< AfaI
                                >< Fnu4HI
                                >< AluI
TGTAGTAGCT GCTATCATTA CAAGAGAGAT TGGTTTCATA GTGCCTGGCT TACCGGGTAC TGTGCTGAGA
8760      8770      8780      8790      8800      8810      8820

                                > < MaeIII >< HphI
                                >< MnlI >< BspWI
GCAATCAATG GTGACTTCTT GCATTTTCTA CCTCGTGTTC TTAGTGCTGT TGGCAACATT TGCTACACAC
8830      8840      8850      8860      8870      8880      8890

                                Tru9I >
                                SfaNI ><
                                >< RsaI
                                MseI >
                                >< BspWI >< Fnu4HI >< Csp6I
                                >< BbvI>< MnlI >< DdeI >< AfaI
CTTCCAACT CATTGAGTAT AGTGATTTTG CTACCTCTGC TTGCGTTCTT GCTGCTGAGT GTACAATTTT
8900      8910      8920      8930      8940      8950      8960

                                > < RmaI
                                >< MnlI
                                > < MaeI
TAAGGATGCT ATGGGCAAAC CTGTGCCATA TTGTTATGAC ACTAATTTGC TAGAGGGTTC TATTTCTTAT
8970      8980      8990      9000      9010      9020      9030

                                ScrFI >
                                MvaI >
                                MnlI ><
                                EcoRII ><
                                Ecl136I >
                                DsaV ><
                                BstOI >
                                BstNI >
                                >< NlaIV >< FokI
                                >< BscBI >< ApyI >
                                >< AluI
                                >< BscBI
AGTGAGCTTC GTCCAGACAC TCGTTATGTG CTTATGGATG GTTCCATCAT ACAGTTTCCT AACACTTACC
9040      9050      9060      9070      9080      9090      9100

                                >< RsaI
                                >< SfcI >< NspI
                                >< ScaI >< NspHI
                                >< RsaI >< NlaIII
                                >< SfaNI
                                > < MaeIII >< Csp6I >< NlaIII
                                >< GsuI >< AfaI >< Csp6I
                                >< BpmI >< DdeI >< AccI >< AfaI
TGGAGGGTTC TGTTAGAGTA GTAACAACTT TTGATGCTGA GTACTGTAGA CATGGTACAT GCGAAAGGTC
9110      9120      9130      9140      9150      9160      9170

                                >< SstI
                                >< SduI
                                >< SacI
                                NspII ><
                                HgiAI ><
                                Eco24I ><
                                Bsp1286I ><

```

FIGURE 13.21

## FIGURE 13.22

238

EP 1 694 829 B1

```

    >< AfaI          >< AfaI          >< Alw26I
CTTTGTGTAC CTTTTTGCTC AACAAGGAAA TGTACCTAAA ATTGCGTAGC GAGACACTGT TGCCACTTAC
    9740          9750          9760          9770          9780          9790          9800

                                >< NlaIV
                                >< DdeI
                                >< BscBI
                                >< BfrI    AluI ><
ACAGTATAAC AGGTATCTTG CTCTATATAA CAAGTACAAG TATTTTCAGTG GAGCCTTAGA TACTACCAGC
    9810          9820          9830          9840          9850          9860          9870

    >< Fnu4HI
    >< DdeI
    >< BfrI
    >< Fnu4HI
    >< BbvI    >< AluI    >< BbvI          >< DdeI >< AlwNI
TATCGTGAAG CAGCTTGCTG CCACTTAGCA AAGGCTCTAA ATGACTTTAG CAACTCAGGT GCTGATGTTG
    9880          9890          9900          9910          9920          9930          9940

                                >< SfcI          >< BsmI
                                >< PstI          >< BscCI
TCTACCAACC ACCACAGACA TCAATCACTT CTGCTGTTCT GCAGAGTGGT TTTAGGAAAA TGGCATTCCC
    9950          9960          9970          9980          9990          10000          10010

    >< RsaI
    >< NlaIII
    >< MaeIII
    >< Csp6I
    >< AfaI
    >< Tru9I
    >< MseI
GTCAGGCAAA GTTGAAGGGT GCATGGTACA AGTAACCTGT GGAACACAA CTCTTAATGG ATTGTGGTTG
    10020          10030          10040          10050          10060          10070          10080

                                XhoII ><
                                Sau3AI ><
                                >< Tru9I    NdeII ><
                                >< NspI      MflI ><
                                >< NspHI     MboI ><
                                >< NlaIII    DpnII ><
                                >< MseI      BstYI ><
                                >< MboII    BspAI ><
                                >< BbsI      BglII ><
GATGACACAG TATACTGTCC AAGACATGTC ATTTGCACAG CAGAAGACAT GCTTAATCCT AACTATGAAG
    10090          10100          10110          10120          10130          10140          10150

                                Pali >
                                MscI >
                                HaeIII >
                                EaeI ><
                                BsuRI >
                                BshI >
                                Bali >
>< DpnI >< MboII
>< Bsp143I
ATCTGCTCAT TCGCAAATCC AACCATAGCT TTCTTGTTCA GGCTGGCAAT GTTCAACTTC GTGTTATTGG
    10160          10170          10180          10190          10200          10210          10220

                                >< DdeI> < Tru9I
                                >< BfrI> < MseI
                                >< DdeI
CCATTCTATG CAAAATTGTC TGCTTAGGCT TAAAGTTGAT ACTTCTAACC CTAAGACACC CAAGTATAAA
    10230          10240          10250          10260          10270          10280          10290

    >< ScrFI
    >< MvaI
    >< EcoRII
    >< Ecl136I
                                >< SphI

```

FIGURE 13.23

# EP 1 694 829 B1

```

>< DsaV
>< BstOI
>< BstNI
>< BsiLI
>< ApyI
TTTGTCCGTA TCCAACCTGG TCAAACATTT TCAGTTCTAG CATGCTACAA TGGTTCACCA TCTGGTGT TT
10300 10310 10320 10330 10340 10350 10360

>< PaeI
>< NspI
>< NspHI
>< RmaI >< NlaIII
>< MaeI >< HphI

>< Sau3AI
>< NdeII
>< MboI>< NlaIII
>< DpnII
>< Eco31I
>< BsmAI
>< BsaI>< NlaIII
>< Alw26I
ATCAGTGTGC CATGAGACCT AATCATACCA TTAAAGGTTT TTTCTTAAT GGATCATGTG GTAGTGT TGG
10370 10380 10390 10400 10410 10420 10430

>< Zsp2I
>< Ppu10I
>< NsiI>< SfaNI
>< NdeI
>< Mph1103I RsaI ><
>< EcoT22I Csp6I ><
>< Tru9I
>< MseI
>< AvaIII >< AluI AfaI ><
TTTTAACATT GATTATGATT GCGTGTCTTT CTGCTATATG CATCATATGG AGCTTCCAAC AGGAGTACAC
10440 10450 10460 10470 10480 10490 10500

>< SinI
>< Sau96I
>< NspIV
>< NspHII
>< Eco47I
>< Cfr13I
>< BsiZI
>< Bme18I >< HindII
>< AvaII >< HincII
>< AsuI>< BsgI >< BbvI >< BspMI AfaI ><
GCTGGTACTG ACTTAGAAGG TAAATTCTAT GGTCCATTTG TTGACAGACA AACTGCACAG GCTGCAGGTA
10510 10520 10530 10540 10550 10560 10570

>< Tru9I
>< MseI
>< BbvI
>< Fnu4HI
CAGACACAAC CATAACATTA AATGTTTTGG CATGGCTGTA TGCTGCTGTT ATCAATGGTG ATAGGTGGTT
10580 10590 10600 10610 10620 10630 10640

>< Tru9I
>< TfiI
>< MseI
>< HphI
>< HinfI
>< Tru9I
>< MseI
>< RsaI
>< Csp6I
>< AfaI
TCTTAATAGA TTCACCACTA CTTTGAATGA CTTTAACCTT GTGGCAATGA AGTACAATA TGAACCTTTG
10650 10660 10670 10680 10690 10700 10710

>< SinI
>< Sau96I
>< PssI
>< Psp5II
>< PpuMI
>< NspIV
>< NspHII
>< NlaIV

```

FIGURE 13. 24

EP 1 694 829 B1

```

                                >< EcoO109I
                                >< Eco47I
    >< Sau3AI                    >< DraII
    >< NdeII                     >< Cfr13I
    >< MboI                      >< Bsi2I
    >< DpnII>< NlaIII            >< BscBI
    >< DpnI >< HindII            >< Bme18I
    >< BspAI >< HincII           >< AvaII
    >< Bsp143I                  >< AsuI
ACACAAGATC ATGTTGACAT ATTGGGACCT CTTTCTGCTC AACAGGAAT TGCCGTCTTA GATATGTGTG
10720      10730      10740      10750      10760      10770      10780

                                >< StyI
                                >< RsaI
                                >< EcoT14I
                                >< Eco130I
                                >< SfcI
                                >< Csp6I
    >< Fnu4HI                    >< Fnu4HI
    >< BbvI                      >< Fnu4HI
    >< BbvI                      >< BssT1I
    >< BbvI                      >< BsaJI
    >< AluI                      >< PstI
    >< PstI                      >< AfaI
CTGCTTTGAA AGAGCTGCTG CAGAATGGTA TGAATGGTCG TACTATCCTT GGTAGCACTA TTTTAGAAGA
10790      10800      10810      10820      10830      10840      10850

                                >< StyI
                                >< EcoT14I
                                >< Eco130I
                                >< BssT1I
    >< MboII                    >< MaeIII>< BsaJI
TGAGTTTACA CCATTTGATG TTGTTAGACA ATGCTCTGGT GTTACCTTCC AAGGTAAGTT CAAGAAAATT
10860      10870      10880      10890      10900      10910      10920

                                >< SfaNI
                                >< SduI
                                >< NspII
    >< Tru9I> < Bsp1286I        >< Tru9I
    >< Tru9I> < Bsp1286I        >< MseI
    >< MseI> < BmyI             >< FokI
GTTAAGGGCA CTCATCATTG GATGCTTTTA ACTTTCTTGA CATCACTATT GATTCTTGTT CAAAGTACAC
10930      10940      10950      10960      10970      10980      10990

                                >< XmnI
                                >< BsmI
                                >< BscCI
    >< MaeIII                    >< Asp700I
AGTGGTCACT GTTTTCTTTT GTTTACGAGA ATGCTTTCTT GCCATTTACT CTTGGTATTA TGGCAATTGC
11000      11010      11020      11030      11040      11050      11060

                                >< NspI
                                >< NspHI
                                >< NlaIII
    >< NspI                      >< Tru9I
    >< NspHI                    >< MseI
    >< NlaIII                    >< BsmI
    >< BspWI >< Fnu4HI>< BspWI >< BscCI
    >< BspWI >< Fnu4HI>< BspWI >< BscCI
TGCAATGTGCT ATGCTGCTTG TTAAGCATAA GCACGCATTC TTGTGCTTGT TTCTGTTACC TTCTCTTGCA
11070      11080      11090      11100      11110      11120      11130

                                >< SfaNI
                                >< RmaI
                                >< NspI
                                >< NlaIII
    >< NspI                      >< MamI
    >< NlaIII                    >< HphI
    >< NheI                      >< BspHI
    >< Tru9I                    >< MaeI
    >< BspWI >< MseI >< AccI> < NspHI>< AluI >< BsaBI >< NlaIII
ACAGTTGCTT ACTTTAATAT GGTCTACATG CCTGCTAGCT GGGTGATGCG TATCATGACA TGGCTTGAAT
11140      11150      11160      11170      11180      11190      11200

```

FIGURE 13.25



EP 1 694 829 B1

```

                >< DrdI >< Alw26I
GTGTTTATGA CTACTTGGTC TCTACACAAG AATTTAGGTA TATGAACTCC CAGGGGCTTT TGCCTCCTAA
    11630      11640      11650      11660      11670      11680      11690

                >< Tru9I
                >< MseI
>< SfaNI          > < HindIII> < Tru9I
    >< MnlI          >< AluI > < MseI > < MnlI          > < NlaIII
GAGTAGTATT GATGCTTTCA AGCTTAACAT TAAGTTGTTG GGTATTGGAG GTAAACCATG TATCAAGGTT
    11700      11710      11720      11730      11740      11750      11760

                >< VneI
                >< SnoI
                    >< SduI
                    >< NspII
                    >< HgiAI
                    >< Bsp1286I
                    >< BmyI >< RsaI
                >< RsaI          >< ApaLI          >< MboII
                >< Csp6I          >< Alw44I          >< Csp6I          DdeI >
                >< AfaI          >< MaeII          >< Alw21I >< AfaI          BfrI >
GCTACTGTAC AGTCTAAAT GTCTGACGTA AAGTGCACAT CTGTGGTACT GCTCTCGGTT CTTCAACAAC
    11770      11780      11790      11800      11810      11820      11830

                >< NspII> < RsaI
                >< DraIII
                >< SduI>< Csp6I
                >< Bsp1286I
                >< MboII          >< HinfI >< PfuI          >< BmyI > < AfaI          >< MboII
TTAGAGTAGA GTCATCTTCT AAATTGTGGG CACAATGTGT ACAACTCCAC AATGATATTC TTCTTGCAAA
    11840      11850      11860      11870      11880      11890      11900

                >< TthHB8I
                >< TaqI
                    >< HindIII          >< MboII          SfcI ><
                    >< AluI          > < Eco57I          >< NlaIII
AGACACAACCT GAAGCTTTTCG AGAAGATGGT TTCTCTTTTG TCTGTTTTGC TATCCATGCA GGGTGCTGTA
    11910      11920      11930      11940      11950      11960      11970

                >< VspI
                >< Tru9I          > < Ksp632I
                >< MseI          >< TthHB8I          > < EarI
                >< AsnI          >< TaqI >< MboII          > < Eam1104I
                >< AseI>< MnlI >< BcgI/a          >< Eco57I          >< Eco57I >< BcgI
GACATTAATA GGTGTGCGA GGAAATGCTC GATAACCGTG CTACTCTTCA GGCTATTGCT TCAGAATTTA
    11980      11990      12000      12010      12020      12030      12040

                >< StuI
                >< ScrFI
                    >< PstI
                    >< MvaI>< HaeIII
                >< EcoRI>< Eco147I
                >< Ecl136I
                >< DsaV >< BsuRI
                >< BstOI
                >< BstNI
                    >< BspWI
                    >< BsiLI
                >< Fnu4HI          >< BsaJI >< BshI          TfiI ><
                >< NdeI          >< BspWI>< MnlI >< BglI          >< SfcI HinfI ><
                >< AciI          >< ApyI>< AatI          > < AluI

```

FIGURE 13. 27

# EP 1 694 829 B1

```

GTTCTTTACC ATCATATGCC GCTTATGCCA CTGCCCAGGA GGCCTATGAG CAGGCTGTAG CTAATGGTGA
12050      12060      12070      12080      12090      12100      12110

    >< XmnI      >< Tru9I      >< SfaNI
    >< HphI      >< MseI      >< DdeI
    >< Asp700I   >< Eco57I   >< BbvI Fnu4HI ><
TTCTGAAGTC GTTCTCAAAA AGTTAAAGAA ATCTTTGAAT GTGGCTAAAT CTGAGTTTGA CCGTGATGCT
12120      12130      12140      12150      12160      12170      12180

                                XhoII ><
                                Sau3AI ><
                                NdeII ><
                                MnlI >
                                >< MnlI
                                >< MflI
                                >< MboI
                                DpnII ><
                                DpnI ><
                                DdeI ><
                                BstYI ><
                                >< BspWI      >< RsaIBspAI ><
                                >< BspAI      >< Csp6IBsp143I ><
    >< NlaIII      >< Bsp143I      >< AfaIBglII ><
GCCATGCAAC GCAAGTTGGA AAAGATGGCA GATCAGGCTA TGACCCAAAT GTACAAACAG GCAAGATCTG
12190      12200      12210      12220      12230      12240      12250

                                >< SpeI      >< Ksp632I > < HindIII
                                >< RmaI      >< DdeI >< SfaNI
                                >< MaeIII      >< MboII      >< Eam1104I >< BspWI
                                >< MaeI      >< BspWI      >< EarI>< BfrI >< AluI
AGGACAAGAG GGCAAAAGTA ACTAGTGCTA TGCAAACAAT GCTCTTCACT ATGCTTAGGA AGCTTGATAA
12260      12270      12280      12290      12300      12310      12320

                                >< ThaI
                                >< MvnI
                                >< HinPII
                                >< Hin6I
                                >< HhaI
                                >< CfoI
                                >< BstUI
                                >< Bsp50I
                                >< AccII      SfcI ><
TGATGCACTT AACAACATTA TCAACAATGC GCGTGATGGT TGTGTTCCAC TCAACATCAT ACCATTGACT
12330      12340      12350      12360      12370      12380      12390

                                >< RsaI
                                >< NlaIV
                                >< Eco64I
                                >< Csp6I
                                >< BslI
                                >< BsiYI>< KpnI
                                >< BscBI
                                >< BanI
                                >< Asp718
                                >< AfaI
                                >< AccB1I      >< MaeIII
                                >< Fnu4HI >< BbvI      >< Acc65I      BsgI ><
ACAGCAGCCA AACTCATGGT TGTTGTCCCT GATTATGGTA CCTACAAGAA CACTTGTGAT GGTAACACCT
12400      12410      12420      12430      12440      12450      12460

    >< Zsp2I
    >< Ppu10I

```

FIGURE 13.28



# EP 1 694 829 B1

```

>< NsiI
>< Mph1103I
>< NdeI>< EcoT22I
>< AvaIII >< SfaNI
>< SfaNI
>< AciI
DdeI ><
BfrI ><
TTACATATGC ATCTGCACTC TGGGAAATCC AGCAAGTTGT TGATGCGGAT AGCAAGATTG TTCAACTTAG
12470 12480 12490 12500 12510 12520 12530

>< Pali
>< HaeIII >< MnlI >< DdeIDdeI ><
>< Tru9I>< NlaIII
>< BsuRI >< MaeIII >< BspWI
>< MseI>< HphI >< XcmI>< BshI >< AluI BspWI ><
TGAAATTAAC ATGGACAATT CACCAAATTT GGCTTGGCCT CTTATTGTTA CAGCTCTAAG AGCCAACTCA
12540 12550 12560 12570 12580 12590 12600

RsaI ><
NlaIV ><
KpnI ><
>< Fnu4HI
Eco64I ><
Csp6I ><
BscBI ><
Asp718 ><
AfaI ><
>< AciI>< BanI
AccBI ><
>< AluI >< SfcI >< DdeI>< BsrI >< PshAI Acc65I ><
GCTGTAAAC TACAGAATAA TGAAGTACT CCAGTAGCAC TACGACAGAT GTCCTGTGCG GCTGGTACCA
12610 12620 12630 12640 12650 12660 12670

>< TthHB8I
>< TaqI
>< SfuI
>< NspV
>< MnlI
>< LspI
>< Csp45I
>< BstBI
>< Bsp119I
>< BsiCI
>< Bpu14I
>< AsuII
>< RsaI
>< Csp6I
>< AluI
>< AfaI
CACAAACAGC TTGTACTGAT GACAATGCAC TTGCCTACTA TAACAATTCG AAGGGAGGTA GGTTTGTGCT
12680 12690 12700 12710 12720 12730 12740

>< XhoII
>< Sau3AI
>< NdeII
>< MflI
>< MboI
>< DpnII
>< DpnI
>< BstYI >< TfiI >< RsaI
>< BspAI >< RmaI >< Csp6I
>< Bsp143I >< HinfI >< Csp6I>< RsaI
>< BglII >< MaeI >< DdeI >< AfaI>< AfaI
GGCATTACTA TCAGACCACC AAGATCTCAA ATGGGCTAGA TTCCCTAAGA GTGATGGTAC AGGTACAATT
12750 12760 12770 12780 12790 12800 12810

>< Sau96I
>< PssI
>< Pali
>< NspIV

```

FIGURE 13.29

# EP 1 694 829 B1

```

                                >< HaeIII
                                >< EcoO109I
                                >< DraII
                                >< Cfr13I
                                >< BsuRI
                                >< BsiZI      RsaI >
                                >< BshI      Csp6I ><
                                >< AsuI      AfaI >
TACACAGAAC TGGAACCACC TTGTAGGTTT GTTACAGACA CACCAAAAGG GCCTAAAGTG AAATACTTGT
12820      12830      12840      12850      12860      12870      12880

                                >< SfcI
                                > < MboII
                                MaeII ><
                                >< Fnu4HI >< RsaI
                                >< Eco57I >< Csp6I
                                > < BbsI
                                >< Tru9I
                                >< MseI >< MnlI
                                >< BbvI      >< AluI      >< AfaI
ACTTCATCAA AGGCTTAAAC AACCTAAATA GAGGTATGGT GCTGGGCAGT TTAGCTGCTA CAGTACGTCT
12890      12900      12910      12920      12930      12940      12950

                                >< RsaI
                                >< SfcI >< Csp6I
                                >< BspWI      >< AfaI      >< BspMI
TCAGGCTGGA AATGCTACAG AAGTACCTGC CAATTCAACT GTGCTTTCCT TCTGTGCTTT TGCAGTAGAC
12960      12970      12980      12990      13000      13010      13020

                                >< RmaI
                                >< MnlI
                                >< MaeI      >< HphI
CCTGCTAAAG CATATAAGGA TTACCTAGCA AGTGGAGGAC AACCAATCAC CAACTGTGTG AAGATGTTGT
13030      13040      13050      13060      13070      13080      13090

                                >< SinI
                                >< Sau96I
                                >< NspIV
                                >< NspHII
                                >< NlaIII
                                >< Eco47I
                                >< Eam1105I
                                >< Cfr13I
                                >< BsiZI
                                >< Bme18I >< XcmI
                                >< AvaII      PleI ><
                                >< AfaI      >< AfaI      >< MaeIII      >< AluI      >< AsuI >< HinfI
GTACACACAC TGGTACAGGA CAGGCAATTA CTGTAACACC AGAAGCTAAC ATGGACCAAG AGTCCTTTGG
13100      13110      13120      13130      13140      13150      13160

                                >< TfiI
                                >< SfaNI
                                >< NlaIII      >< FokI
                                >< HinfI
TGGTGCTTCA TGTGTCTGT ATTGTAGATG CCACATTGAC CATCCAAATC CTAAAGGATT CTGTGACTTG
13170      13180      13190      13200      13210      13220      13230

                                > < RsaI
                                >< MaeII
                                >< Csp6I
                                > < AfaI
                                >< BsrI      >< DdeI
                                >< BfrI
AAAGGTAAGT ACGTCCAAAT ACCTACCACT TGTGCTAATG ACCCAGTGGG TTTTACACTT AGAAACACAG
13240      13250      13260      13270      13280      13290      13300

                                >< ThaI

```

FIGURE 13.30

# EP 1 694 829 B1

```

                                >< SfaNI
                                >< MvnI
                                >< BstUI
                                >< Bsp50I
                                >< AciI
>< RsaI
>< Csp6I
>< AfaI >< AciI                                >< SfcI >< MaeIII                                >< AccIISfaNI ><
TCTGTACCGT CTGCGGAATG TGGAAAGGTT ATGGCTGTAG TTGTGACCAA CTCCGCGAAC CCTTGATGCA
13310      13320      13330      13340      13350      13360      13370

                                >< Zsp2I
                                > < SfaNI
                                >< Mph1103I>< Tru9I
>< Ppu10I>< MaeII                                Fnu4HI ><
                                >< NsiI> < FokI                                BsgI ><
                                >< EcoT22I >< MseI                                >< BbvI
>< AciI>< AvaIII                                >< DraI                                >< AciI                                >< Fnu4HI                                AciI ><
GTCTGCGGAT GCATCAACGT TTTTAAACGG GTTTGCGGTG TAAGTGAGC CCGTCTTACA CCGTGCGGCA
13380      13390      13400      13410      13420      13430      13440

>< SpeI
>< ScaI
>< RsaI
>< RmaI
>< MaeI
> < Csp6I                                >< SfcI                                >< BspWI
>< BspWI >< AfaI                                >< AccI                                >< BcgI/a                                BcgI >
CAGGCACTAG TACTGATGTC GTCTACAGGG CTTTTGATAT TTACAACGAA AAAGTTGCTG GTTTTGCAAA
13450      13460      13470      13480      13490      13500      13510

                                >< ScrFI
                                >< MvaI
                                >< MnlI
>< EcoRII
>< Ecl136I
>< BstOI
>< BstNI
>< BslI
>< DsaV >< BsiYI
>< BsiLI
>< ApyI                                >< PleI
                                > < FokI >< HinfI
GTTCTAATAA ACTAATTGCT GTCGCTTCCA GGAGAAGGAT GAGGAAGGCA ATTTATTAGA CTCTTACTTT
13520      13530      13540      13550      13560      13570      13580

                                >< NlaIII
                                >< Ksp632I
                                >< EarI
                                >< Eam1104I
>< Tru9I
>< MseI
>< MnlI                                >< BsmAI                                >< Tru9I
>< MnlI                                >< Alw26I                                >< MboII                                >< MseI
GTAGTTAAGA GGCATACTAT GTCTAACTAC CAACATGAAG AGACTATTTA TAACTTGGTT AAAGATTGTC
13590      13600      13610      13620      13630      13640      13650

                                >< RsaI
                                >< NlaIV
> < NlaIII
                                >< KpnI
                                >< HphI
> < Eco64I
>< Csp6I
>< BscBI
> < BanI
> < Asp718

```

FIGURE 13.31

# EP 1 694 829 B1

```

                                >< MaeIII >< AfaI
                                > < AccBII MaeII ><
>< NspBII                                >< Acc65I > < HgaI
>< AciI                                >< NlaIII
CAGCGGTTGC TGTCCATGAC TTTTTCAGT TTAGAGTAGA TGGTGACATG GTACCACATA TATCACGTCA
13660      13670      13680      13690      13700      13710      13720

                                >< MnlI
                                >< MaeII
GCGTCTAACT AAATACACAA TGGCTGATTT AGTCTATGCT CTACGTCATT TTGATGAGGG TAATTGTGAT
13730      13740      13750      13760      13770      13780      13790

>< Tru9I
>< MseI                                >< MaeIII >< MunI
ACATTAAAAG AAATACTCGT CACATACAAT TGCTGTGATG ATGATTATTT CAATAAGAAG GATTGGTATG
13800      13810      13820      13830      13840      13850      13860

                                >< ThaI
                                >< MvnI
                                >< MluI
                                >< BstUI
                                >< Bsp50I
                                >< RsaI
                                >< HphI
                                >< Csp6I Tru9I ><
>< TfiI                                >< AflIII >< DdeI
>< HinfI                                >< AccII >< BfrI
ACTTCGTAGA GAATCCTGAC ATCTTACGCG TATATGCTAA CTTAGGTGAG CGTGACGCC AATCATTATT
13870      13880      13890      13900      13910      13920      13930

                                >< XhoII >
                                >< Sau3AI >
                                >< NdeII >
                                >< MflI >
                                >< RsaI
                                >< Csp6I
                                >< BspWI
                                >< AfaI
                                >< BspAI >
> < SfaNI                                >< RsaI
>< RsaI                                > < Csp6I
>< Csp6I                                >< BspWI
>< AfaI                                >< AfaI
AAAGACTGTA CAATTCTGCG ATGCTATGCG TGATGCAGGC ATTGTAGGCG TACTGACATT AGATAATCAG
13940      13950      13960      13970      13980      13990      14000

                                > < ScrFI
                                > < MvaI
                                >< Fnu4HI
                                >< EcoRII
                                > < Ecl136I
                                > < BstOI
                                > < BstNI
>< Tru9I                                >< RsaI
>< MseI                                >< BslI
>< DpnI                                > < HphI
>< Bsp143I >< Csp6I >< Csp6I >< BsiLI
>< BsrI                                > < BbvI > < ApyI
>< AlwI                                >< AfaI >< DsaV >< AciI
GATCTTAATG GGAAGTGGTA CGATTTCGGT GATTTCGTAC AAGTAGCACC AGGCTGCGGA GTTCCTATTG
14010      14020      14030      14040      14050      14060      14070

                                >< SfaNI
                                > < HinfI
                                >< RmaI
                                >< MnlI
                                >< Fnu4HIpleI ><
>< TfiI                                >< BsiBI
>< HinfI                                >< BsaBI
>< FokI                                >< BbvI
TGGATTCATA TTACTCATTG CTGATGCCCA TCCTCACTTT GACTAGGGCA TTGGCTGCTG AGTCCCATAT
14080      14090      14100      14110      14120      14130      14140

>< Sau3AI
>< NdeII

```

FIGURE 13.32

# EP 1 694 829 B1

```

>< MboI
>< MamI
>< DpnII
>< DpnI
>< BspWI
>< BspAI
>< BspI43I
>< BsiBI
>< BsaBI >< FokI
GGATGCTGAT CTCGCAAAAC CACTTATTAA GTGGGATTTG CTGAAATATG ATTTTACGGA AGAGAGACTT
14150      14160      14170      14180      14190      14200      14210

>< SinI
>< Sau96I
>< NspIV
>< NspHII
>< NlaIV
>< FokI
>< McrI
>< Eco47I
>< Ksp632I
>< Cfr13I
>< EarI
>< BsiZI
>< Eam1104I >< SspI>< BscBI
>< BsmAI >< Tru9I >< Bme18I
>< MboII >< BsiEI>< MseI >< AvaII
>< Alw26I >< DraI >< AsuI >< MunI >< MseI
TGTCTCTTCG ACCGTATTAT TAAATATTGG GACCAGACAT ACCATCCCAA TTGTATTAACT TGTTCGGATG
14220      14230      14240      14250      14260      14270      14280

>< SinI ><
>< Sau96I ><
>< NspIV ><
>< NspHII >
>< Eco47I ><
>< Cfr13I ><
>< BsiZI ><
>< Bme18I ><
>< AvaII ><
>< AsuI ><
>< FokI
>< MseI
ATAGGTGTAT CCTTCATTGT GCAAACCTTA ATGTGTTATT TTCTACTGTG TTTCACCTA CAAGTTTGGG
14290      14300      14310      14320      14330      14340      14350

>< SpeI
>< RmaI
>< MaeI >< SspI
ACCACTAGTA AGAAAAATAT TTGTAGATGG TGTTCCTTTT GTTGTTCCTT CAAGTATACCA TTTTCGTGAG
14360      14370      14380      14390      14400      14410      14420

>< ThaI>< Esp3I
>< DdeI
>< BstUI
>< Bsp50I >< BsmBI
>< MvnI>< BsmAI
>< HgaI>< AluI >< Alw26I
>< FokI >< AccII >< BbvI
TTAGGAGTCG TACATAATCA GGATGTAAAC TTACATAGCT CGCGTCTCAG TTTCAAGGAA CTTTGTAGTG
14430      14440      14450      14460      14470      14480      14490

>< Zsp2I
>< SphI
>< Ppu10I
>< PaeI
>< NspI

```

FIGURE 13.33

```

    >< Sau3AI          >< NspHI
    >< NdeII           >< NsiI
    >< MboI            >< NlaIII
    >< DpnII           >< Mph1103I
    > < DpnI          >< Fnu4HI
    >< Fnu4HI>< BspWI  >< EcoT22I
    >< BspAI           >< BspWI
    > < Bsp143I> < AvaIII > < AlwNI
    >< AlwI            >< AluI          >< AluI >< BbvI >< MaeI
    ATGCTGCTGA TCCAGCTATG CATGCAGCTT CTGGCAATTT ATTGCTAGAT AAACGCACTA CATGCTTTTC
    14500      14510      14520      14530      14540      14550      14560

                                >< ScrFI
                                >< NciI
                                >< MspI
                                >< HpaII
                                >< HapII
    >< Fnu4HI
    >< AlwNI
    >< AluI
    AGTAGCTGCA CTAACAAACA ATGTTGCTTT TCAAACTGTC AAACCCGGTA ATTTTAATAA AGACTTTTAT
    14570      14580      14590      14600      14610      14620      14630

                                >< Tru9I
                                >< MseI
                                >< MboII
                                >< BbvI
    GACTTTGCTG TGTCTAAAGG TTTCTTTAAG GAAGGAAGTT CTGTTGAACT AAAACACTTC TTCTTTGCTC
    14640      14650      14660      14670      14680      14690      14700

                                >< FokI
                                >< Fnu4HI
                                >< EcoRV
                                >< Eco32I
    AGGATGGCAA CGCTGCTATC AGTGATTATG ACTATTATCG TTATAATCTG CCAACAATGT GTGATATCAG
    14710      14720      14730      14740      14750      14760      14770

                                >< VspI
                                >< Tru9I
                                >< MseI
                                >< AsnI
                                >< AseI
    ACAACTCCTA TTCGTAGTTG AAGTTGTTGA TAAATACTTT GATTGTTACG ATGGTGGCTG TATTAATGCC
    14780      14790      14800      14810      14820      14830      14840

    >< Tru9I
    >< MseI
    >< HpaI
    >< HindII
    >< HincII
    >< PvuII
    >< Psp5I
    >< NspBII
    >< AluI
    >< XcmI
    >< Tru9I
    >< MseI
    >< RmaI
    >< MaeI
    AACCAAGTAA TCGTTAACAA TCTGGATAAA TCAGCTGGTT TCCCATTTAA TAAATGGGGT AAGGCTAGAC
    14850      14860      14870      14880      14890      14900      14910

    >< SfaNI
    >< Sau3AI
    >< NdeII
    >< MboI
    >< DpnII
    >< DpnI
    >< Bsp143I
    >< PleI
    >< HinfI>< MnlI
    >< BspAI
    >< AlwI
    >< ThaI
    >< MvnI
    >< BstUI
    >< Bst1107I
    >< BspWI
    >< FokI
    >< Bsp50I
    >< AccII>< DdeI
    >< AccI
    TTTATTATGA CTCAATGAGT TATGAGGATC AAGATGCACT TTTCGCGTAT ACTAAGCGTA ATGTCATCCC
    14920      14930      14940      14950      14960      14970      14980

                                >< SstI
                                >< SduI
                                >< SacI

```

FIGURE 13.34

```

                                >< NspII
                                >< HgiAI
                                >< Eco24I
                                > < Ecl136II
                                >< Bsp1286I
                                >< BmyI
                                >< BanII
                                >< Alw21I
                                > < AluI
                                >< AluI
TACTATAACT CAAATGAATC TTAAGTATGC CATTAGTGCA AAGAATAGAG CTCGCACCGT AGCTGGTGTC
14990      15000      15010      15020      15030      15040      15050

                                >< ScaI
                                >< SfcI>< RsaI
                                >< BsmAI >< Csp6I
                                >< Alw26I >< AfaI
TCTATCTGTA GTACTATGAC AAATAGACAG TTTCATCAGA AATTATTGAA GTCAATAGCC GCCACTAGAG
15060      15070      15080      15090      15100      15110      15120

                                >< Tru9I
                                >< MseI
GAGCTACTGT GGTAATTGGA ACAAGCAAGT TTTACGGTGG CTGGCATAAT ATGTTAAAAA CTGTTTACAG
15130      15140      15150      15160      15170      15180      15190

                                NspI ><
                                NspHI ><
                                NlaIII ><
                                >< NlaIII
                                DdeI ><
                                BspWI ><
                                >< MaeIII
                                BfrI ><
TGATGTAGAA ACTCCACACC TTATGGGTTG GGATTATCCA AAATGTGACA GAGCCATGCC TAACATGCTT
15200      15210      15220      15230      15240      15250      15260

                                > < PalI
                                > < HaeIII
                                > < BsuRI
                                > < BshI
                                >< MnlI
                                >< MaeIII
                                SfcI ><
AGGATAATGG CCTCTCTTGT TCTTGCTCGC AAACATAACA CTTGCTGTAA CTTATCACAC CGTTTCTACA
15270      15280      15290      15300      15310      15320      15330

                                Tru9I ><
                                ScrFI >
                                MvaI >
                                >< MseI
                                FokI ><
                                EcoRII ><
                                Ecl136I >
                                DsaV ><
                                BstOI >
                                >< NlaIII
                                > < Fnu4HI
                                BstNI >
                                BsiLI >
                                ApyI >
                                >< AluI
                                >< AviII >< MseI
                                >< AciI
                                >< AluI ><
GGTTAGCTAA CGAGTGTGCG CAAGTATTAA GTGAGATGGT CATGTGTGGC GGCTCACTAT ATGTTAAACC
15340      15350      15360      15370      15380      15390      15400

                                > < SfaNI
                                >< MspI
                                >< HpaII
                                >< HapII
                                >< HphI
                                >< BspWI
                                >< Tru9I
                                >< MseI
                                MaeIII ><
                                AluI ><

```

FIGURE 13.35

# EP 1 694 829 B1

```

AGGTGGAACA TCATCCGGTG ATGCTACAAC TGCTTATGCT AATAGTGTCT TTAACATTTG TCAAGCTGTT
15410      15420      15430      15440      15450      15460      15470

                                >< DrdI
>< BspWI                                >< AluI                                > < AciI
ACAGCCAATG TAAATGCACT TCTTTCAACT GATGGTAATA AGATAGCTGA CAAGTATGTC CGCAATCTAC
15480      15490      15500      15510      15520      15530      15540

                                >< Sau3AI
                                >< NdeII
                                >< MboI
> < MamI
                                >< FbaI
                                >< DpnII
                                >< DpnI
                                >< BspHI
                                >< BspAI
                                >< BspI43I
                                >< BsiQI
                                >< SfcI                                > < BsiBI>< NlaIII
                                >< BsmAI                                > < BsaBI>< FokI
                                >< Alw26I                                >< BclI>< EcoRI                                FokI ><
AACACAGGCT CTATGAGTGT CTCTATAGAA ATAGGGATGT TGATCATGAA TTCGTGGATG AGTTTTACGC
15550      15560      15570      15580      15590      15600      15610

                                >< TfiI
                                >< SfaNI
                                >< NlaIII
                                >< BspMI                                >< HinfI                                >< MaeIII
TTACCTGCGT AAACATTTCT CCATGATGAT TCTTTCTGAT GATGCCGTTG TGTGCTATAA CAGTAACTAT
15620      15630      15640      15650      15660      15670      15680

                                > < RmaI
                                >< NheI >< Tru9I
>< Fnu4HI                                > < MaeI                                >< Tru9I
>< AciI                                >< AluI >< MseI                                >< MseI                                MnlI ><
GCGGCTCAAG GTTTAGTAGC TAGCATTAAG AACTTTAAGG CAGTTCTTTA TTATCAAAAT AATGTGTTCA
15690      15700      15710      15720      15730      15740      15750

                                >< SinI
                                >< Sau96I
                                >< PssI
                                >< Psp5II
                                >< PpuMI
                                >< NspIV
                                >< NspHII
                                >< EcoO109I
                                >< Eco47I
                                >< DraII
                                >< Cfr13I
                                >< BsiZI
                                >< Bme18I
                                >< AvaII
                                >< AsuI                                >< MnlI
                                >< DdeI
                                >< NlaIII                                >< BsmAI
                                >< DdeI                                >< Alw26I
TGTCTGAGGC AAAATGTTGG ACTGAGACTG ACCTTACTAA AGGACCTCAC GAATTTTGCT CACAGCATAC
15760      15770      15780      15790      15800      15810      15820

                                >< XhoII
                                >< Sau3AI
                                >< NdeII
                                >< MflI
                                >< MboI

```

FIGURE 13. 36



```

                                >< RsaI          >< DpnII
                                >< MaeII          >< DpnI          > < SspI
                                >< Csp6I          >< BstYI          HinPII ><
                                >< BsaAI          >< BspMI          Hin6I ><
                                >< AflIII          >< BspAI          HhaI ><
>< BspWI>< MseI          >< AfaI          >< AlwI>< Bsp143I          CfoI ><
AATGCTAGTT AAACAAGGAG ATGATTACGT GTACCTGCCT TACCCAGATC CATCAAGAAT ATTAGGCGCA
15830      15840      15850      15860      15870      15880      15890

                                >< RsaI          >< SfaNI
                                >< TthHB8I          >< Csp6I          >< MaeIII
                                >< TaqI          >< AfaI          BsrI ><
GGCTGTTTTG TCGATGATAT TGTCAAAACA GATGGTACAC TTATGATTGA AAGGTTCTGT TCACTGGCTA
15900      15910      15920      15930      15940      15950      15960

> < FokI
>< BspWI
TTGATGCTTA CCCACTTACA AAACATCCTA ATCAGGAGTA TGCTGATGTC TTCACTTGT ATTTACAATA
15970      15980      15990      16000      16010      16020      16030

                                >< Van91I
                                >< PflMI
                                >< NspI
                                > < PalI>< NspHI
                                > < MscI>< NlaIII
                                > < HaeIII
                                > < BsuRI
                                >< BsrI
                                >< EaeI          >< BslI          >< NspI
                                > < BshI>< BsiYI          >< NspHI
                                >< NlaIII          >< AflIII          >< AflIII
>< MaeIII          >< AluI > < BalI>< AccB7I          >< NlaIII
CATTAGAAAG TTACATGATG AGCTTACTGG CCACATGTTG GACATGTATT CCGTAATGCT AACTAATGAT
16040      16050      16060      16070      16080      16090      16100

                                >< RsaI> < NlaIV
                                >< MnlI
                                >< Csp6I          >< DdeI
                                >> BsrI >< MnlI          >< RsaI
                                >< AfaI> < BscBI          >< Csp6I
                                AACACCTCAC GGTACTGGGA ACCTGAGTTT TATGAGGCTA TGTACACACC ACATACAGTC TTGCAGGCTG
16110      16120      16130      16140      16150      16160      16170

                                >< NlaIV
                                >> EcoNI
                                >> Eco31I
                                >> Eco64I>< BsmAI
                                >> BscBI >< BslI
                                >> BanI          >< BsiYI
                                >< AciI          >< BsaI
                                >> AccB1I>< Alw26I          BbvI ><
>< BspWI
TAGGTGCTTG TGTATTGTGC AATTACACAGA CTTCACTTCG TTGCGGTGCC TGTATTAGGA GACCATTCCT
16180      16190      16200      16210      16220      16230      16240

                                >< Tth111I
                                >< Fnu4HI          >< NlaIII          > < Tru9I
                                >< BspWI          >< AspI          > < MseI
ATGTTGCAAG TGCTGCTATG ACCATGTCAT TTCAACATCA CACAAATTAG TGTGTCTGT TAATCCCTAT
16250      16260      16270      16280      16290      16300      16310

                                >< ScrFI
                                >< MvaI

```

FIGURE 13.37

# EP 1 694 829 B1

```

>< EcoRII
>< Ecl136I
>< DsaV
>< BstOI
>< BstNI
>< BsiLI
>< BsaJI
>< ApyI
>< MaeIII >< MaeIII
>< MaeI
>< RmaI
>< MnlI
>< BspWI ><
>< AluI
GTTTGCAATG CCCCAGGTTG TGATGTCAC TATGTGACAC AACTGTATCT AGGAGGTATG AGCTATTATT
16320 16330 16340 16350 16360 16370 16380

>< MaeIII
>< MnlI
GCAAGTCACA TAAGCCTCCC ATTAGTTTTT CATTATGTGC TAATGGTCAG GTTTTTGGTT TATACAAAAA
16390 16400 16410 16420 16430 16440 16450

>< NspI
>< NspHI
>< NlaIII>< MaeIII>< MaeIII
>< AflIII
>< AspI
>< AflIII
CACATGTGTA GGCAGTGACA ATGTCAC TCAATGCG ATAGCAACAT GTGATTGGAC TAATGCTGGC
16460 16470 16480 16490 16500 16510 16520

>< RsaI
>< P1eI
>< DdeI
>< Csp6I
>< BsmAI >< HinfI
>< Alw26I >< HindIII
>< AfaI >< AluI >< Fnu4HI >< BbvI
GATTACATAC TTGCCAACAC TTGTACTGAG AGACTCAAGC TTTTCGCAGC AGAAACGCTC AAAGCCACTG
16530 16540 16550 16560 16570 16580 16590

>< ThaI
>< ScaI
>< RsaI >< RsaI
>< MvnI
>< Csp6I >< Csp6I
>< BstUI
>< Bsp50I
>< AfaI >< AfaI
>< MseI >< NdeI
>< AluI
>< AccII
MnlI >
AGGAAACATT TAAGCTGTCA TATGGTATTG CCACTGTACG CGAAGTACTC TCTGACAGAG AATTGCATCT
16600 16610 16620 16630 16640 16650 16660

MaeIII ><
>< MaeIII
>< EcoO65I
>< Eco91I
>< BstPI
>< BstEII
>< BsrI
>< SfaNI
>< NlaIII
>< RmaI
>< MaeI
TTCATGGGAG GTTGGAAAAC CTAGACCACC ATTGAACAGA AACTATGTCT TTAGTGGTTA CCGTGTAAC T
16670 16680 16690 16700 16710 16720 16730

RsaI ><
>< MnlI
>< HphI
>< RsaI
>< RsaI
>< SfaNI
>< Csp6I >< Csp6I
>< AfaI >< AfaI
>< MaeIII
>< HphI AfaI ><
AAAAATAGTA AAGTACAGAT TGGAGAGTAC ACCTTTGAAA AAGGTGACTA TGGTGATGCT GTTGTGTACA
16740 16750 16760 16770 16780 16790 16800

```

FIGURE 13. 38

# EP 1 694 829 B1

```

    >< RsaI
    >< Csp6I
    >< AfaI
GAGGTACTAC GACATACAAG TTGAATGTTG GTGATTACTT TGTGTTGACA TCTCACA CTG TAATGCCACT
    16810      16820      16830      16840      16850      16860      16870

    >< HphI
    >< HindII
    >< HincII
    DdeI ><
    BfrI ><

    >< VneI
    >< SnoI
    >< SduI
    >< NspII
    >< HgiAI
    > < SduI
    > < NspII
    >< DraIII
    >< Bsp1286I
    >< BmyI
    >< BspWI
    >< DraIII
    >< RsaI
    >< Csp6I
    >< ApaLI
    >< RmaI
    >< Bsp1286I
    >< Csp6I
    >< Alw44I
    >< MaeI
    >< BmyI
    >< BsrI
    >< Alw21I
    >< AfaI
    DdeI >
TAGTGACCT ACTCTAGTGC CACAAGAGCA CTATGTGAGA ATTACTGGCT TGTACCCAAC ACTCAACATC
    16880      16890      16900      16910      16920      16930      16940

    StyI ><
    SinI >
    Sau96I >
    NspIV >
    EcoT14I ><
    Eco47I >
    Eco130I ><
    >< ScaI Cfr13I >
    BssT1I ><
    >< SphI >< RsaI Bsi2I >
    >< PaeI BsaJI ><
    >< NlaIII Bme18I >
    >< NspI >< Csp6I AvaII >
    >< NspHI >< AfaI AsuI >
TCAGATGAGT TTTCTAGCAA TGTGCAAAT TATCAAAAGG TCGGCATGCA AAAGTACTCT AACTCCAAG
    16950      16960      16970      16980      16990      17000      17010

    >< ScrFI
    >< RsaI
    >< MvaI
    >< EcoRII
    >< Ecl136I
    > < Csp6I
    >< BstOI
    >< BstNI
    >< XcmI
    >< BslI
    >< NspHII
    >< BsiYI
    >< BsiLI
    >< ApyI
    >< BsrI
    >< DsaV >< AfaI > < HinfI >< PleI
GACCACCTGG TACTGGTAAG AGTCATTTTG CCATCGGACT TGCTCTCTAT TACCCATCTG CTCGCATAGT
    17020      17030      17040      17050      17060      17070      17080

    >< SfaNI
    >< SphI
    >< PvuII
    >< PaeI
    >< Psp5I
    >< NspI
    >< NspBII
    >< NspHI
    >< Fnu4HI
    > < Tru9I
    >< Bst1107I
    > < NlaIII >< BspWI
    >< SspI
    >< AccI
    >< NlaIII
    >< AluI
    >< BbvI
    > < MseI
GTATACGGCA TGCTCTCATG CAGCTGTTGA TGCCCTATGT GAAAAGGCAT TAAAATATTT GCCCATAGAT
    17090      17100      17110      17120      17130      17140      17150

```

FIGURE 13.39

```

> < Thai
>< Thai
> < MvnI
>< MvnI >< Thai
> < HinPII
>< HinPII
>< HinPII >< MvnI
> < Hin6I
>< Hin6I
> < HhaI
>< HhaI >< HhaI
> < CfoI
>< CfoI >< CfoI
> < BstUI
>< BstUI >< BstUI
>< BssHII
>< BspMI
> < Bsp50I
>< Bsp50I>< Bsp50I
>< TfiI >< Hin6I> < AccII RmaI >
>< HinfI >< AccII >< AccII MaeI >
> < EcoRI
AAATGTAGTA GAATCATACC TGCGCGTGCG CGCGTAGAGT GTTTTGATAA ATTCAAAGTG AATTCAACAC
17160 17170 17180 17190 17200 17210 17220

>< Zsp2I
>< Ppu10I
>< NsiI
>< Mph1103I
>< EcoT22I
>< BsgI > < AvaIII >< DrdI
TAGAACAGTA TGT TTTCTGC ACTGTAAATG CATTGCCAGA AACAACTGCT GACATTGTAG TCTTTGATGA
17230 17240 17250 17260 17270 17280 17290

>< RmaI
>< MaeI >< MaeII
AATCTCTATG GCTACTAATT ATGACTTGAG TGTGTCAAT GCTAGACTTC GTGCAAAACA CTACGTCTAT
17300 17310 17320 17330 17340 17350 17360

>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< DpnI
>< BspAI >< RmaI
>< AlwI>< Bsp143I > < AciI >< MaeI SspI ><
ATTGGCGATC CTGCTCAATT ACCAGCCCCC CGCACATTGC TGAATAAAGG CACACTAGAA CCAGAAATATT
17370 17380 17390 17400 17410 17420 17430

>< SinI
>< Sau96I
>< NspIV >< StyI
>< NspHII >< NspI
>< Eco47I >< NspHI
>< Cfr13I >< NlaIII
>< BsiZI >< EcoT14I
>< BsgI >< Eco130I
>< Bme18I >< BssT1I
>< AvaII >< BsaJI
>< AsuI> < AflIII
TTAATTCAGT GTGCAGACTT ATGAAAACAA TAGGTCCAGA CATGTTCCCTT GGAAGTGTGTC GCCGTTGTCC
17440 17450 17460 17470 17480 17490 17500

```

FIGURE 13. 40

```

                >< HindII
                >< HincII
                >< AluI
TGCTGAAATT GTTGACACTG TGAGTGCTTT AGTTTATGAC AATAAGCTAA AAGCACACAA GGATAAGTCA
  17510      17520      17530      17540      17550      17560      17570

>< AluI
GCTCAATGCT TCAAAATGTT CTACAAAGGT GTTATTACAC ATGATGTTTC ATCTGCAATC AACAGACCTC
  17580      17590      17600      17610      17620      17630      17640

                >< MnlI
>< EcoNI
    >< BslI
    >< BsiYI
    >< HphI
    >< AluI
AAATAGGCGT TGTAAGAGAA TTTCTTACAC GCAATCCTGC TTGGAGAAAA GCTGTTTTTA TCTCACCTTA
  17650      17660      17670      17680      17690      17700      17710

                >< SfcI
                >< DdeI
                >< TfiI
                >< AluI
                >< BfrI
                >< HinfI
TAATTCACAG AACGCTGTAG CTTCAAAAAT CTTAGGATTG CCTACGCAGA CTGTTGATTG ATCACAGGGT
  17720      17730      17740      17750      17760      17770      17780

                >< Tth111I
                >< AspI
                >< HindII
                >< HincII
                >< AciI
TCTGAATATG ACTATGTCAT ATTACACAA ACTACTGAAA CAGCACACTC TTGTAATGTC AACCGCTTCA
  17790      17800      17810      17820      17830      17840      17850

                >< XhoII
                >< Sau3AI
                >< NdeII
                >< MflI
                >< MboI
                >< MamI
                >< DpnII
                >< DpnI
                >< BstYI
                >< BspAI
                >< Bsp143I
                >< BsiBI
                >< BsaBI
                >< BglII
                >< BspWI
ATGTGGCTAT CACAAGGGCA AAAATTGGCA TTTTGTGCAT AATGTCTGAT AGAGATCTTT ATGACAAACT
  17860      17870      17880      17890      17900      17910      17920

                >< XbaI
                >< RmaI
                >< MaeI
                >< MaeII
                >< MaeIII
                >< BsrI
GCAATTTACA AGTCTAGAAA TACCACGTCG CAATGTGGCT ACATTACAAG CAGAAAATGT AACTGGACTT
  17930      17940      17950      17960      17970      17980      17990

                >< Sau3AI
                >< NdeII
                >< MboII
                >< MboI
                >< FokI
                >< DpnII
                >< DpnI
                >< BspAI
                >< Bsp143I
                >< NlaIV
                >< Eco64I
                >< BscBI
                >< BanI
                >< MnlI
>< Tru9I
>< MseI>< SfcI
                >< BbsI
                >< BsrI
                >< AccBI
                >< DdeI

```

FIGURE 13. 41

## FIGURE 13.42

258

# EP 1 694 829 B1

```

TGTGACACT GAAAATAACA CAGAATTCAC CAGAGTTAAT GCAAAACCTC CACCAGGTGA CCAGTTTAAA
18350      18360      18370      18380      18390      18400      18410

      >< ScrFI
      >< MvaI
      >< EcoRII
      >< Ecl136I
      >< DsaV
      >< BstOI
      >< BstNI
      >< BsiLI
      >< BsaJI
      >< NlaIII
      >< ApyI
      >< Tru9I>< Csp6I
      >< MseI >< AfaI
      >< DdeI ><
CATCTTATAC CACTCATGTA TAAAGGCTTG CCCTGGAATG TAGTGCGTAT TAAGATAGTA CAAATGCTCA
18420      18430      18440      18450      18460      18470      18480

      >< NlaIII
      >< HinPII
      >< Tth111I
      >< HinfI
      >< AspI
      >< P1eI
      >< CfoI
      >< AluI
GTGATACACT GAAAGGATTG TCAGACAGAG TCGTGTTCGT CCTTTGGGCG CATGGCTTTG AGCTTACATC
18490      18500      18510      18520      18530      18540      18550

      >< SinI
      >< Sau96I
      >< NspIV
      >< NspHII
      >< Eco47I
      >< Cfr13I
      >< ScaI
      >< Bsi2I
      >< RsaI
      >< Bme18I
      >< Csp6I
      >< AvaII
      >< MaeII
      >< AfaI
      >< AsuI
      >< AflIII
      >< MaeIII>< MaeII
AATGAAGTAC TTTGTCAAGA TTGGACCTGA AAGAACGTGT TGTCTGTGTG ACAAACGTGC AACTTGCTTT
18560      18570      18580      18590      18600      18610      18620

      >< TfiI
      >< Tth111I
      >< HinfI
      >< AspI
TCTACTTCAT CAGATACTTA TGCCTGCTGG AATCATTCTG TGGGTTTTGA CTATGTCTAT AACCCATTTA
18630      18640      18650      18660      18670      18680      18690

      >< ScrFI
      RsaI ><
      >< MvaI
      >< EcoRII
      Ecl136I ><
      >< DsaV
      Csp6I ><
      BstXI ><
      >< MaeIII
      >< EcoO65I
      >< Eco91I
      >< BstPI
      >< BstEII
      >< MaeIII >< NlaIII
      >< AfaI ><
TGATTGATGT TCAGCAGTGG GGCTTTACGG GTAACCTTCA GAGTAACCAT GACCAACATT GCCAGGTACA
18700      18710      18720      18730      18740      18750      18760

      >< SfaNI
      >< RmaI
      >< NspI
      >< NspHI

```

FIGURE 13.43

```

                >< NlaIII                >< RmaI
                >< MaeI                >< NlaIII                Tru9I ><
    >< NlaIII    >< BspWI                >< MaeI                >< NlaIII
                > < AflIII                >< BspHI                MseI ><
TGGAAATGCA CATGTGGCTA GTTGTGATGC TATCATGACT AGATGTTTAG CAGTCCATGA GTGCTTTGTT
    18770      18780      18790      18800      18810      18820      18830

    >< ThaI
    >< MvnI
    >< HinPII
    >< Hin6I
    >< HhaI
    >< CfoI
    >< BstUI
    >< Bsp50I                >< EcoNI> < MnlI
    >< AccII                >< BslI                >< Tru9I
                >< BsiYI                >< DdeI >< MseI
AAGCGCGTTG ATTGGTCTGT TGAATACCCT ATTATAGGAG ATGAAGTGA GGTAAATTCT GCTTGCAGAA
    18840      18850      18860      18870      18880      18890      18900

    >< RsaI
    >< Csp6I
    >< AfaI    >< NlaIII                >< BspWI                >< MboII                > < NlaIII
                >< BsrI    >< BspHI
AAGTACAACA CATGGTTGTG AAGTCTGCAT TGCTTGCTGA TAAGTTTCCA GTTCTTCATG ACATTGGAAA
    18910      18920      18930      18940      18950      18960      18970

                >< SauI
                >< MstII
                >< Eco81I
                >< DdeI                NlaIII ><
                >< CvnI                >< EspI
                >< Bsu36I                >< Eco57I MaeIII ><
                >< Bse21I                >< DdeI
                >< AxyI                >< CelII
                >< AocI    >< MnlI    >< SfaNI                >< Bpu1102I
TCCAAAGGCT ATCAAGTGTG TGCCTCAGGC TGAAGTAGAA TGGAAGTTCT ACGATGCTCA GCCATGTAGT
    18980      18990      19000      19010      19020      19030      19040

                >< MnlI                >< Ksp632I
    >< HindIII                >< EarI
    >< AluI    >< MboII                >< Eam1104I
GACAAAGCTT ACAAATAGA GGAAGTCTTC TATTCTTATG CTACACATCA CGATAAATTC ACTGATGGTG
    19050      19060      19070      19080      19090      19100      19110

                >< Sau3AI
                >< NdeII
                >< MboI
    >< MaeII> < MaeIII
                >< DpnII
                >< DpnI
                >< BspAI                HinFI >
                >< MaeIII >< Bsp143I                >< MunI                DrdI ><
TTTGTTTGTT TTGGAATTGT AACGTTGATC GTTACCCAGC CAATGCAATT GTGTGTAGGT TTGACACAAG
    19120      19130      19140      19150      19160      19170      19180

                Zsp2I ><
                >< SphI
                > < Ppu10I
                >< PaeI
                >< NspI
                >< NspHI
                >< NlaIII
    >< ScrFI
    >< MvaI
    >< EcoRII                Mph1103I ><

```

FIGURE 1344



# EP 1 694 829 B1

```

                >< Ecl136I
                >< DsaV
                >< BstOI
                >< BstNI
                >< BsiLI
                >< ApyI
                >< PleI
AGTCTTGTCA AACTTGAAC TACCAGGCTG TGATGGTGGT AGTTTGTATG TGAATAAGCA TGCATTCCAC
19190      19200      19210      19220      19230      19240      19250

                >< Tru9I
                > < MunI
                >< TthHB8I
                >< MseI
>< BcgI/a >< TaqI
                >< DraI
                >< AluI
                >< BcgI
ACTCCAGCTT TCGATAAAAG TGCATTTACT AATTAAAGC AATTGCCTTT CTTTTACTAT TCTGATAGTC
19260      19270      19280      19290      19300      19310      19320

                >< PleI
                >< NlaIII
                >< BsmAI
                >< HinfI>< Alw26I
                SfaNI ><
                >< MaeII
                BsaAI ><
                AflIII ><
CTTGTGAGTC TCATGGCAA CAAGTAGTGT CCGATATTGA TTATGTTCCA CTCAAATCTG CTACGTGTAT
19330      19340      19350      19360      19370      19380      19390

                Zsp2I >
                >< ScaI
                Ppu10I ><
                >< RsaINsiI >
                Mph1103I >
                >< SfaNEcoT22I >
                > < RsaI >< Csp6I
                >< Csp6I
                >< NlaIII> < AfaI >< AfaI
                AvaIII ><
TACACGATGC AATTTAGGTG GTGCTGTTTG CAGACACCAT GCAAATGAGT ACCGACAGTA CTTGGATGCA
19400      19410      19420      19430      19440      19450      19460

                >< FokI
TATAATATGA TGATTTCTGC TGGATTTAGC CTATGGATTT ACAAACAATT TGATACTTAT AACCTGTGGA
19470      19480      19490      19500      19510      19520      19530

                >< ScrFI
                >< MvaI
                >< MaeIII
                >< EcoRII
                >< Ecl136I
                >< DsaV
                >< BstOI
                >< BstNI
                >< BsiLI
                >< ApyI
                >< Tru9I
                >< MseI
ATACATTTAC CAGGTTACAG AGTTTAGAAA ATGTGGCTTA TAATGTTGTT AATAAAGGAC ACTTTGATGG
19540      19550      19560      19570      19580      19590      19600

                >< SgrAI
                >< NaeI
                >< MspI
                >< HpaII
                >< HapII
                >< Cfr10I
                >< BspWI
                > < VspI
                > < Tru9I
                > < MseI
                > < AsnI
                > < AseI
ACACGCCGGC GAAGCACCTG TTTCCATCAT TAATAATGCT GTTTACACAA AGGTAGATGG TATTGATGTG
19610      19620      19630      19640      19650      19660      19670

```

FIGURE 13. 45

# EP 1 694 829 B1

```

>< XhoII
>< Sau3AI
>< NdeII
>< MflI
>< MboI
>< DpnII
  >< DpnI
  >< MaeIII
>< BstYI
  >< EspI
>< BspAI
  >< DdeITru9I ><
  >< CelIIMseI ><
  >< Bsp143I
    >< Tru9I
    >< MseI
    >< AluI
    >< Bpu1102I
  >< BglIII
GAGATCTTTG AAAATAAGAC AACACTTCCT GTTAATGTTG CATTTGAGCT TTGGGCTAAG CGTAACATTA
19680      19690      19700      19710      19720      19730      19740

                                >< Fnu4HI
  >< BsrI
    >< Tru9I
    >< MseI
    >< BbvI
    >< EcoRV
    >< Eco32I
AACCAGTGCC AGAGATTAAG ATACTCAATA ATTTGGGTGT TGATATCGCT GCTAATACTG TAATCTGGGA
19750      19760      19770      19780      19790      19800      19810

  >< NspI
  >< NspHI
  >< NlaIII
  >< BsgI
    >< AflIII
CTACAAAAGA GAAGCCCCAG CACATGTATC TACAATAGGT GTCTGCACAA TGA CTGACAT TGCCAAGAAA
19820      19830      19840      19850      19860      19870      19880

  >< DdeI>< MboII
CCTACTGAGA GTGCTTGTTT TCACTTACT GTCTTGTTTG ATGGTAGAGT GGAAGGACAG GTAGACCTTT
19890      19900      19910      19920      19930      19940      19950

                                SinI ><
                                Sau96I ><
                                NspIV ><
                                NspHII ><
                                NlaIV ><
                                Eco47I ><
                                Cfr13I ><
                                >< BslI
                                Bsi2I ><
                                >< BsiYI
                                BscBI ><
                                Bmel8I ><
                                AvaII ><
                                AsuI ><
  >< Tru9I
  >< MseI
TTAGAAACGC CCGTAATGGT GTTTTAATAA CAGAAGGTTT AGTCAAAGGT CTAACACCTT CAAAGGGACC
19960      19970      19980      19990      20000      20010      20020

  >< VspI
  >< Tru9I
  >< PleI
  >< MseI
    >< RmaI
    >< NheI
    >< MaeI
    >< AsnI
    >< TfiI
    >< HgaI>< AluI
    >< HinfI>< AseI
    >< HinfI
    >< MseI
AGCACAAAGCT AGCGTCAATG GAGTCACATT AATTGGAGAA TCAGTAAAAA CACAGTTTAA CTACTTTAAG
20030      20040      20050      20060      20070      20080      20090

                                >< DdeI
                                >< MnlI
                                Tru9I ><
                                >< BsmAI
                                >< DdeI

```

FIGURE 1346

# EP 1 694 829 B1

```

>< AccI                                     >< Alw26I >< BfrIMseI ><
AAAGTAGACG GCATTATTCA ACAGTTGCCT GAAACCTACT TTA CTCAGAG CAGAGACTTA GAGGATTTTA
20100      20110      20120      20130      20140      20150      20160

>< TthHB8I
>< TaqI
>< SstI
>< SduI
>< SacI
> < PaeR7I
> < NspIII
>< NspII
>< HgiAI
> < Eco88I
> < XhoI>< Eco24I
>< Ecl136II
>< XcmI
>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< DpnI
>< BspAI
>< Bsp143I
AGCCCAGATC ACAAATGGAA ACTGACTTTC TCGAGCTCGC TATGGATGAA TTCATACAGC GATATAAGCT
20170      20180      20190      20200      20210      20220      20230

>< TthHB8I
>< TaqI
>< SfuI
>< NspV
>< LspI
>< Csp45I
>< BstBI
>< Bsp119I
>< BsiCI
>< Bpu14I
>< AsuII >< BcgI
CGAGGGCTAT GCCTTCGAAC ACATCGTTTA TGGAGATTTC AGTCATGGAC AACTTGGCGG TCTTCATTTA
20240      20250      20260      20270      20280      20290      20300

>< HphI
>< HinPII
>< Hin6I
>< EspI > < HhaI >< TfiI
>< DdeI >< HaeII
>< CelII >< Eco47III >< Tru9I
>< Bpu1102I > < CfoI >< HinfI >< MseI
>< BfrI >< Bsp143II >< MnlI
ATGATAGGCT TAGCCAAGCG CTCACAAGAT TCACCACTTA AATTAGAGGA TTTTATCCCT ATGGACAGCA
20310      20320      20330      20340      20350      20360      20370

>< MstI
>< HinPII
>< Hin6I
>< HhaI
>< FspI
>< FdiII
>< CfoI
>< SfaNI >< AviII
CAGTGAAAAA TTA CTCATA ACAGATGCGC AAACAGGTTC ATCAAAATGT GTGTGTTCTG TGATTGATCT
20380      20390      20400      20410      20420      20430      20440

>< TthHB8I

```

FIGURE 13.47

# EP 1 694 829 B1

```

>< Tth111I
  >< TaqI
    >< AspI          > < MaeIII          MaeIII ><
TTTACTTGAT GACTTTGTCG AGATAATAAA GTCACAAGAT TTGTCACTGA TTTCAAAAGT GGTCAAGGTT
20450      20460      20470      20480      20490      20500      20510

                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< FokI

>< MunI          > < NlaIII          >< AflIII
ACAATTGACT ATGCTGAAAT TTCATTCATG CTTTGGTGTA AGGATGGACA TGTTGAAACC TTCTACCCAA
20520      20530      20540      20550      20560      20570      20580

                                >< SfaNI
                                >< ScrFI
                                >< MvaI
                                >< EcoRII
                                >< Ecl136I
                                >< DsaV
                                >< BstOI          >< SfaNI
                                >< BstNI          >< RsaI   BspWI ><
                                >< BsiLI          > < Csp6I   BsmI >
                                >< BspWI          >< ApyI          >< AfaI   BscCI ><
AACTACAAGC AAGTCAAGCG TGGCAACCAG GTGTTGCGAT GCCTAACTTG TACAAGATGC AAAGAATGCT
20590      20600      20610      20620      20630      20640      20650

>< Eco57I >< MaeIII          >< HphI
TCTTGAAAAG TGTGACCTTC AGAATTATGG TGAAAATGCT GTTATACCAA AAGGAATAAT GATGAATGTC
20660      20670      20680      20690      20700      20710      20720

                                > < RsaI
                                >< Csp6I

>< Bst1107I          >< Tru9I          >< AluI
>< AccI          >< MseI          > < AfaINlaIII ><
GCAAAGTATA CTCAACTGTG TCAATACTTA AATACACTTA CTTTAGCTGT ACCCTACAAC ATGAGAGTTA
20730      20740      20750      20760      20770      20780      20790

                                >< ScrFI
                                >< RsaI
                                >< MvaI
                                >< EcoRII >< NspBII
                                >< Ecl136I          >< SduI
                                > < Csp6I          >< NspII
                                >< BstOI >< PvuII>< HgiAI
                                >< BstNI          >< DdeI
                                >< BsiLI >< Psp5I>< Bsp1286I
                                >< ApyI >< AluI >< BmyI
                                >< DsaV>< AfaI          >< Alw21I
TTCACTTTGG TGCTGGCTCT GATAAAGGAG TTGCACCAGG TACAGCTGTG CTCAGACAAT GGTGCCCAAC
20800      20810      20820      20830      20840      20850      20860

                                >< XhoII
                                >< Tru9I
                                >< Sau3AI
                                >< NdeII
>< TthHB8I >< MseI
>< MflI
>< MboI
>< MamI
>< DpnII
>< TfiI >< DpnI

```

FIGURE 13. 48

# EP 1 694 829 B1

```

>< BstYI          > < TfiI
>< BspAI          > < HinfI
>< HinfI>< Bsp143I    >< Esp3I      >< Tru9I
>< BsiBI          >< Tth111I  >< BsmBI      >< MseI
>< BsaBI          >< BsmAI      > < BsmAI
>< BsrI          >< TaqI >< BglII  >< AspI      >< Alw26I >< HgaI> < Alw26I
TGGCACACTA CTTGTCTGATT CAGATCTTAA TGACTTCGTC TCCGACGCAG ATTCTACTTT AATTGGAGAC
20870      20880      20890      20900      20910      20920      20930

>< StyI
>< SinI
>< Sau96I
> < SinI
> < Sau96I
>< PssI
>< Psp5II
> < PpuMI
> < NspIV
>< NspHII
>< NlaIV
> < EcoO109I
> < Eco47I
> < DraII
> < Cfr13I
> < Bsi2I
>< BscBI
>< RsaI
> < Csp6I
>< AfaI
>< Bme18I
> < AvaII
> < AsuI
>< BlnI
>< AvrII
>< AvaII
>< AsuI
AflIII ><
TGTGCAACAG TACATACGGC TAATAAATGG GACCTTATTA TTAGCGATAT GTATGACCCT AGGACCAAAC
20940      20950      20960      20970      20980      20990      21000

>< NspI
>< NspHI
>< NlaIII >< PleI
>< MaeIII >< HinfI
ATGTGACAAA AGAGAATGAC TCTAAAGAAG GGTTTTTCAC TTATCTGTGT GGATTATATAA AGCAAAAAC
21010      21020      21030      21040      21050      21060      21070

>< ScrFI
>< MvaI
>< EcoRII
>< Ecl136I
>< DsaV
>< BstOI
>< BstNI
>< BsiLI
>< BsaJI
>< BsaJI >< SfcI
>< ApyI > < AluI >< BscCI
>< BscCIHindIII ><> < AluI
AGCCCTGGGT GGTCTATAG CTGTAAAGAT AACAGAGCAT TCTTGAATG CTGACCTTTA CAAGCTTATG
21080      21090      21100      21110      21120      21130      21140

>< Zsp2I
>< Ppu10I
>< NsiI
>< Mph1103I Tru9I ><
>< Eco22I
>< BshI >< NlaIII>< AluI >< BcgI >< AvaIII >< SfaNIBcgI/a ><
GGCCATTTCT CATGGTGGAC AGCTTTTGT ACAAATGTAA ATGCATCATC ATCGGAAGCA TTTTAAATTG
21150      21160      21170      21180      21190      21200      21210

```

FIGURE 13.49

```

                                >< Zsp2I
                                >< SphI
                                >< Ppu10I
                                >< PaeI
                                >< NspI
                                >< NspHI
                                >< NsiI
                                >< NlaIII
                                > < NlaIII
                                >< Mph1103I
                                >< EcoT22I
                                > < AvaIII >< MnlI
GGGCTAACTA TCTTGGCAAG CCGAAGGAAC AAATTGATGG CTATACCATG CATGCTAACT ACATTTTCTG
    21220      21230      21240      21250      21260      21270      21280

                                Tru9I ><
                                >< Tru9I
                                >< GsuI      MseI ><
                                >< BsrI      >< MseI
                                >< BpmI      MnlI ><
                                >< BbsI      >< NlaIII >< MnlI
GAGGAACACA AATCCTATCC AGTTGTCTTC CTATTCACCTC TTTGACATGA GCAAATTTCC TCTTAAATTA
    21290      21300      21310      21320      21330      21340      21350

                                >< Tru9I
                                >< MseI
                                >< Esp4I> < TfiI
                                >< BsmAI
                                >< Alw26I
                                >< AflIII> < HinfI
AGAGGAACTG CTGTAATGTC TCTTAAGGAG AATCAAATCA ATGATATGAT TTATTCTCTT CTGGAAAAAG
    21360      21370      21380      21390      21400      21410      21420

                                >< Tru9I
                                >< MseI
                                >< HindII
                                >< HincII
                                >< HpaI AflIII >
GTAGGCTTAT CATTAGAGAA AACAACAGAG TTGTGGTTTC AAGTGATATT CTTGTTAACA ACTAAACGAA
    21430      21440      21450      21460      21470      21480      21490

                                >< VneI
                                >< SnuI
                                >< SduI
                                >< NspII
                                >< HpaII
                                >< HgiAI
                                >< HapII
                                >< Cfr10I
                                >< Bsp1286I
                                >< MspI>< BmyI
                                >< ApaLI
                                >< Alw44I
                                >< AgeI >< Alw21I
                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< SpeI
                                >< RmaI
                                >< MaeI >< MaeIII >< AgeI >< Alw21I
CATGTTTATT TTCTTATTAT TTCTTACTCT CACTAGTGGT AGTGACCTTG ACCGGTGCAC CACTTTTGTAT
    21500      21510      21520      21530      21540      21550      21560

                                > < AluI
                                >< MnlI
GATGTTCAAG CTCCTAATTA CACTCAACAT ACTTCATCTA TGAGGGGGGT TTACTATCCT GATGAAATTT
    21570      21580      21590      21600      21610      21620      21630

                                >< Sau3AI

```

FIGURE 13. 50

```

>< NdeII
>< MboI
>< DpnII
  >< DpnI          >< Tru9I
>< BspAI          >< MseI > < MboII
  >< Bsp143I       >< DdeI          >< MaeIII
TTAGATCAGA CACTCTTTAT TTAAGTCAGG ATTTATTTCT TCCATTTTAT TCTAATGTTA CAGGGTTTCA
  21640      21650      21660      21670      21680      21690      21700

  >< VspI
  >< Tru9I
  >< MseI
  >< AsnI
  >< AseI >< MaeII          >< Tru9I          >< FokI
  >< MseI >< BbvI          >< Fnu4HI
TACTATTAAT CATACGTTTG GCAACCCTGT CATACCTTTT AAGGATGGTA TTTATTTTGC TGCCACAGAG
  21710      21720      21730      21740      21750      21760      21770

          >< BslI
          >< DsaI>< BsiYI          >< NlaIII
          >< BsaJI          >< MaeIII
AAATCAAATG TTGTCCGTGG TTGGGTTTTT GGTTCCTACCA TGAACAACAA GTCACAGTCG GTGATTATTA
  21780      21790      21800      21810      21820      21830      21840

          >< NspI
          >< NspHI
          >< NlaIII
          >< MaeIII
          >< MaeIII
TTAACAATTC TACTAATGTT GTTATACGAG CATGTAACCT TGAATTGTGT GACAACCCCT TCTTTGCTGT
  21850      21860      21870      21880      21890      21900      21910

  >< StyI          >< Zsp2I
  >< NlaIII          >< Tru9I
  >< NcoI >< RsaI          >< Ppu10I TthHB8I ><
  >< EcoT14I          >< NsiI          >< TaqI
  >< Eco130I          >< MseI          SfaNI ><
  >< DsaI>< Csp6I          >< Mph1103I RsaI ><
  >< BssT1I          >< TthHB8I >< EcoT22I Csp6I ><
  >< BsaJI>< AfaI          >< TaqI >< AvaIII AfaI ><
TTCTAAACCC ATGGGTACAC AGACACATAC TATGATATTC GATAATGCAT TTAATTGCAC TTTCGAGTAC
  21920      21930      21940      21950      21960      21970      21980

          >< Tru9I
          >< MseI
          >< DraI
ATATCTGATG CCTTTTCGCT TGATGTTTCA GAAAAGTCAG GTAATTTTAA AACTTACGA GAGTTTGTGT
  21990      22000      22010      22020      22030      22040      22050

          >< Sau3AI
          >< NdeII
          >< MboI
          >< DpnII
          >< DpnI
          >< BspAI
>< Tru9I
>< MseI
  >< DraI          >< SfcI Bsp143I ><
TTAAAAATAA AGATGGGTTT CTCTATGTTT ATAAGGGCTA TCAACCTATA GATGTAGTTC GTGATCTACC
  22060      22070      22080      22090      22100      22110      22120

          >< Tru9I
          >< Tru9I          >< MseI
          >< MseI          >< MnlI
TTCTGGTTTT AACACTTTGA AACCTATTTT TAAGTTGCCT CTTGGTATTA ACATTACAAA TTTTAGAGCC
  22130      22140      22150      22160      22170      22180      22190

```

FIGURE 13.51

```

> < SduI>< SfcI
>< PvuII
>< Psp5I
> < NspII
>< NspBII
> < MaeII > < Fnu4HI
> < Bsp1286I >< PstI Tru9I >
>< BspMI >< BmyI>< Fnu4HI MseI >
>< HphI >< BbvI >< AluI >< BbvI
ATTCTTACAG CCTTTTCACC TGCTCAAGAC ATTTGGGGCA CGTCAGCTGC AGCCTATTTT GTTGGCTATT
22200 22210 22220 22230 22240 22250 22260

>< SfaNI
>< RsaI
> < Csp6I
>< DraI >< AfaI >< AlwNI
TAAAGCCAAC TACATTTATG CTCAAGTATG ATGAAAATGG TACAATCACA GATGCTGTTG ATTGTTCTCA
22270 22280 22290 22300 22310 22320 22330

> < Tru9I
> < MseI
>< AluI
AAATCCACTT GCTGAACCTCA AATGCTCTGT TAAGAGCTTT GAGATTGACA AAGGAATTTA CCAGACCTCT
22340 22350 22360 22370 22380 22390 22400

>< SauI
>< MstII
>< Eco81I
>< DdeI
>< CvnI
>< Bsu36I
>< Bse21I
>< AxyI >< TfiI
>< MnlI >< AocI >< MnlI >< HinfI >< SspI >< MnlI
AATTTTCAGGG TTGTTCCCTC AGGAGATGTT GTGAGATTCC CTAATATTAC AAAGTTGTGT CTTTGGGAG
22410 22420 22430 22440 22450 22460 22470

>< Zsp2I
>< Ppu10I
>< NsiI
> < NlaIII
>< Mph1103I
>< EcoT22I
>< Tru9I
>< MseI >< AvaIII
AGGTTTTTAA TGCTACTAAA TTCCCTTCTG TCTATGCATG GGAGAGAAAA AAAATTTCTA ATTGTGTTGC
22480 22490 22500 22510 22520 22530 22540

>< SduI
>< NspII
>< HgiAI
>< Bsp1286I
>< BmyI >< Tru9I
>< Alw21I >< MseI DdeI >>
TGATTACTCT GTGCTCTACA ACTCAACATT TTTTCAACC TTAAAGTGCT ATGGCGTTTC TGCCACTAAG
22550 22560 22570 22580 22590 22600 22610

>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< DpnI

```

FIGURE 13.52



# EP 1 694 829 B1

```

    >< BspAI
    >< Bsp143I
    TTGAATGATC TTTGCTTCTC CAATGTCTAT GCAGATTCTT TTGTAGTCAA GGGAGATGAT GTAAGACAAA
    22620      22630      22640      22650      22660      22670      22680

    >< ScrFI
    >< MvaI
    >< HinPII
    >< Hin6I
    >< HhaI
    >< HaeII
    >< EcoRII
    >< Ecl136I
    >< DsaV
    >< CfoI
    >< BstOI
    >< BstNI
    >< Bsp143II
    >< BsiLI
    >< ApyI
    >< BsrI
    TAGCGCCAGG ACAAACTGGT GTTATTGCTG ATTATAATTA TAAATTGCCA GATGATTTCa TGGGTTGTGT
    22690      22700      22710      22720      22730      22740      22750

    >< SfaNI
    >< RmaI
    >< MaeI
    CCTTGCTTGG AATACTAGGA ACATTGATGC TACTTCAACT GGTAATTATA ATTATAAATA TAGGTATCTT
    22760      22770      22780      22790      22800      22810      22820

    >< Sau96I
    >< Pali
    >< NspIV
    >< HindIII
    >< HaeIII
    >< Eco0109I
    >< DraII
    >< DdeI
    >< Cfr13I
    >< BsuRI
    >< BsiZI
    >< BshI
    >< BfrI >< PssI
    >< NlaIII >< AsuI>< BsmAI
    >< AluI >< Alw26I
    AGACATGGCA AGCTTAGGCC CTTTGAGAGA GACATATCTA ATGTGCCTTT CTCCCCTGAT GGCAAACCTT
    22830      22840      22850      22860      22870      22880      22890

    >< Tru9I
    >< Pali
    >< MscI
    >< HaeIII
    >< EaeI>< MseI
    >< Tru9I >< BsuRI
    >< MseI >< BshI
    >< BspMI >< BalI
    GCACCCACCC TGCTCTTAAT TGTTATTGGC CATTAAATGA TTATGGTTTT TACACCACTA CTGGCATTGG
    22900      22910      22920      22930      22940      22950      22960

    Sau96I ><
    >< PalINspIV ><
    >< MspI NspHII ><
    >< HaeIII

```

FIGURE 13.53

```

                                > < HpaII Eco47I ><
                                >< DsaI
                                > < HapII Cfr13I ><
                                >< BsuRISinI ><
                                >< GdiII Bsi2I ><
                                >< BsaJI
                                >< Tru9I >< EaeI Bme18I ><
                                >< MseI >< Cfr10I AvaII ><
                                >< DraI >< BshI AsuI ><
CTACCAACCT TACAGAGTTG TAGTACTTTC TTTTGAACCT TTAATGCAC CGGCCACGGT TTGTGGACCA
  22970      22980      22990      23000      23010      23020      23030

                                >< Tru9I >< RsaI
                                >< Tru9I >< Csp6I
                                >< PleI >< BsrI ><
                                >< MseI >< BsrI
                                > < Tru9I
                                > < MseI>< BsrI
                                >< MseI >< HinfI >< AfaI
AAATTATCCA CTGACCTTAT TAAGAACCAG TGTGTCAATT TTAATTTTAA TGGACTCACT GGTACTGGTG
  23040      23050      23060      23070      23080      23090      23100

>< Tru9I >< Pali
>< MseI >< HaeIII
>< MboII >< GdiII
>< HpaI >< EaeI
>< HindII >< BsuRI TffiI ><
>< HincII >< BshI HinfI ><
TGTTAACCTCC TTCTTCAAAG AGATTTCAAC CATTCAACA ATTTGGCCGT GATGTTTCTG ATTTCACTGA
  23110      23120      23130      23140      23150      23160      23170

> < XhoII
>< TthHB8I
>< TaqI
> < Sau3AI
> < NdeII
> < MflI
> < MboI
> < DpnII
>< DpnI
> < BstYI
> < BspAI > < SspI
>< AlwI >< Bsp143I >< HphI
TTCCGTTTCA GATCCTAAAA CATCTGAAAT ATTAGACATT TCACCTTGCT CTTTGGGGG TGTAAGTGTA
  23180      23190      23200      23210      23220      23230      23240

>< ScrFI
>< MvaI
>< EcoRII
>< Ecl136I >< Tru9I
>< DsaV >< MseI
>< BstOI >< HpaI
>< BstNI >< HindII
>< BsiLI >< Eco57I
>< ApyI >< BsgI >< HincII
ATTACACCTG GAACAAATGC TTCATCTGAA GTTGCTGTTC TATATCAAGA TGTTAACTGC ACTGATGTTT
  23250      23260      23270      23280      23290      23300      23310

>< Sau3AI
>< NlaIII
>< NdeII
>< MboI
>< DpnII
>< DpnI >< HinfII

```

FIGURE 13. 54

# EP 1 694 829 B1

```

>< BspWI                >< Hin6I
>< BspAI                > < HhaI                PleI ><
>< SfcI                >< Bsp143I        >< AluI> < CfoI                >< BsrI
CTACAGCAAT TCATGCAGAT CAACTCACAC CAGCTTGGCG CATATATTCT ACTGGAAACA ATGTATTCCA
23320      23330      23340      23350      23360      23370      23380

>< TthHB8I
>< TaqI
>< SalI
>< RtrI
>< NspI
>< EspI >< NspHI
>< DdeI >< NlaIII
>< CelII >< HindII
>< Bpu102I>< HincII
>< HinfI                >< AluI                >< AccI
GACTCAAGCA GGCTGTCTTA TAGGAGCTGA GCATGTCGAC ACTTCTTATG AGTGGGACAT TCCTATTGGA
23390      23400      23410      23420      23430      23440      23450

> < SnaBI
>< ScaI
>< RsaI
>< RmaI
>< MaeII >< MaeI
> < Eco105I
>< Csp6I
>< RmaI
>< MaeIII
> < BsaAI
>< AfaI
>< AluI                >< MaeI
GCTGGCATTG GTGCTAGTTA CCATACAGTT TCTTTATTAC GTAGTACTAG CCAAAAATCT ATTGTGGCTT
23460      23470      23480      23490      23500      23510      23520

>< MunI
ATACTATGTC TTTAGGTGCT GATAGTTCAA TTGCTTACTC TAATAACACC ATTGCTATAC CTACTAACTT
23530      23540      23550      23560      23570      23580      23590

RsaI ><
>< MnlI
Csp6I ><
AfaI ><
>< SfcI
TTCAATTAGC ATTACTACAG AAGTAATGCC TGTTTCTATG GCTAAACCT CCGTAGATTG TAATATGTAC
23600      23610      23620      23630      23640      23650      23660

> < TfiI
> < HinfI
>< AciI                > < AluI
ATCTGCGGAG ATTCTACTGA ATGTGCTAAT TTGCTTCTCC AATATGGTAG CTTTTCGACA CAACTAAATC
23670      23680      23690      23700      23710      23720      23730

>< VneI
>< SduI
>< NspII
>< HgiAI
>< SnoI>< DdeI                >< Sau3AI                >< PmlI
>< Bsp1286I                >< NdeII                >< PmaCI
>< BmyI                >< MboI                >< MaeII
>< BbvI                >< DpnI                >< Eco72I
>< ApaLI                >< Bsp143I                >< BsaAI
>< Alw44I                >< BbrPI
>< Alw21I                >< DpnII >< AlwI
>< Alw21I                >< Fnu4HI                >< BspAI                >< AflIII
GTGCACTCTC AGGTATTGCT GCTGAACAGG ATCGCAACAC ACGTGAAGTG TTCGCTCAAG TCAAACAAAT
23740      23750      23760      23770      23780      23790      23800

```

FIGURE 13.55

EP 1 694 829 B1

```

>< RsaI
>< Csp6I
>< AfaI
GTACAAAACC CCAACTTTGA AATATTTTGG TGGTTTTAAT TTTTCACAAA TATTACCTGA CCCTCTAAAG
23810 23820 23830 23840 23850 23860 23870

>< MnlI
>< MnlI
>< DdeI >< MnlI
>< Tru9I >< SfaNI >< HphI NlaIII ><
>< MseI >< MaeIII BspHI ><
CCAACATAAGA GGTCTTTTAT TGAGGACTTG CTCTTTAATA AGGTGACACT CGCTGATGCT GGCTTCATGA
23880 23890 23900 23910 23920 23930 23940

>< XhoII
>< Sau3AI
>< StyI
>< RmaI
>< MaeI
>< EcoT14I
>< Eco130I
>< BssT1I
>< BsmI
>< BscCI
>< BsaJI
>< BlnI
>< AvrII
>< VspI
>< Tru9I
>< MseI
>< AsnI
>< AseI
>< RmaI
>< NdeII
>< MflI
>< MboI
>< MaeI
>< DpnII
>< HphI> < DpnI
>< BstYI
>< BspAI
>< Bsp143I
>< BglII
>< MstI
>< HinPII
>< Hin6I
>< HhaI
>< FspI
>< FdiII
>< CfoI
>< AviII
AGCAATATGG CGAATGCCTA GGTGATATTA ATGCTAGAGA TCTCATTTGT GCGCAGAAGT TCAATGGACT
23950 23960 23970 23980 23990 24000 24010

>< RmaIRsaI ><
>< MnlI
>< Fnu4HI
>< Fnu4HI Csp6I ><
>< BspWI >< BbvI
>< BbvI >< BspWI >< MaeIAfaI ><
TACAGTGTGG CCACCTCTGC TCACTGATGA TATGATTGCT GCCTACACTG CTGCTCTAGT TAGTGGTACT
24020 24030 24040 24050 24060 24070 24080

>< MboII
>< HinPII
>< Hin6I
>< HhaI
>< HaeII
>< Fnu4HI >< Ksp632I
>< CfoI >< EarI
>< FokI >< BspWI >< Eam1104I
>< BbvI
>< Bsp143II
GCCACTGCTG GATGGACATT TGGTGCTGGC GCTGCTCTTC AAATACCTTT TGCTATGCAA ATGGCATATA
24090 24100 24110 24120 24130 24140 24150

>< MaeIII
Tru9I ><
MseI ><
GGTTCAATGG CATTTGAGTT ACCCAAAATG TTCTCTATGA GAACCAAAAA CAAATCGCCA ACCAATTTAA
24160 24170 24180 24190 24200 24210 24220

>< MaeII ><
>< TfiI
>< HinfI
>< Fnu4HI
>< BbvI
>< AluI
CAAGGCGATT AGTCAAATTC AAGAATCACT TACAACAACA TCAACTGCAT TGGGCAAGCT GCAAGACGTT
24230 24240 24250 24260 24270 24280 24290

>< Tru9I
>< MseI
>< HpaI
>< HindII
>< HincII><
>< BsmI >< Tru9I
>< BsmI >< Tru9I >< BfrI
>< MseI >< AluI

```

FIGURE 13. 56

EP 1 694 829 B1

```

GTTAACCAGA ATGCTCAAGC ATTAAACACA CTTGTTAAAC AACTTAGCTC TAATTTTGGT GCAATTTCAA
  24300      24310      24320      24330      24340      24350      24360

      >< ThaI
      >< SpoI
      >< NruI
      >< MvnI
      >< BstUI      >< TthHB8I
      >< Bsp68I      >< TaqI      >< RsaI
      >< EcoRV      >< Bsp50I      >< MnlI      >< Csp6I      >< Tru9I
      >< Eco32I      >< AccII      >< MnlI      >< AciI      >< AfaI      >< MseI
GTGTGCTAAA TGATATCCTT TCGCGACTTG ATAAAGTCGA GGCGGAGGTA CAAATTGACA GGTTAATTAC
  24370      24380      24390      24400      24410      24420      24430

      >< MaeIII      >< BbvI      >< Fnu4HI      BbvI ><
AGGCAGACTT CAAAGCCTTC AAACCTATGT AACACAACAA CTAATCAGGG CTGCTGAAAT CAGGGCTTCT
  24440      24450      24460      24470      24480      24490      24500

      >< Fnu4HI      >< HindII
      >< BspWI      >< DdeI      >< HincII
GCTAATCTTG CTGCTACTAA AATGTCTGAG TGTGTTCTTG GACAATCAAA AAGAGTTGAC TTTTGTGGAA
  24510      24520      24530      24540      24550      24560      24570

      > < NspI
      > < NspHI
      > < NlaIII
      >< MaeIII
      >< NlaIII      >< MaeII
      >< MboII      >< FokI
      >< Fnu4HI      >< BbsI      BsaAI ><
      >< AciI      >< BbvI      >< AflIII
AGGGCTACCA CCTTATGTCC TTCCACAAG CAGCCCCGCA TGGTGTGTC TTCCTACATG TCACGTATGT
  24580      24590      24600      24610      24620      24630      24640

      >< ScrFI
      >< MvaI
      >< EcoRII
      >< Ecl136I
      >< BstOI
      >< BstNI      >< HinPII
      >< MnlI      >< BslI      >< Hin6I
      >< DsaV      >< BsiYI      >< HhaI
      >< BsiLI      >< HaeII
      >< BsaJI      >< HphI      >< CfoI      >< NlaIII
      >< ApyI      >< Bsp143II      >< BspHI      EcoNI ><
GCCATCCCAG GAGAGGAACT TCACCACAGC GCCAGCAATT TGTCATGAAG GCAAAGCATA CTTCCCTCGT
  24650      24660      24670      24680      24690      24700      24710

      >< MnlI
      >< BslI      >< Tru9I
      >< BsiYI      >< MseI      >< MnlI
GAAGGTGTTT TTGTGTTTAA TGGCACTTCT TGGTTTATTA CACAGAGGAA CTTCTTTTCT CCACAAATAA
  24720      24730      24740      24750      24760      24770      24780

      >< DdeI      >< Tru9I
      >< BsmAI      >< SfaNI
      >< SfcI      >< Alw26I      >< MseI      >< AlwI ><
TTACTACAGA CAATACATTT GTCTCAGGAA ATTGTGATGT CGTTATTGGC ATCATTAAACA ACACAGTTTA
  24790      24800      24810      24820      24830      24840      24850

      >< Sau3AI
      >< NdeII

```

FIGURE 13.57

```

>< MboI          >< P1eI          > < ScaI
>< DpnII         >< MnlI          > < Ksp632I      > < RsaI
  >< DpnI         >< DdeI   >< HinfI      >< MboII
>< BspAI         >< BspWI         > < Eam1104I      >< Csp6I
  >< Bsp143I      >< AluI          > < EarI   > < AluI   > < AfaI   > < HphI
TGATCCTCTG CAACCTGAGC TTGACTCATT CAAAGAAGAG CTGGACAAGT ACTTCAAAAA TCATACATCA
24860      24870      24880      24890      24900      24910      24920

  >< Sau3AI
  >< NdeII
  >< MboI
>< MamI
  >< DpnII
  >< DpnI
  >< BspAI
  >< Bsp143I
  >< BsiBI          >< Tru9I          >< HindII
  >< BsaBI          >< MseI          >< HincII      AciI ><
CCAGATGTTG ATCTTGCGCA CATTTCAGGC ATTAACGCTT CTGTCGTCAA CATTCAAAAA GAAATTGACC
24930      24940      24950      24960      24970      24980      24990

  >< Tru9I
  > < TfiI
  >< MnlI          >< SwaI
  >< EcoNI         >< MseI
  >< BslI          > < HinfI
>< MnlI>< BsiYI   >< DraI
GCCTCAATGA GGTCTGCTAAA AATTTAAATG AATCACTCAT TGACCTTCAA GAATTGGGAA AATATGAGCA
25000      25010      25020      25030      25040      25050      25060

  >< StyI
  >< P1eI
  >< HaeIII
  >< EcoT14I
  >< Ecol30I
  >< BsuRI
  >< BssT1I
  >< Tru9I>< BshI          NlaIII ><
  >< MseI >< BsaJI          MaeIII ><
  >< BstXI
ATATATTAAA TGGCCTTGGT ATGTTTGGCT CGGCTTCATT GCTGGACTAA TTGCCATCGT CATGGTTACA
25070      25080      25090      25100      25110      25120      25130

  > < SphI
  > < PaeI
  >< SpeI          > < NspI
  > < RmaI        > < NspHI
  >< NlaIII       > < NlaIII
  > < MaeI        >< MnlI>< BbvI Fnu4HI ><
ATCTTGCTTT GTTGCATGAC TAGTTGTTGC AGTTGCCTCA AGGGTGCATG CTCTTGTGGT TCTTGCTGCA
25140      25150      25160      25170      25180      25190      25200

  >< FokI
  >< DdeI
>< MnlI >< P1eI>< HinfI >< BsrI
AGTTTGATGA GGATGACTCT GAGCCAGTTC TCAAGGGTGT CAAATTACAT TACACATAAA CGAACTTATG
25210      25220      25230      25240      25250      25260      25270

  >< Sau3AI
  >< NdeII
  >< MboI
  >< DpnII
  > < DpnI

```

FIGURE 13.58

# EP 1 694 829 B1

```

                >< BspAI
                  > < BspI43I
                    >< BsgI      >< AlwI      >< BsrI      BspWI >
GATTTGTTTA TGAGATTTT TACTCTTGGG TCAATTACTG CACAGCCAGT AAAAATTGAC AATGCTTCTC
25280      25290      25300      25310      25320      25330      25340

    >< ScaI
    >< RsaI
    >< Csp6I      >< SfcI
    >< AfaI      >< NlaIII      >< AciI      >< MnlI      FokI >
CTGCAAGTAC TGTTTCATGCT ACAGCAACGA TACCGCTACA AGCCTCACTC CCTTTCGGAT GGCTTGTAT
25350      25360      25370      25380      25390      25400      25410

                > < HinPII
                > < Hin6I
                  >< HhaI      RmaI ><
                  >< HaeII      >< HinPII      NheI ><
                  >< Eco47III      >< Hin6I      MaeI ><
                  >< CfoI      >< HhaI      Fnu4HI ><
    >< BspWI      >< BspI43II      >< CfoI      AluI ><
TGGCGTTGCA TTTCTTGCTG TTTTTCAGAG CGCTACCAA ATAATTGCGC TCAATAAAAG ATGGCAGCTA
25420      25430      25440      25450      25460      25470      25480

    >< EcoNI
    >< BslI
    >< BsiYI
    >< BbvI      >< BsrI      >< BbvI      > < Fnu4HI      BbvI ><
GCCCTTTATA AGGGCTTCCA GTTCATTTGC AATTACTGC TGCTATTTGT TACCATCTAT TCACATCTTT
25490      25500      25510      25520      25530      25540      25550

                Zsp2I ><
                Ppu10I ><
    > < SfcI      >< HinPII
    >< PstI      >< Hin6I      >< RsaI      NsiI ><
    > < Fnu4HI      >< HhaI      >< Csp6I      Mph1103I ><
    >< BspMI      >< MnlI      >< CfoI      >< AfaI      >< MnlI      AvaIII ><
TGCTTGTCGC TGCAGGTATG GAGGCGCAAT TTTTGTACCT CTATGCCTTG ATATATTTTC TACAATGCAT
25560      25570      25580      25590      25600      25610      25620

    >< SfaNI
    >< NspI
    >< NspHI
    >< NlaIII      >< SfaNI
CAACGCATGT AGAATTATTA TGAGATGTTG GCTTTGTTGG AAGTGCAAAT CCAAGAACCC ATTACTTTAT
25630      25640      25650      25660      25670      25680      25690

                >< Bst1107I
                >< AccI      MaeIII ><
GATGCCAACT ACTTTGTTTG CTGGCACACA CATAACTATG ACTACTGTAT ACCATATAAC AGTGTCACAG
25700      25710      25720      25730      25740      25750      25760

                >< MboII
                BstXI ><
    >< MunI >< MaeIII >< MaeIII      >< Eco57I      >< BbsI MnlI >
ATACAATTGT CGTTACTGAA GGTGACGGCA TTTCAACACC AAAACTCAAA GAAGACTACC AAATTGGTGG
25770      25780      25790      25800      25810      25820      25830

                >< RsaI
                > < NlaIII
                >< HphI
    >< Tru9I >< Tth111I >< Csp6I
    >< DdeI      >< DdeI      >< MseI >< AspI      >< AfaI

```

FIGURE 13.59

```

TTATTCTGAG GATAGGCACT CAGGTGTTAA AGACTATGTC GTTGTACATG GCTATTTTAC CGAAGTTTAC
25840      25850      25860      25870      25880      25890      25900

      > < HinfI>< P1eI      >> BsrI      Tru9I ><
      >< AluI >< AccI      >< SfcI >< AlwNI      >< MboII      HindIII >
TACCAGCTTG AGTCTACACA AATTACTACA GACACTGGTA TTGAAAATGC TACATTCTTC ATCTTTAACA
25910      25920      25930      25940      25950      25960      25970

      >< Tru9I      >< TthHB8I
      >< MseI      >< TaqI      >< Ksp632I
      >< AluI      >< MboII      >< EarI BspWI ><
AGCTTGTTAA AGACCCACCG AATGTGCAAA TACACACAAT CGACGGCTCT TCAGGAGTTG CTAATCCAGC
25980      25990      26000      26010      26020      26030      26040

      >< XhoII
      >< Sau3AI
      >< NlaIV
      >< NdeII
      >< MflI
      >< MboI
      >< DpnII
      >< DpnI
      >< BstYI
      >< BstI
      >< BspAI
      >< Bsp143I
      >< BscBI      >< RmaI      RsaI ><
      >< BamHI >< AlwI      >< MaeI      Csp6I ><
AATGGATCCA ATTTATGATG AGCCGACGAC GACTACTAGC GTGCCTTTGT AAGCACAAGA AAGTGAGTAC
26050      26060      26070      26080      26090      26100      26110

      >< Tru9I
      >< RsaI
      >< MseI
      >< MboII
      >< RsaI      >< MaeII      >< RsaI
      >< Csp6I      >< Csp6I      >< Tru9I >< Csp6I
      >< AfaI      >< AfaI      >< MseI      >< AfaI
GAACTTATGT ACTCATTCGT TTCGGAAGAA ACAGGTACGT TAATAGTTAA TAGCGTACTT CTTTTCTTGT
26120      26130      26140      26150      26160      26170      26180

      >< TthHB8I
      >< TaqI
      >< RmaI      >< HinfI      >< RsaI
      >< MaeIII      >< HinfI      Fnu4HI ><
      >< MaeI >< RmaI      >< HhaI      >< Csp6I
      >< FokI >< MaeI      >< CfoI >< BbvI >< AfaI
CTTTCGTGGT ATTCTTGCTA GTCACACTAG CCATCCTTAC TGCCTTCGA TTGTGTGCGT ACTGCTGCAA
26190      26200      26210      26220      26230      26240      26250

      >< Tru9I
      >< MseI      >< ThaI
      >< SspI >< MaeII      >< MvnI
      >< HpaI      >< MseI
      >< HindII      >< BstUI      Ksp632I >
      >< HincII      >< MaeII >< Bsp50I >< MboII EarI >
      >< AccI >< AccII      Eam1104I >
TATTGTAAAC GTGAGTTTAG TAAAACCAAC GGTTCACGTC TACTCGCGTG TTAAAAATCT GAACTCTTCT
26260      26270      26280      26290      26300      26310      26320

```

FIGURE 13.60



```

>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< MboII>< DpnI
>< XmnI >< BspAI> < Eco57I
>< Asp700I>< Bsp143I
GAAGGAGTTC CTGATCTTCT GGTCTAAACG AACTAACTAT TATTATTATT CTGTTTGAA CTTTAACATT
26330      26340      26350      26360      26370      26380      26390

>< ScrFI
>< MvaI
>< EcoRII
>< Ecl136I
>< DsaV NlaIV ><
>< RsaI
>< MnlI
>< Tru9I
>< BstOI
>< Csp6I
>< MseI
>< BstNI RmaI ><
>< AfaI
>< BsiLI MaeI ><
> < NlaIII >< AluI
>< ApyIBscBI ><
GCTTATCATG GCAGACAACG GTACTATTAC CGTTGAGGAG CTTAAACAAC TCCTGGAACA ATGGAACCTA
26400      26410      26420      26430      26440      26450      26460

>< ScrFI
>< RmaI
>< MvaI
>< MaeI
>< EcoRII
>< Ecl136I
>< DsaV
>< BstOI
>< BstNI
>< BsiLI
>< ApyI >< MaeIII
GTAATAGGTT TCCTATTCCT AGCCTGGATT ATGTTACTAC AATTGCGCTA TTCTAATCGG AACAGGTTTT
26470      26480      26490      26500      26510      26520      26530

>< Pali
>< MscI
>< MnlI >< MaeIII
>< HaeIII
>< EaeI
>< BsuRI
>< BsrI
>< RsaI
>< BspWI
>< Csp6I >< HindIII
>< BshI
>< AfaI >< AluI
>< BalI
>< BbvI Fnu4HI ><
TGTACATAAT AAAGCTTGTT TTCCTCTGGC TCTTGTGGCC AGTAACACTT GCTTGTTTTG TGCTTGCTGC
26540      26550      26560      26570      26580      26590      26600

>< VspI
>< Tru9I
>< MseI
>< HphI
>< SfcI >< AsnI
>< BsrI
>< AccI >< AseI>< MaeIII>< AciI
TGTCTACAGA ATTAATTGGG TGAATGGCGG GATTGCGATT GCAATGGCTT GTATTGTAGG CTTGATGTGG
26610      26620      26630      26640      26650      26660      26670

>< EspI
>< Eco57I
>< DdeI
>< CeiII
>< Bpu1102I
>< RsaI
>< Csp6I

```

FIGURE 13.61

EP 1 694 829 B1

```

>< BfrI
    >< AluI
CTTAGCTACT TCGTTGCTTC CTTGAGGCTG TTTGCTCGTA CCCGCTCAAT GTGGTCATTC AACCCAGAAA
    26680      26690      26700      26710      26720      26730      26740

    >< AfaI
    >< AciI
    MboII >
    >< ScrFI
    >< NciI
    >< MspI
    >< HpaII
    >< HapII
    >< DsaV>< MnlI
    >< BslI
    >< BsiYI
    >< BsaJI >< MunI
    >< BcnI >< MaeIII >< AciI >< NlaIII
    >< XcmI
CAAACATTCT TCTCAATGTG CCTCTCCGGG GGACAATTGT GACCAGACCG CTCATGGAAA GTGAACCTTG
    26750      26760      26770      26780      26790      26800      26810

    Tru9I ><
    SinI >
    Sau96I >
    PpuMI >
    NspIV >
    MseI ><
    >< MaeIII
    >< Sau3AI
    >< NdeII
    >< MboI
    >< FbaI
    >< DpnII
    >< DpnI
    >< BspAI
    >< Bsp143I
    >< BsiQI
    >< BclI
    >< MaeIII
    >< Pali
    >< MspI
    >< HpaII
    >< HapII
    >< HaeIII
    >< GdiII
    >< EaeI
    >< BsuRI
    >< BshI
    >< RmaI >< HaeII
    >< MaeI >< Eco0109I >
    >< HinPIIEco47I >
    >< StyI>< Hin6I DraII >
    >< EcoT14I >< Cfr13I >
    >< Eco130I>< Bsp143II
    >< BssT1I >< BsiZI >
    >< BsaJI >< Bme18I >
    >< BlnI >< HhaI AvaII >
    >< AvrII >< CfoI AsuI >
CATTGGTGCT GTGATCATTC GTGGTCACTT GCGAATGGCC GGACACTCCC TAGGGCGCTG TGACATTAAG
    26820      26830      26840      26850      26860      26870      26880

    >< Sau3AI
    >< NdeII
    >< MboI
    >< DpnII
    >< DpnI
    >< PssI >< BspMI
    >< Psp5II >< BspAI
    >< NspHII >< Bsp143I
    GACCTGCCAA AAGAGATCAC TGTGGCTACA TCACGAACGC TTTCTTATTA CAAATTAGGA GCGTCGCAGC
    26890      26900      26910      26920      26930      26940      26950

    >< TfiI
    >< HinfI
    >< BbvI
    >< BbvI
    >< Fnu4HI >< AciI
    GTGTAGGCAC TGATTGAGGT TTTGCTGCAT ACAACCGCTA CCGTATTGGA AACTATAAAT TAAATACAGA
    26960      26970      26980      26990      27000      27010      27020

    >< MspI
    >< HpaII
    >< HapII
    >< Cfr10I
    >< BcgI/a
    >< RsaI
    >< RmaI
    >< Csp6I
    >< MaeI>< BcgI
    >< AfaI >< MaeIII
    HindII ><
    HincII ><

```

FIGURE 13.62

# EP 1 694 829 B1

```

CCACGCCGGT AGCAACGACA ATATTGCTTT GCTAGTACAG TAAGTGACAA CAGATGTTTC ATCTTGTTGA
  27030      27040      27050      27060      27070      27080      27090

    >< ScrFI
    >< MvaI
      >< MaeIII
    >< EcoRII
      >< Ecl136I
    >< DsaV
      >< BstOI
      >< BstNI
      >< BsiLI
      >< ApyI
      >< MnlI
      >< TfiI
      >< HinfI ><
CTTCCAGGTT ACAATAGCAG AGATATTGAT TATCATTATG AGGACTTTCA GGATTGCTAT TTGGAATCTT
  27100      27110      27120      27130      27140      27150      27160

      >< BsmAI
      >< Tru9I
      > < MnlI
    >< MaeII
      >< Alw26I
      >< MseI
      >< DdeI
      >< MboII
GACGTTATAA TAAGTTCAAT AGTGAGACAA TTATTTAAGC CTCTAACTAA GAAGAATTAT TCGGAGTTAG
  27170      27180      27190      27200      27210      27220      27230

      >< Ksp632I
      >< MboII
      >< EarI
      >< MboII
      >< NlaIIIEam1104I ><
ATGATGAAGA ACCTATGGAG TTAGATTATC CATAAACGA ACATGAAAT TATTCTCTTC CTGACATTGA
  27240      27250      27260      27270      27280      27290      27300

      > < RsaI >< RsaI
      >< Csp6I >< Csp6I
      > < AluI
      >< MnlI
      > < AfaI >< AfaI
TTGTATTTAC ATCTTGCGAG CTATATCACT ATCAGGAGTG TGTTAGAGGT ACGACTGTAC TACTAAAAGA
  27310      27320      27330      27340      27350      27360      27370

      >< MnlI
      >< HphI
      >< HphI
      >< MnlI
ACCTTGCCCA TCAGGAACAT ACGAGGGCAA TTCACCATTT CACCCTCTTG CTGACAATAA ATTTGCACTA
  27380      27390      27400      27410      27420      27430      27440

      >< RmaI
      >< MaeI
      >< RmaI
      >< MaeI
      >< BbvI
      >< AfaI
      > < AluI
ACTTGCACTA GCACACACTT TGCTTTTGCT TGTGCTGACG GTACTCGACA TACCTATCAG CTGCGTGCAA
  27450      27460      27470      27480      27490      27500      27510

      >< SstI
      >< SduI
      >< SacI
      >< NspII
      >< HgiAI
      >< Eco24I
      > < Ecl136II
      >< BspWI
      >< Bsp1286I
      >< BmyI
      >< BanII
      >< Alw21I
    >< HphI
    >< DpnI
      >< MnlI

```

FIGURE 13. 63

```

>< Bsp143I          >< MnlI          > < AluI      BbvI ><
GATCAGTTTC ACCAAACTT TTCATCAGAC AAGAGGAGGT TCAACAAGAG CTCTACTCGC CACTTTTTCT
27520      27530      27540      27550      27560      27570      27580

                                SstI ><
                                SduI ><
                                SacI ><
                                NspII ><
                                HgiAI ><
                                Eco24I ><
                                Ecl136II ><
                                Bsp1286I ><
                                BmyI ><
                                BanII ><
                                Alw21I ><
                                AluI ><
>< Fnu4HI          >< RmaI      >< Tru9I
>< MaeI          >< MseI          >< Tru9I
CATTGTTGCT GCTCTAGTAT TTTTAATACT TTGCTTCACC ATTAAGAGAA AGACAGAATG AATGAGCTCA
27590      27600      27610      27620      27630      27640      27650

>< Tru9I          >< Tru9I
>< MseI          >< MseI
CTTTAATTGA CTTCTATTTG TGCTTTTTAG CCTTCTGCT ATTCCTTGTT TTAATAATGC TTATTATATT
27660      27670      27680      27690      27700      27710      27720

                                >< XhoII
                                >< XbaI
                                > < ScrFI
                                >< Sau3AI
                                >< RmaI
                                >< NdeII
                                > < MvaI
                                >< MflI
                                >< MboI
                                >< EcoRII>< MaeI
                                > < Ecl136I
                                >< DpnII
                                >< DpnI
                                >< BstYI
                                > < BstOI
                                > < BstNI
                                >< TthHB8I >< BspAI          > < RsaI
                                >< DsaV>< Bsp143I      >< MboII
                                > < BsiLI          >< Csp6I
                                >< TaqI > < ApyI > < AlwI > < AfaI          >< NlaIII
TTGGTTTTCA CTCGAAATCC AGGATCTAGA AGAACCTGT ACCAAAGTCT AAACGAACAT GAAACTTCTC
27730      27740      27750      27760      27770      27780      27790

                                >< HinPII
                                >< Hin6I
                                >< HhaI
                                >< RsaI >< HaeII
                                >< SfcI          >< Eco47III
                                >< Csp6I>< CfoI SfaNI ><
                                >< NdeI          >< AfaI >< Bsp143II
ATTGTTTTGA CTTGTATTTT TCTATGCAGT TGCATATGCA CTGTAGTACA GCGCTGTGCA TCTAATAAAC
27800      27810      27820      27830      27840      27850      27860

                                >< XhoII
                                >< Sau3AI
                                >< NdeII
                                > < MnlI
                                >< MflI

```

FIGURE 13.64

```

    >< MboI
    >< DpnII
    >< DpnI    >< RsaI
    >< BstYI    >< MboII
    >< NlaIII>< BspAI    >< Csp6I >< RmaI
    >< AlwI >< Bsp143I    >< AfaI >< MaeI
CTCATGTGCT TGAAGATCCT TGTAAGGTAC AACACTAGGG GTAATACTTA TAGCACTGCT TGGCTTTGTG
    27870      27880      27890      27900      27910      27920      27930

    >< SduI
    >< RmaI
    >< NspII
    >< MaeI
    >< HgiAI
    >< Bsp1286I
    >< BmyI
    >< Alw21I
    >< NspI
    >< NspHI
    >< NlaIII >< MaeIII
CTCTAGGAAA GGTTTTACCT TTTCATAGAT GGCACACTAT GGTTCAAACA TGCACACCTA ATGTTACTAT
    27940      27950      27960      27970      27980      27990      28000

    > < XhoII
    > < Sau3AI > < Van91I
    >< PvuII
    >< Psp5I
    > < NdeII > < PflMI
    > < MflI>< NspBII
    > < DpnII    >< HinPII
    >< Bsp143I    >< Hin6I
    > < BstYI > < BslI >< HhaI >< RmaI
    > < BspAI > < BsiYI>< CfoI >< MaeI
    > < MboI>< AluI>< BspWI >< BspWI
    >< AlwI >< DpnI > < AccB7I
    >< AluI
    >< RsaI
    >< NlaIV
    >< KpnI >< NlaIII
    >< Eco64I >< MaeIII
    >< Csp6I>< HphI
    >< BscBI >< EcoO651
    >< BanI >< BspHI
    >< Asp718 >< Eco91I
    >< AfaI >< BstPI
    >< AccB1I >< BstEII
    >< Acc65I >< BbvI
CAACTGTCAA GATCCAGCTG GTGGTGCCT TATAGCTAGG TGTTGGTACC TTCATGAAGG TCACCAAACCT
    28010      28020      28030      28040      28050      28060      28070

    >< SinI
    >< Sau96I
    >< NspIV
    NspHII ><
    NlaIV ><
    >< Eco47I
    >< Cfr13I
    >< BsiZi
    >< RsaI
    >< Fnu4HI >< MaeII
    >< Esp3I >< Csp6I >< Tru9I
    >< BsmAI >< BsmBI >< MseI
    >< Alw26I >< AfaI >< DraI
    >< Tru9I
    >< MseI
    >< BscBI ><
    >< Bme18I
    >< AvaII
    >< AsuI
GCTGCATTTA GAGACGTACT TGTTGTTTTA AATAAACGAA CAAATTAAAA TGTCTGATAA TGGACCCCAA
    28080      28090      28100      28110      28120      28130      28140

    >< SinI
    >< Sau96I
    >< NspIV
    >< NspHII
    >< NlaIV
    >< Eco47I
    >< Cfr13I
    >< BsiZi
    >< SduI
    >< NspII
    >< Bsp1286I
    >< BmyI
    >< MaeII
    >< AciI
    >< BscBI
    >< Bme18I
    >< AvaII >< TfiI
    >< AsuI >< HinfI
    >< MnlI

```

FIGURE 13. 65

# EP 1 694 829 B1

```

TCAAACCAAC GTAGTGCCCC CCGCATTACA TTTGGTGGAC CCACAGATTC AACTGACAAT AACCAGAATG
28150      28160      28170      28180      28190      28200      28210

                >< HinPII >< StyI
                  >< HaeII
                > < Pali >< Hin6I >< EcoT14I
                > < HaeIII >< HhaI>< Eco130I
                  >< BspWI      >< BssTII
                > < BsuRI      >< Bsp143II
                >< HgaI> < BshI      >< CfoI>< BsaJI      >< HgaI
GAGGACGCAA TGGGGCAAGG CCAAAACAGC GCCGACCCCA AGGTTTACCC AATAATACTG CGTCTTGTTT
28220      28230      28240      28250      28260      28270      28280

                                >< TthHB8I
                                  > < ScrFI
                                    >< Pali
                                >< PaeR7I
                                >< NspIII
                                  > < MvaI
                                    >< HaeIII
                                    >< EcoRII
                                >< Eco88I
                                >< XhoI > < Ecl136I
                                  >< DsaV
                                    >< BsuRI
                                >< SlaI > < BstOI
                                >< MnlI>< TaqI> < BstNI
                                  >< CcrI > < BsiLI
                                >< HinfI      >< BshI
                                >< TfiI>< BcoI>< BsaJI
                                  >< MnlI      >< DdeI      >< AvaI > < ApyI
                                >< AluI >< DdeI > < NlaIII >< BfrI      >< Ama87I >< MnlI
CACAGCTCTC ACTCAGCATG GCAAGGAGGA ACTTAGATTC CCTCGAGGCC AGGGCGTTCC AATCAACACC
28290      28300      28310      28320      28330      28340      28350

                >< SinI
                >< Sau96I
                >< NspIV
                >< NspHII
                >< Eco47I
                >< Cfr13I
                >< BsiZI
                >< Bme18I      > < Ksp632I
                >< AvaII      > < Eam1104I
                >< AsuI      > < EarI > < AluI>< MboII      >< MaeIII
AATAGTGGTC CAGATGACCA AATTGGCTAC TACCGAAGAG CTACCCGACG AGTTCGTGGT GGTGACGGCA
28360      28370      28380      28390      28400      28410      28420

                >< SstI
                >< SduI
                >< SacI
                >< NspII
                >< HgiAI
                >< EspI
                >< Eco24I
                >< Ecl136II      >< StyI      >< Sau96I
                >< DdeI      >< RmaI      >< NspIV
                >< CelII      >< MaeI      >< HaeIII
                >< Bsp1286I      >< EcoT14I >< Cfr13I
                >< Bpu1102I      >< Eco130I >< BsuRI
                >< BmyI      >< BssTII      >< BsrI
                >< BanII      >< RsaI      >< BsaJI      >< BsiZI

```

FIGURE 13.66

# EP 1 694 829 B1

```

    >< Alw21I    >< Csp6I    >< BlnI    >< BshI>< HindIII
>< HphI >< AluI    >< AfaI    >< AvrII    >< AsuI    >< AluI
AAATGAAAGA GCTCAGCCCC AGATGGTACT TCTATTACCT AGGAACTGGC CCAGAAGCTT CACTTCCCTA
28430      28440      28450      28460      28470      28480      28490

>< HinPII
>< Hin6I
>< HhaI
>< HaeII
>< CfoI          > < MnlI    >< NlaIV
>< Bsp143II      >< SfaNI >< DdeI >< BscBI
CGGCGCTAAC AAAGAAGGCA TCGTATGGGT TGCAACTGAG GGAGCCTTGA ATACACCCAA AGACCACATT
28500      28510      28520      28530      28540      28550      28560

>< NlaIV
>< Eco64I
>< BscBI
>< BanI
>< AciI
>< AccB1I >< BbvI    >< Fnu4HI          >< MnlI
GGCACCCGCA ATCCTAATAA CAATGCTGCC ACCGTGCTAC AACTTCCTCA AGGAACAACA TTGCCAAAAG
28570      28580      28590      28600      28610      28620      28630

                                >< ThaI
                                >< MnlI
                                >< MaeII >< MvnI
                                BstUI ><
                                >< MnlI
                                >< Ksp632I Bsp50I ><
                                >< Fnu4HI
                                >< BspWI
                                >< EarI    >< BsaAI>< AciI
                                >< Eam1104I AccII ><
>< MnlI    >< MnlI    >< AciI>< MboII    >< Eam1104I AccII ><
GCTTCTACGC AGAGGGAAGC AGAGGCGGCA GTCAAGCCTC TTCTCGCTCC TCATCACGTA GTCGCGGTAA
28640      28650      28660      28670      28680      28690      28700

                                >< ScrFI
                                >< MvaI
                                >< EcoRII
                                >< Ecl136I
                                >< DsaV>< Fnu4HI
                                >< BstOI
                                >< BstNI
                                >< BsiLI
                                >< ApyI
                                >< BbvI
                                >< TaqI
                                >< AciI
                                >< TthHB8I
                                >< RmaI
                                >< NheI
                                >< MnlI
                                >< MaeI
                                > < BspWI
                                >< BbvI
                                >< TaqI
                                >< AciI
TTCAAGAAAT TCAACTCCTG GCAGCAGTAG GGGAAATCT CCTGCTCGAA TGGCTAGCGG AGGTGGTGAA
28710      28720      28730      28740      28750      28760      28770

> < ThaI
> < MvnI
>< HphI >< MnlI
> < HinPII
> < Hin6I
>< HhaI
> < BstUI >< RmaI
> < Bsp50I >< MaeI
>< BbvI >< CfoI>< Fnu4HI
> < AccII>< BspWI
>< AluI
ACTGCCCTCG CGCTATTGCT GCTAGACAGA TTGAACCAGC TTGAGAGCAA AGTTTCTGGT AAAGGCCAAC
28780      28790      28800      28810      28820      28830      28840

                                RsaI ><
                                >< MnlI
                                MaeII ><
                                >< BsuRI
                                >< DdeI
                                >< DdeI
                                Csp6I ><
> < PalI>< MaeIII
> < HaeIII
> < BsuRI >< DdeI

```

FIGURE 13.67

# EP 1 694 829 B1

```

> < BshI   > < BbvI   >< MnlI >< BspWI   >< SfaNI   AfaI ><
AACAACAAGG CCAAACGTGC ACTAAGAAAT CTGCTGCTGA GGCATCTAAA AAGCCTCGCC AAAAACGTAC
28850      28860      28870      28880      28890      28900      28910

                                >< Tth111I
                                >< SinI
                                >< Sau96I
                                >< NspIV
                                >< NspHII
> < MaeII
                                >< Eco47I
                                >< Cfr13I
                                >< BsmBI
                                >< BsiZI   >< StyI
                                >< Bme18I   >< EcoT14I
                                >< MaeII   >< Esp3I   >< AvaII   >< Eco130I
                                >< Csp6I   >< BsmAI   >< AsuI    >< BssT1I
                                >< AfaI   >< Alw26I> < AspI    >< BsaJI
TGCCACAAAA CAGTACAACG TCACTCAAGC ATTTGGGAGA CGTGGTCCAG AACAAACCCA AGGAAATTTC
28920      28930      28940      28950      28960      28970      28980

>< SinI
>< Sau96I
>< NspIV
>< NspHII
>< NlaIV
>< Eco47I
>< Cfr13I
>< BsiZI
>< BscBI
>< Bme18I
>< AvaII
>< AsuI
GGGGACCAAG ACCTAATCAG ACAAGGAAC TATTACAAAC ATTGGCCGCA AATTGCACAA TTTGCTCCAA
28990      29000      29010      29020      29030      29040      29050

                                >< PalI
                                >< HaeIII
                                >< GdiII
                                >< Fnu4HI
                                >< EaeI
                                >< BsuRI
                                >< BshI
                                >< AciI
                                BspWI >
                                >< BspWI
GGGGACCAAG ACCTAATCAG ACAAGGAAC TATTACAAAC ATTGGCCGCA AATTGCACAA TTTGCTCCAA
28990      29000      29010      29020      29030      29040      29050

                                >< BsmI
                                >< NlaIII
>< BscCI >< MnlI >< MaeIII >< MaeIII >< NlaIII
GTGCCTCTGC ATTCTTTGGA ATGTCACGCA TTGGCATGGA AGTCACACCT TCGGGAACAT GGCTGACTTA
29060      29070      29080      29090      29100      29110      29120

                                >< XhoII
                                >< Sau3AI
                                >< NdeII
                                >< MflI
                                >< MboI
                                >< FokI
                                >< Tru9I
                                >< DpnII
                                >< NlaIV > < DpnI
                                >< NlaIII >< BstYI
                                >< MseI   >< BspAI   >< Tth111I
                                >< BscBI >< BstXI>< AlwI> < Bsp143I >< MaeII
                                >< AspI
TCATGGAGCC ATTAAATTGG ATGACAAAGA TCCACAATTC AAAGACAACG TCATACTGCT GAACAAGCAC
29130      29140      29150      29160      29170      29180      29190

                                EspI ><
                                DdeI ><
                                CelII ><
                                Bpu1102I ><
                                AluI ><
                                >< HgaI
ATTGACGCAT ACAAACATT CCCACCAACA GAGCCTAAAA AGGACAAAAA GAAAAAGACT GATGAAGCTC
29200      29210      29220      29230      29240      29250      29260

```

FIGURE 1368



```

                >< P1eI
    >< Fnu4HI
    >< BspWI
    >< BsmAI
    >< Alw26I
    >< AciI
AGCCTTTGCC GCAGAGACAA AAGAAGCAGC CCACTGTGAC TCTTCTTCCT GCGGCTGACA TGGATGATTT
29270      29280      29290      29300      29310      29320      29330

                >< MboII
                >< Ksp632I >< GsuI
                >< EarI>< Fnu4HI
                >< HinfI >< Eam1104I>< BpmI
                >< Fnu4HI >< BbvI
                >< AciI >< NlaIII
                >< NlaIII
                >< HinfI
                NlaIII ><
    >< FokI
    CTCCAGACAA CTTCAAAATT CCATGAGTGG AGCTTCTGCT GATTCAACTC AGGCATAAAC ACTCATGATG
    29340      29350      29360      29370      29380      29390      29400

                >< MaeII
                >< AccI
ACCACACAAG GCAGATGGGC TATGTAAACG TTTTCGCAAT TCCGTTTACG ATACATAGTC TACTCTTGTG
29410      29420      29430      29440      29450      29460      29470

                >< Tru9I
                >< Tru9I
                >< MseI
                >< MseI
    >< XmnI
    >< EcoRI>< MaeIII
    >< Asp700I >< BsgI
    CAGAATGAAT TCTCGTAACT AAACAGCACA AGTAGGTTTA GTTAACTTTA ATCTCACATA GCAATCTTTA
    29480      29490      29500      29510      29520      29530      29540

                XorII >
                TthHB8I >
                TaqI >
                Sau3AI ><
                RsaI ><
                >< ThaIPvuI >
                NdeII ><
                >< MnlI
                >< MvnIMcrI >
                MboI ><
                DpnII ><
                DpnI ><
                Csp6I ><
                >< BstUI
                >< HaeIII BspCI >
                BspAI ><
                >< TthHB8I >< Bsp50I
                >< Pali Bsp143I ><
                >< BsuRI BsiEI >
                >< BshIAfaI ><
                >< MnlI
                >< MaeIII
    ATCAATGTGT AACATTAGGG AGGACTTGAA AGAGCCACCA CATTTTCATC GAGGCCACGC GGAGTACGAT
    29550      29560      29570      29580      29590      29600      29610

                >< TaqI >< AciI
                >< MnlI >< AccII
                >< SduI
                >< NspII
                >< MboII >< VspI
                >< Ksp632I >< Eco24I >< Tru9I
                >< RsaI >< RmaI >< Fnu4HI >< Bsp1286I >< MseI
                >< Csp6I >< MaeI >< EarI >< BmyI >< AsnI
                >< AfaI >< BbvI > < AluI>< Eam1104I >< BanII >< AseI

```

FIGURE 13.69

EP 1 694 829 B1

```
CGAGGGTACA GTGAATAATG CTAGGGAGAG CTGCCTATAT GGAAGAGCCC TAATGTGTAA AATTAATTTT
 29620      29630      29640      29650      29660      29670      29680

                >< Tru9I    >< DdeI
                >< MseI    >< BfrI
                >< NlaIII   > < AluI
AGTAGTGCTA TCCCCATGTG ATTTTAATAG CTTCTTAGGA GAATGACAAA AAAAAAAAAA AAAAAA
 29690      29700      29710      29720      29730      29740
```

FIGURE 13. 70

# Sérologie SRAS : Technique N Indirect (1ère série)

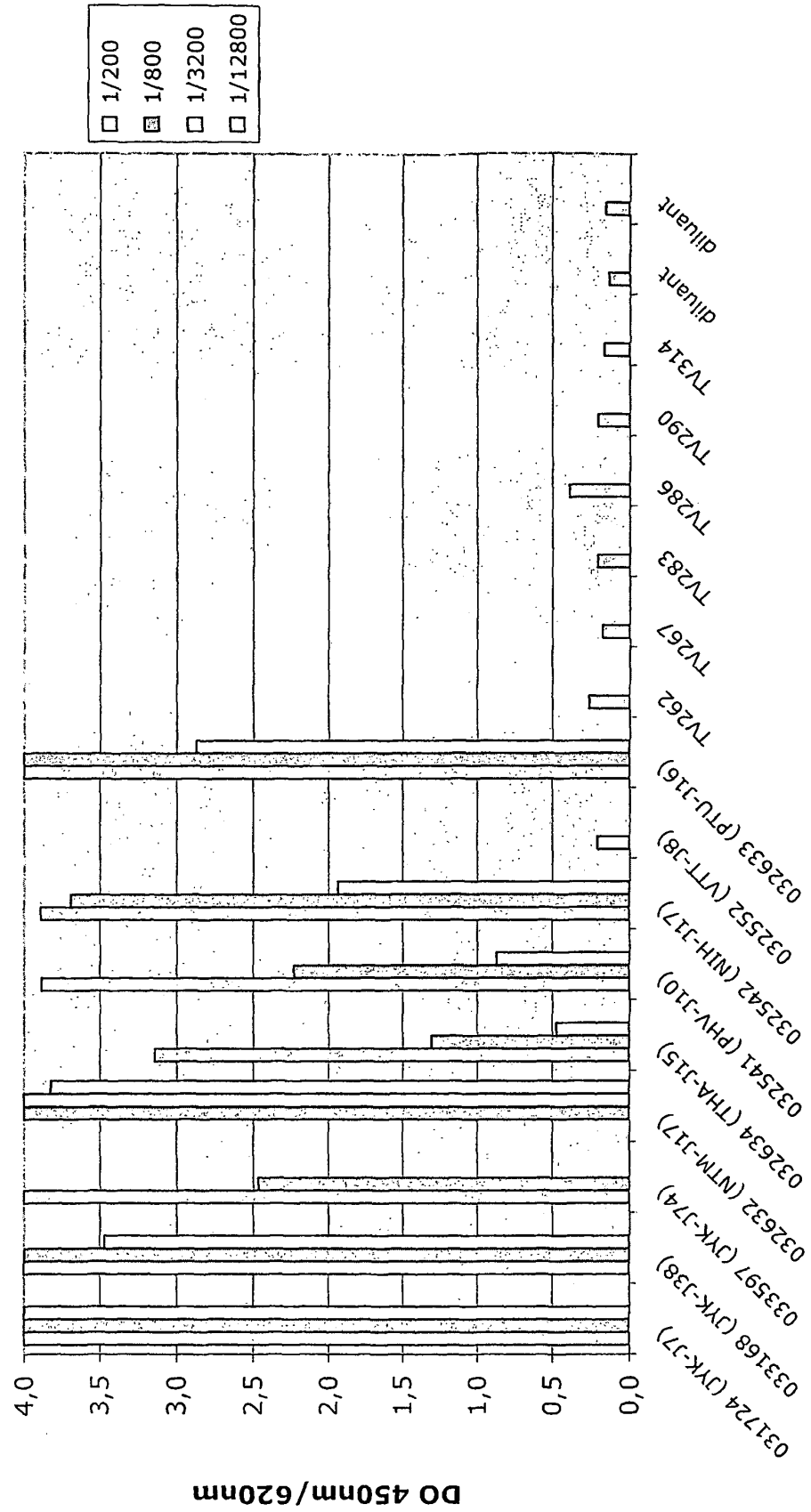


FIGURE 14

Sérologie SRAS : Technique N Indirect (2ème série)

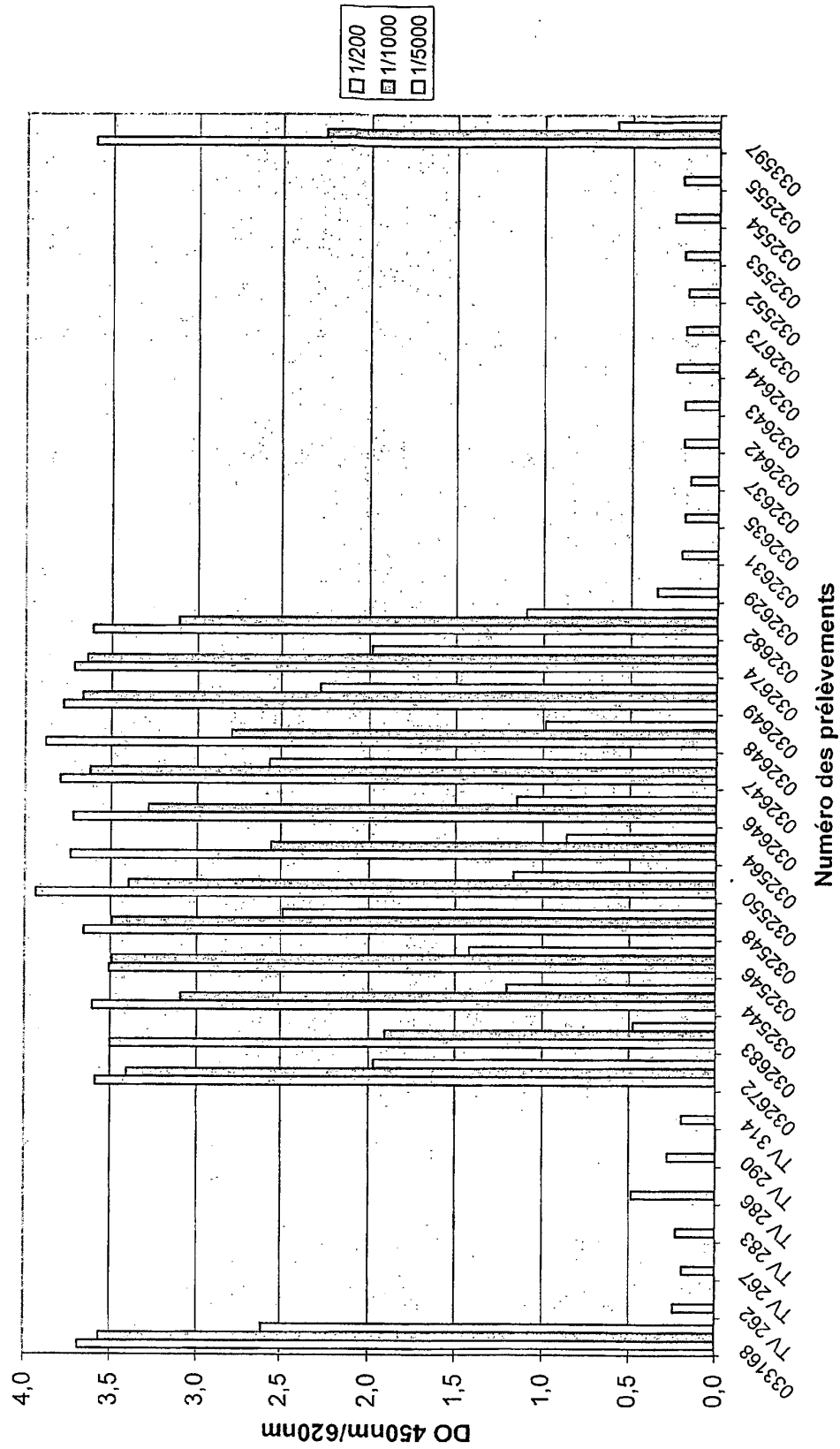


FIGURE 15

# Sérologie SRAS : Technique Double Epitope (1ère série)

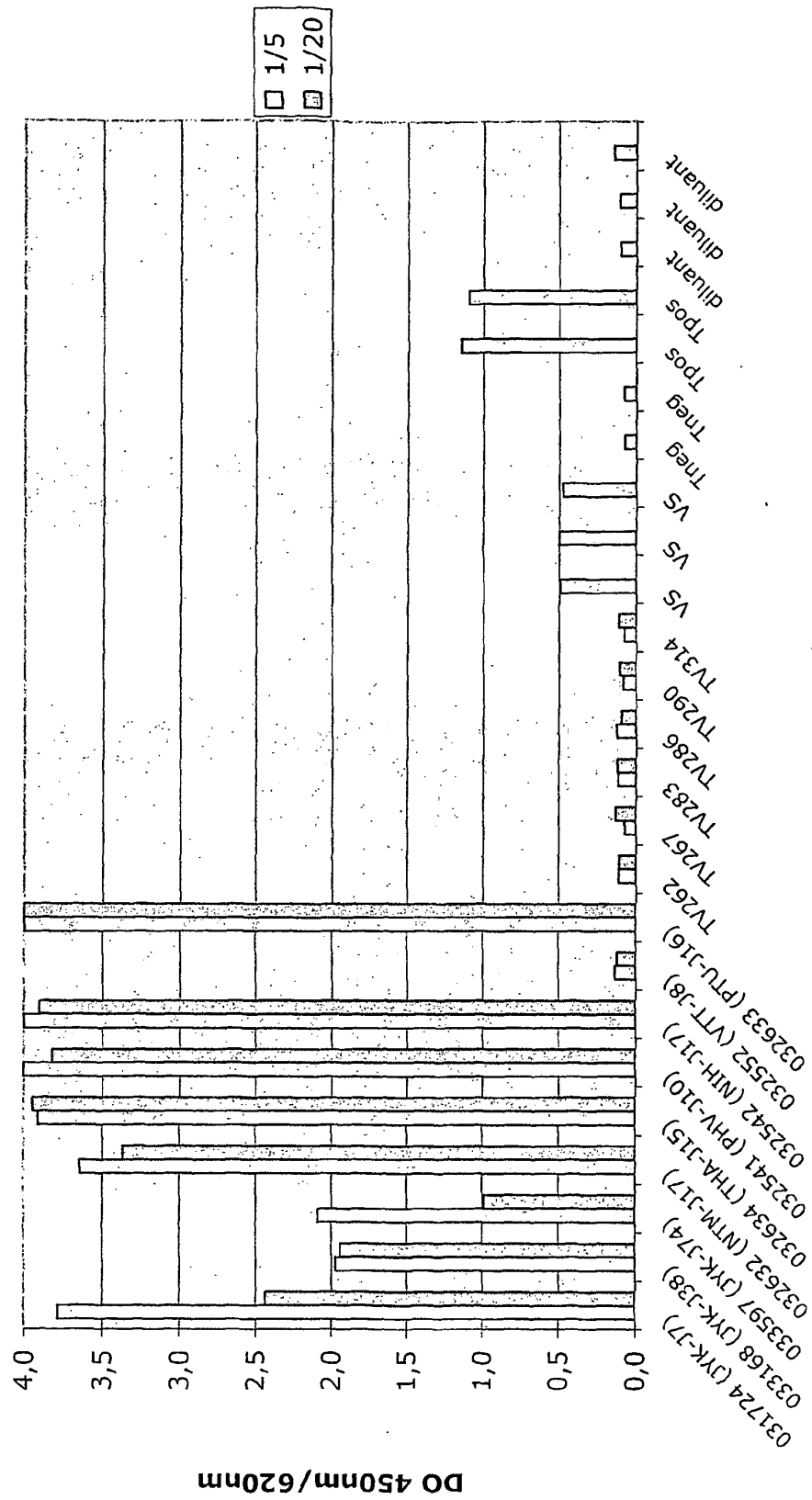
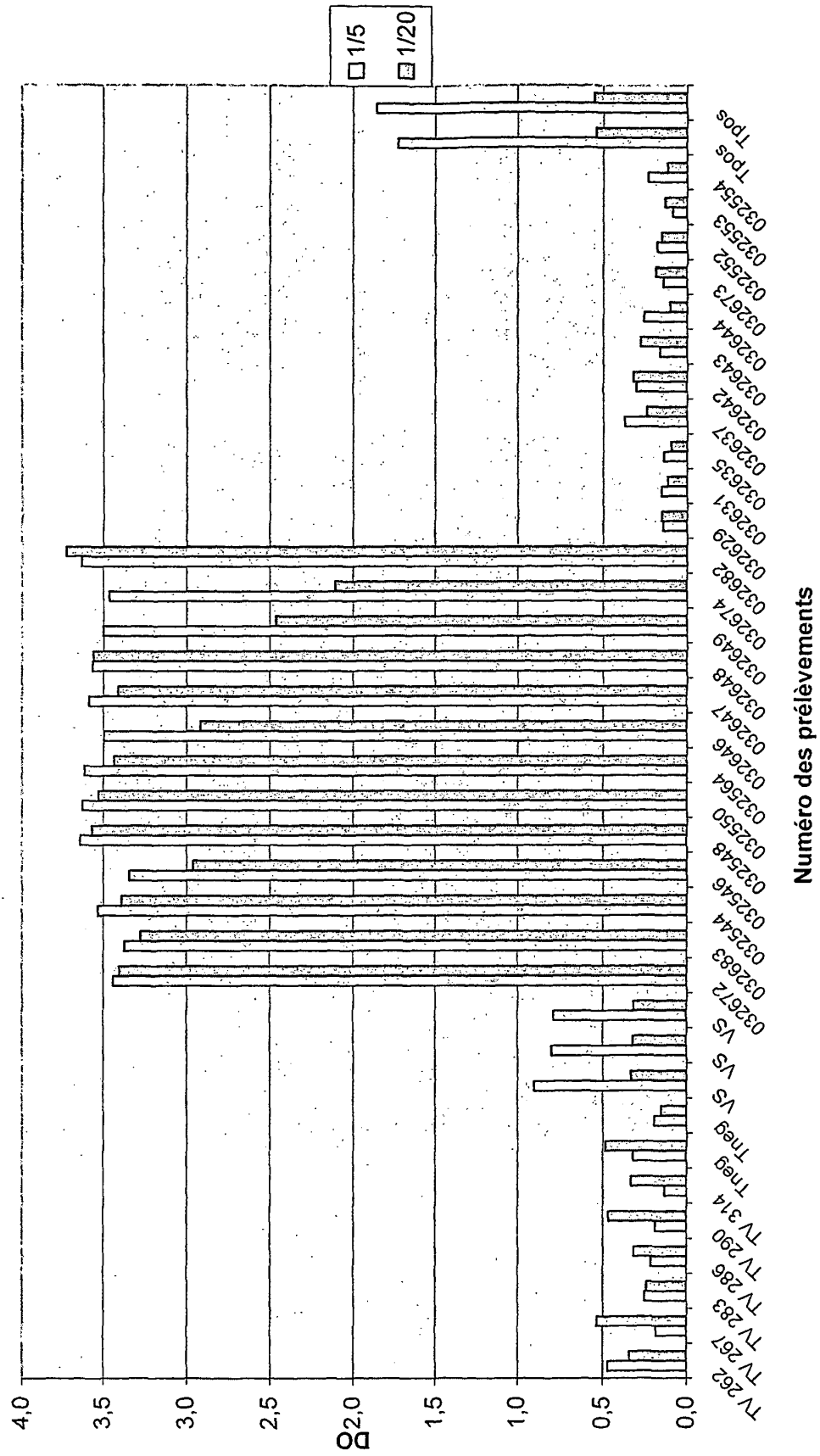


FIGURE 16

## Sérologie SRAS : Technique Double Epitope (2ème série)



**FIGURE 17**

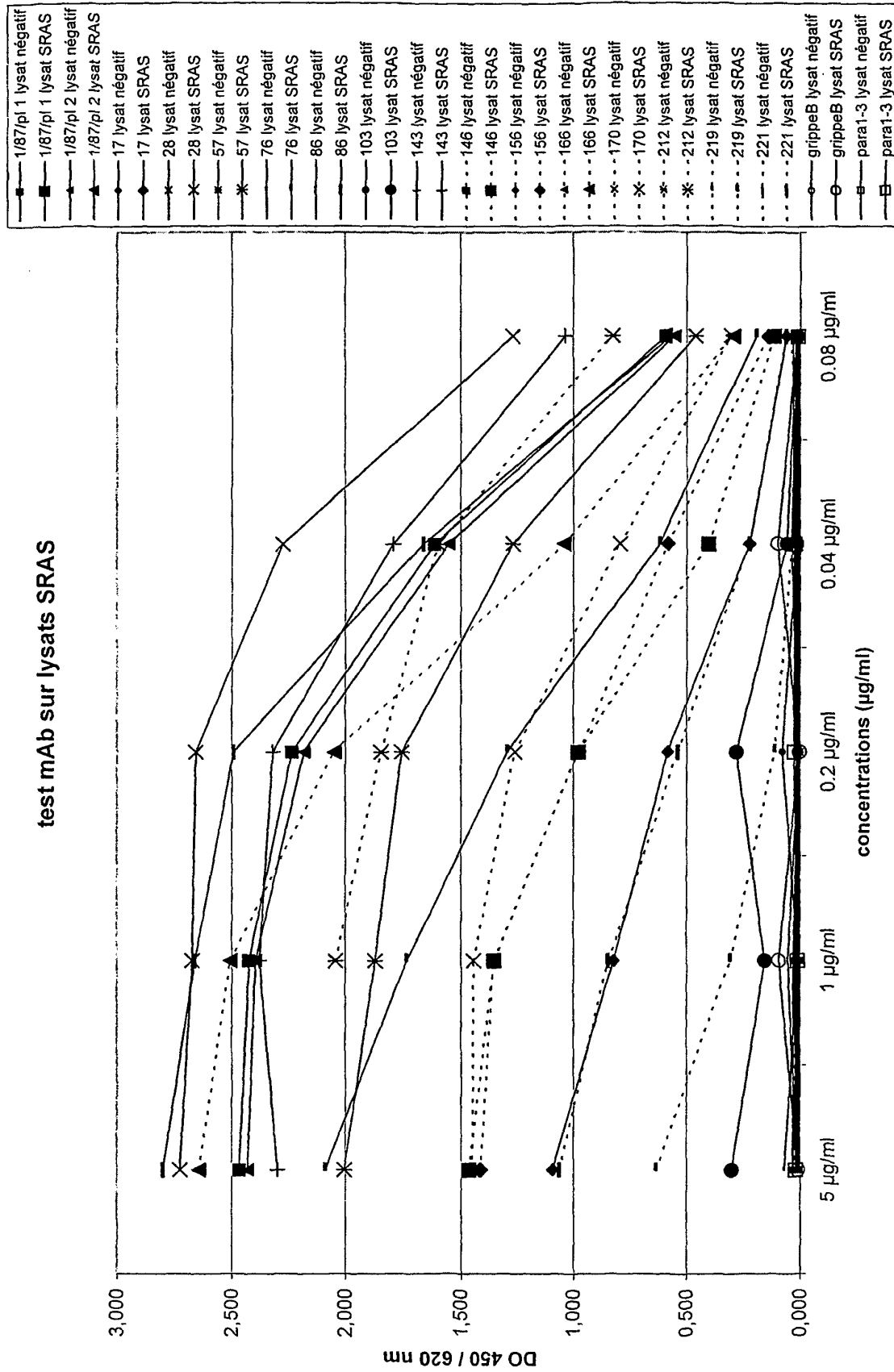


FIGURE 18

Test mAb sur lysats HCoV-229E

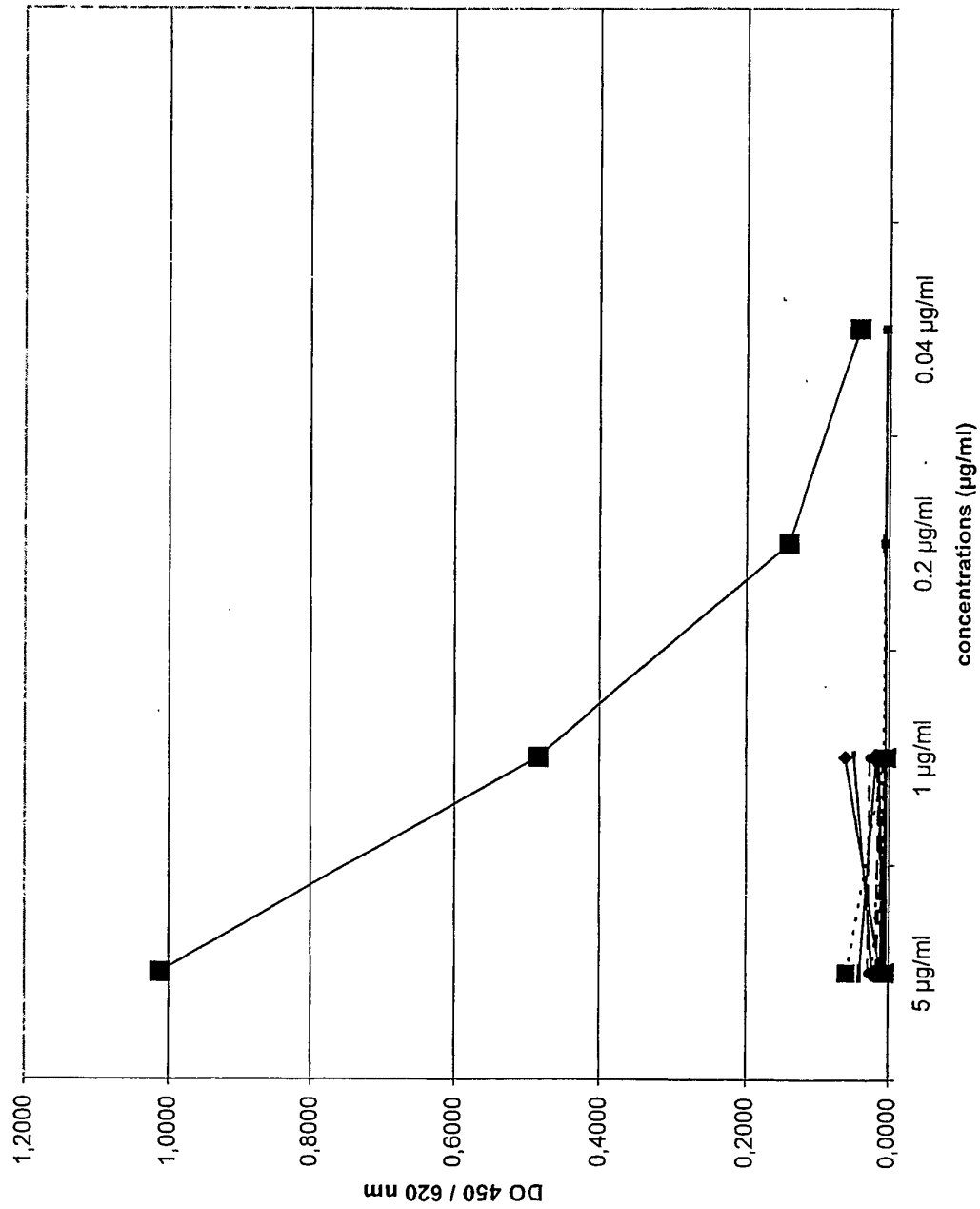


FIGURE 19

- 1/87 lysat négatif
- ▲— 1/87 lysat 229E
- 5-11H.6 lysat négatif
- 5-11H.6 lysat 229E
- 17 lysat négatif
- 17 lysat 229E
- 28 lysat négatif
- 28 lysat 229E
- ◆— 57 lysat négatif
- ◆— 57 lysat 229E
- ▲— 76 lysat négatif
- ▲— 76 lysat 229E
- 86 lysat négatif
- 86 lysat 229E
- 103 lysat négatif
- 103 lysat 229E
- 143 lysat négatif
- 143 lysat 229E
- ◆— 146 lysat négatif
- ◆— 146 lysat 229E
- ▲— 156 lysat négatif
- ▲— 156 lysat 229E
- 166 lysat négatif
- 166 lysat 229E
- 170 lysat négatif
- 170 lysat 229E
- 212 lysat négatif
- 212 lysat 229E
- ◆— 219 lysat négatif
- ◆— 219 lysat 229E
- ▲— 221 lysat négatif
- ▲— 221 lysat 229E
- grippeB lysat négatif
- grippeB lysat 229E
- para1-3 lysat négatif
- para1-3 lysat 229E



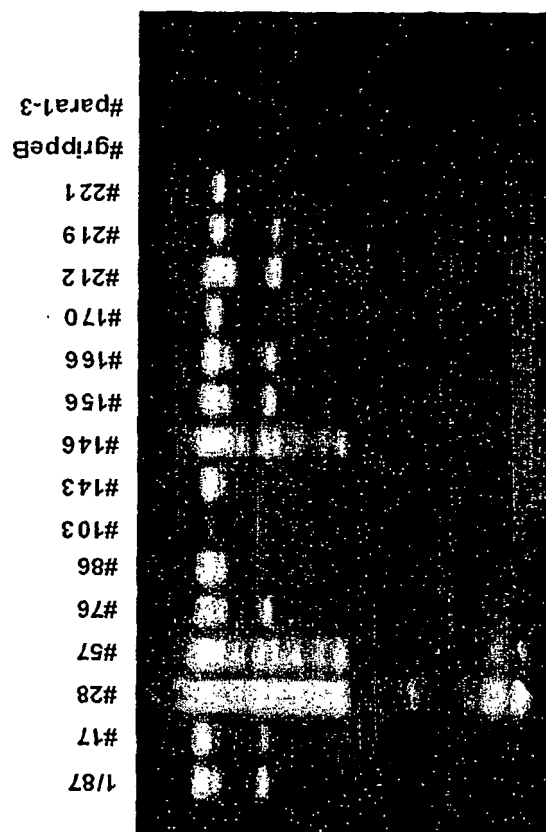


FIGURE 20

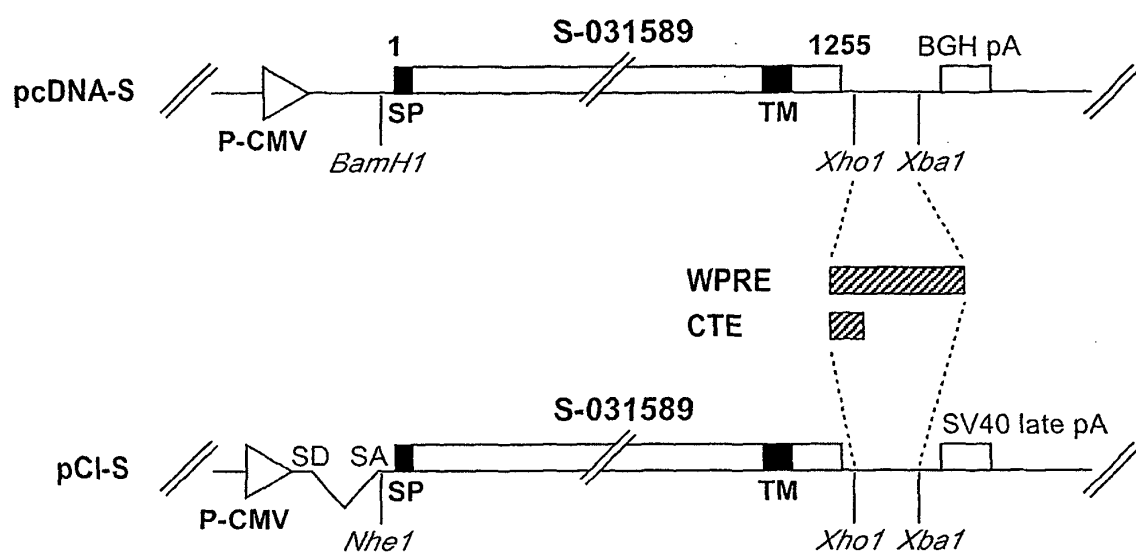


FIGURE 21

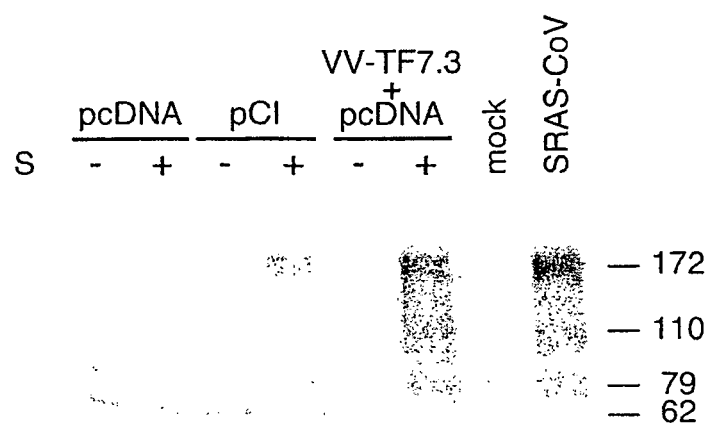
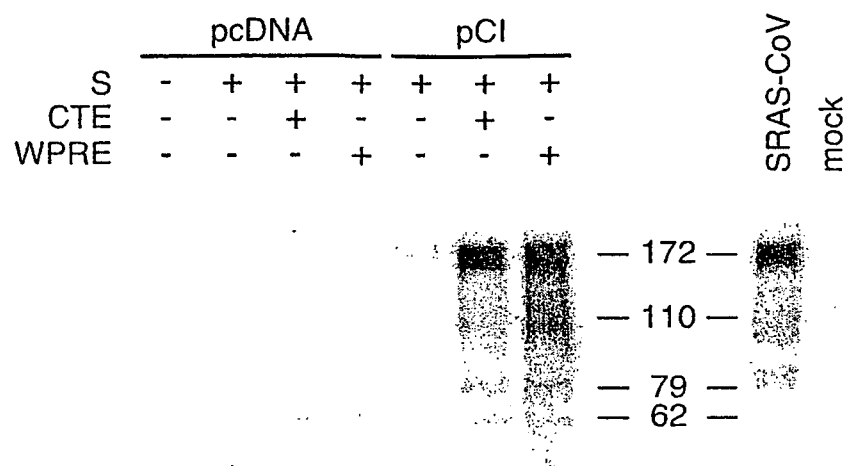


FIGURE 22

A.



B.

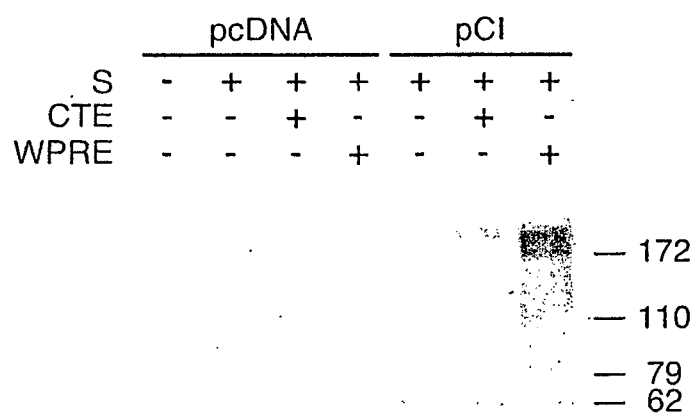


FIGURE 23

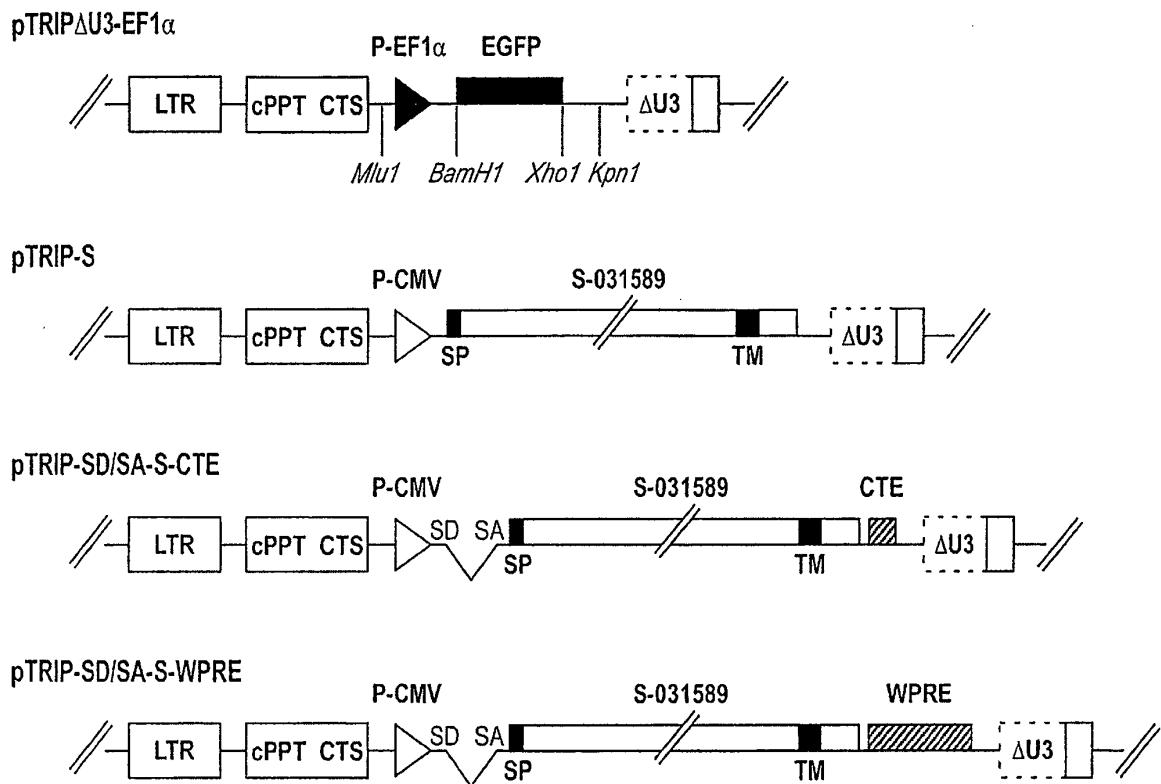


FIGURE 24

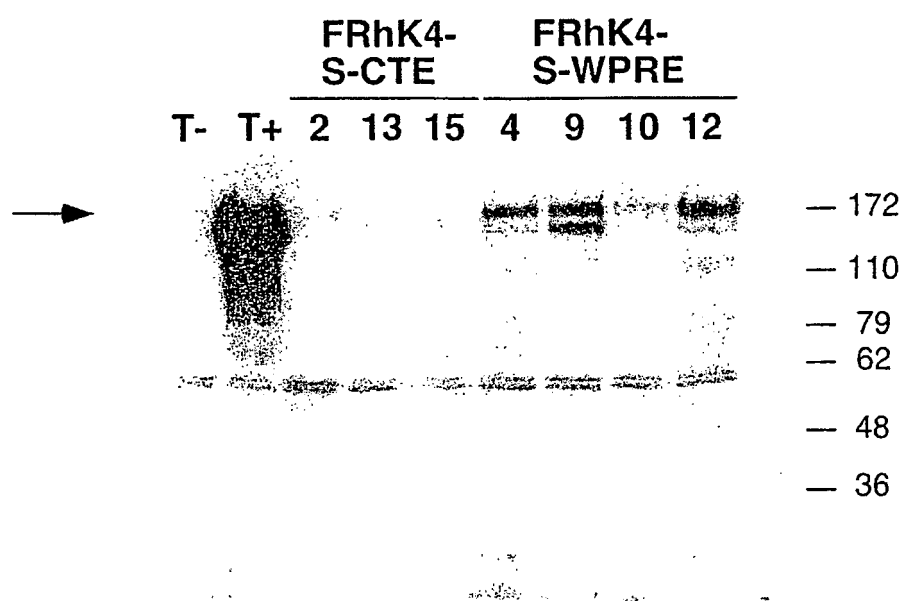


FIGURE 25

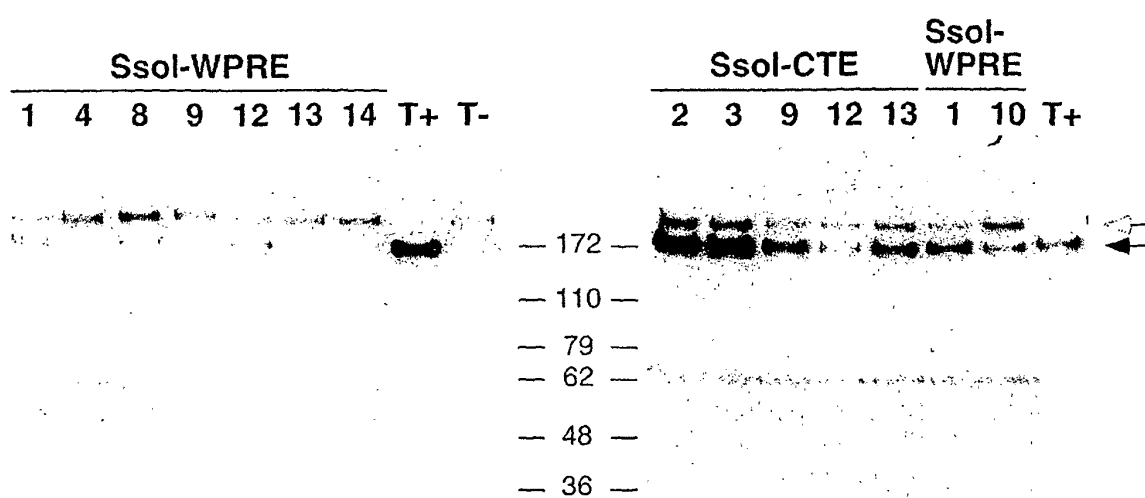
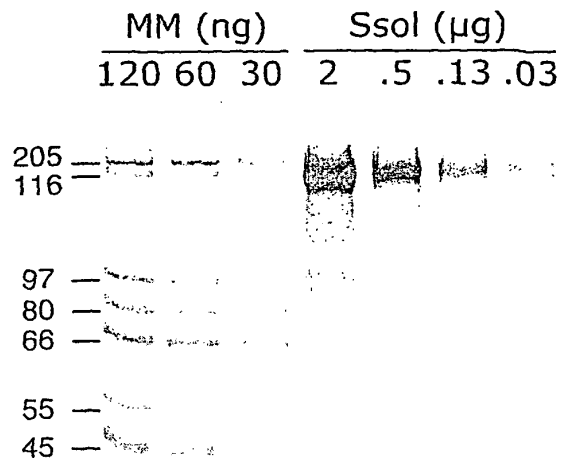
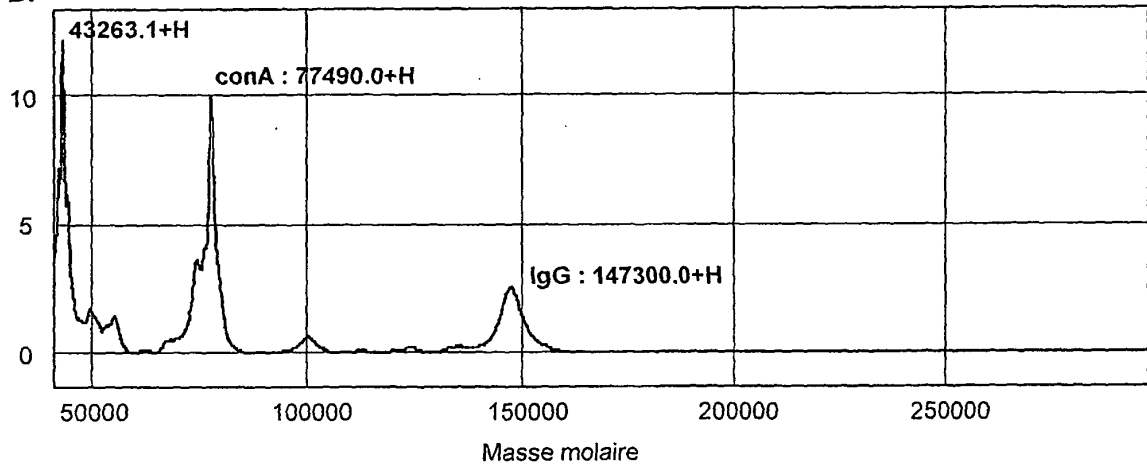


FIGURE 26

A.



B.



C.

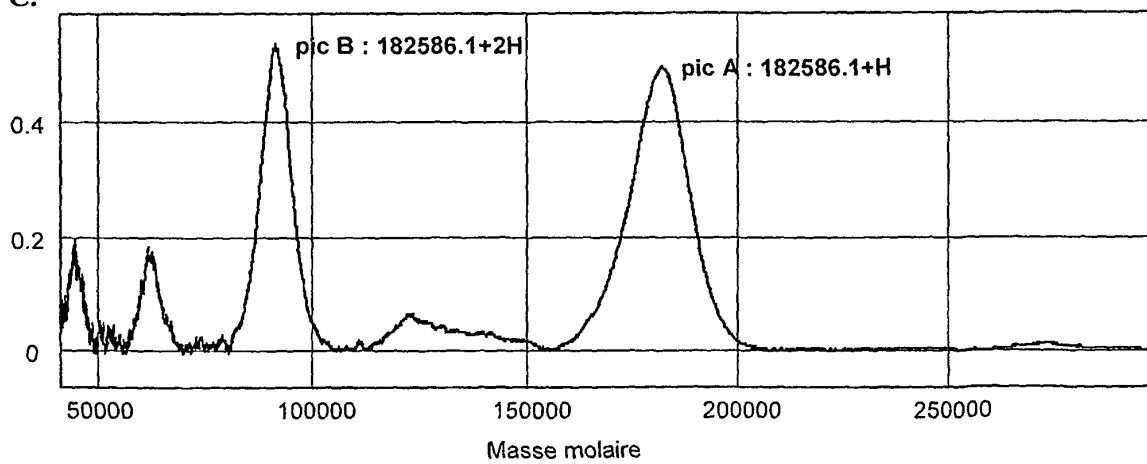


FIGURE 27 A-C



D.

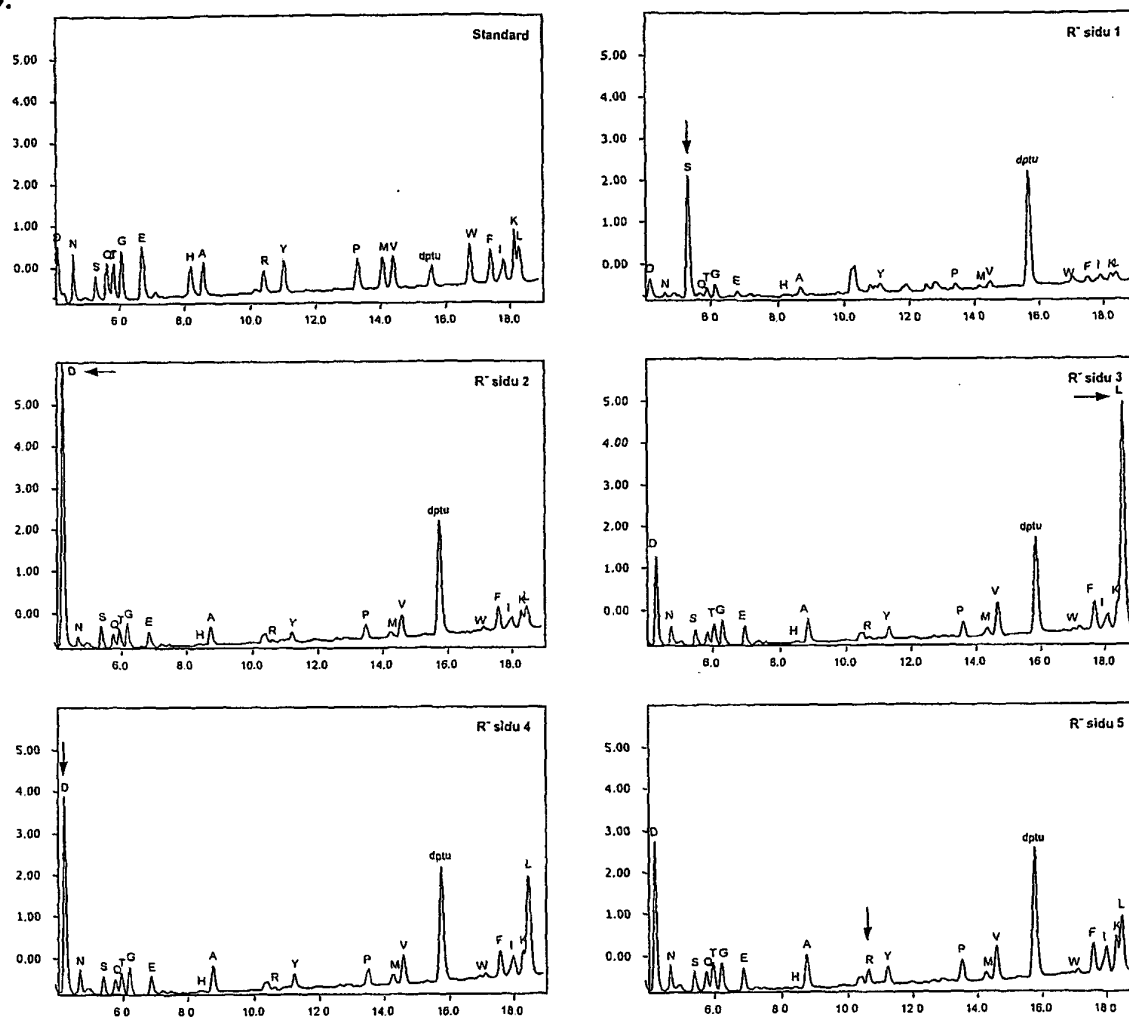
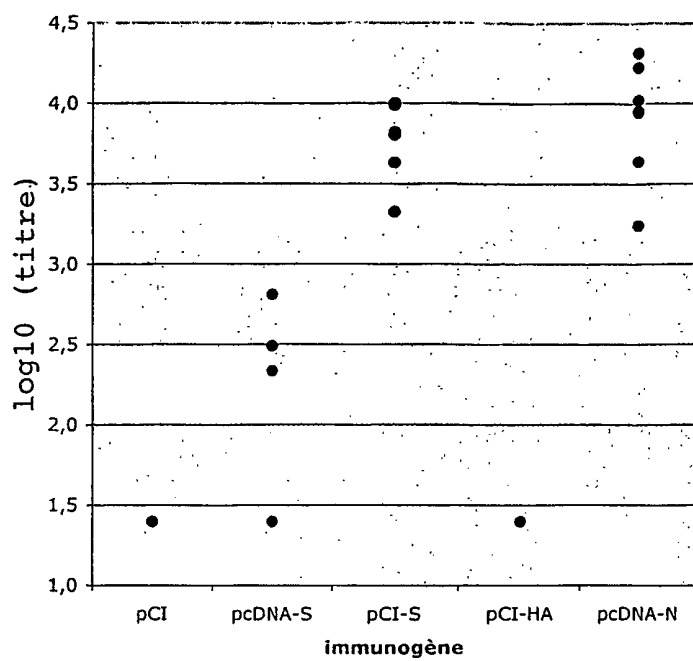


FIGURE 27 D

A.



B.

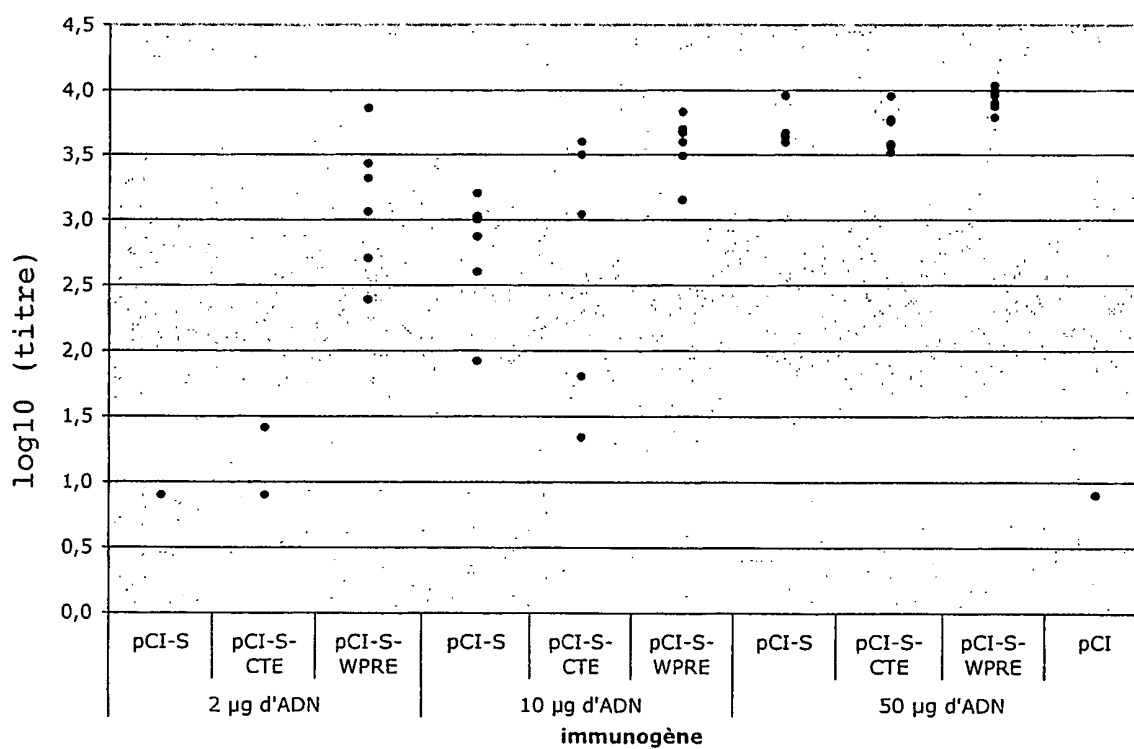
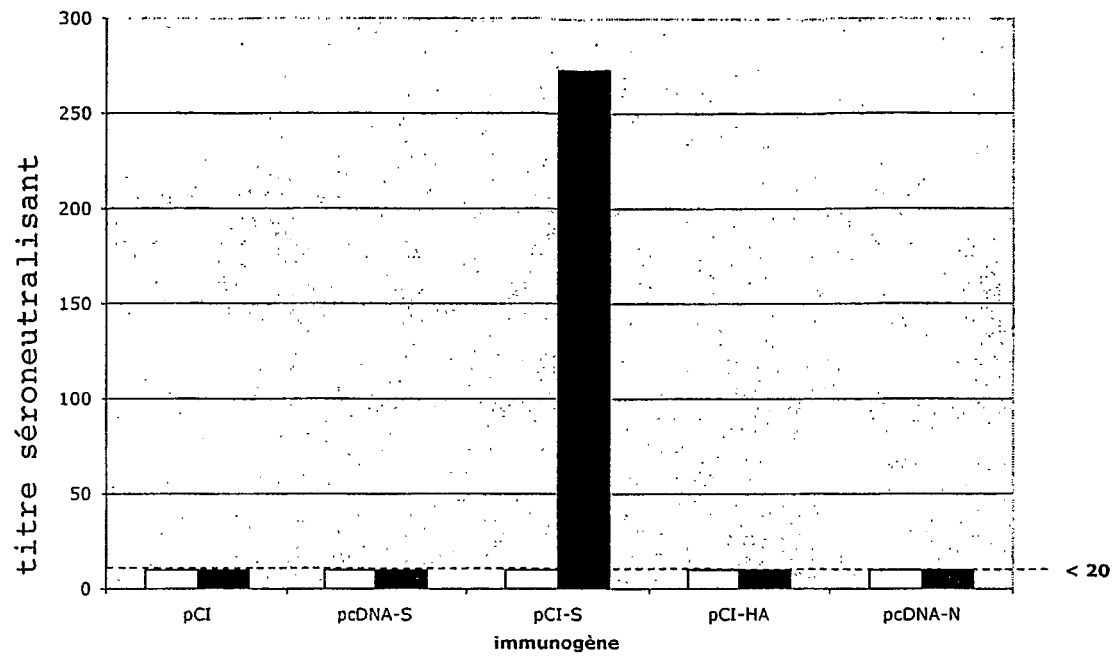


FIGURE 28

A.



B.

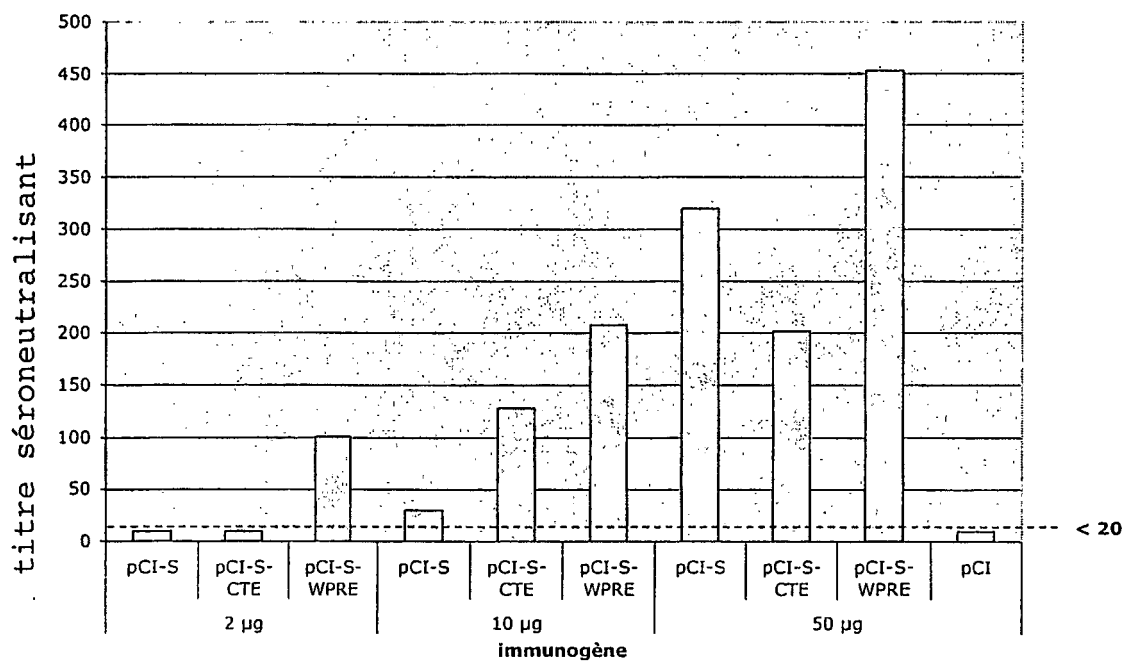


FIGURE 29

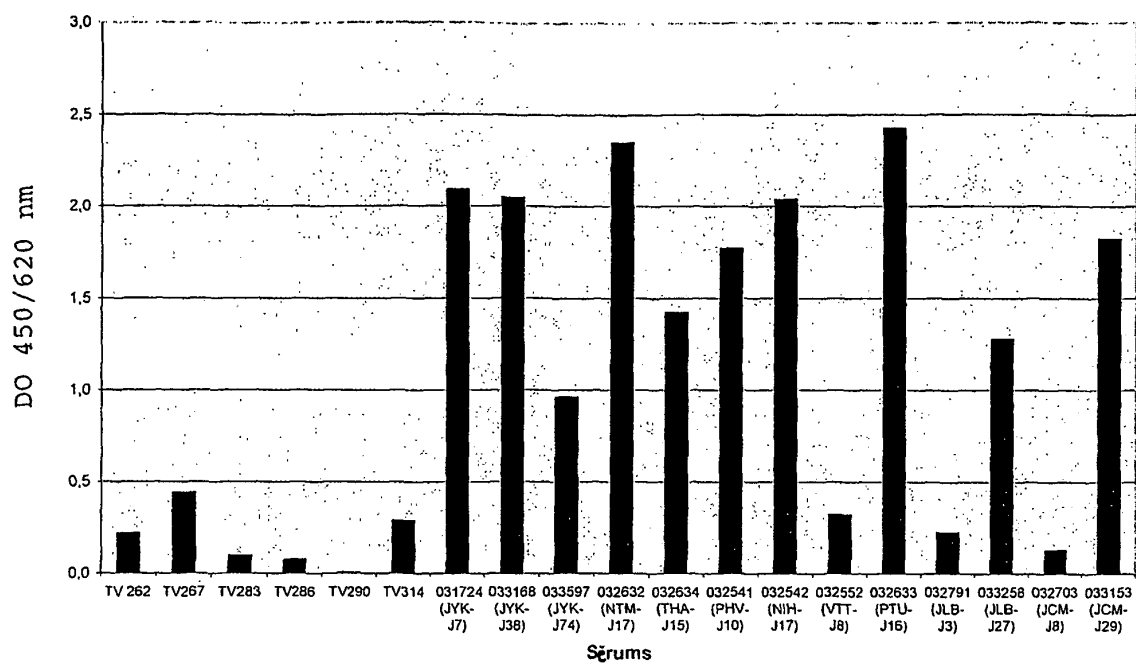


FIGURE 30

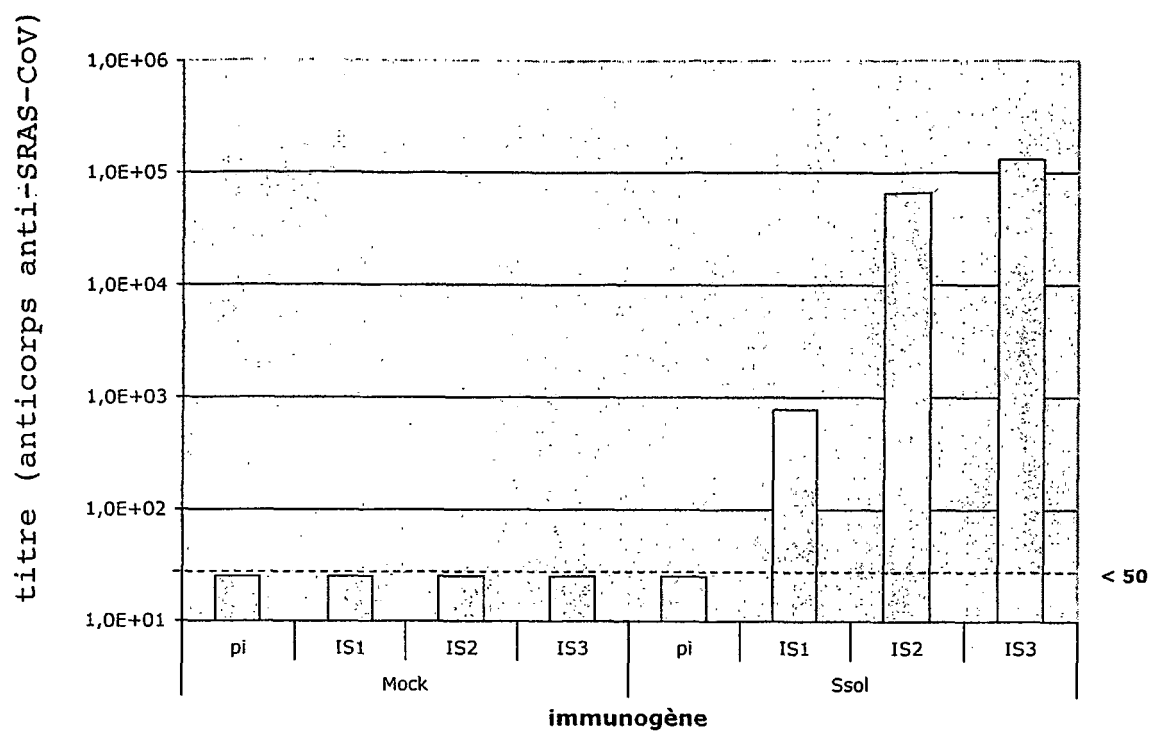


FIGURE 31

[illegible]

**FIGURE 32.1**

[illegible]

### FIGURE 32.2

I-3059	2697	GGTGCTGGCGCTGCTCTTCAAATACCTTTTGCTATGCAAATGGCATATAGGTTCAATGGC
S-040530	2620	"A"C"A"C"C"G"G"C"C"C"C"G"C"C"C"C"
I-3059	2757	ATTGGAGTTACCCAAAATGTTCTCTATGAGAACCAAAACAAATCGCCAACCAATTTAAC
S-040530	2680	"C"C"G"G"G"C"G"G"C"G"G"G"G"G"G"G"G"G"C"
I-3059	2817	AAGGCGATTAGTCAAATCAAGAATCACTTACAACAACATCAACTGCATTGGGCAAGCTG
S-040530	2740	"C"C"C"C"G"C"G"GAGC"G"C"C"CAGC"C"CC"
I-3059	2877	CAAGACGTTGTTAACCAGAATGCTCAAGCATTAAACACACTTGTAAACAACCTAGCTCT
S-040530	2800	"G"G"G"G"C"C"C"G"CC"G"C"G"G"G"G"G"AGC
I-3059	2937	AATTTTGGTGCAATTTCAAGTGTGCTAAATGATATCCTTTTCGCGACTTGATAAAGTCGAG
S-040530	2860	"C"C"C"C"CAGCTC"G"C"C"GAGCA"G"G"C"G"
I-3059	2997	GCGGAGGTACAAATTGACAGGCTAATTACAGGCAGACTTCAAAGCCTTCAAACCTATGTA
S-040530	2920	"C"A"G"G"C"C"G"C"AC"C"G"GTG"G"G"C"G"
I-3059	3057	ACACAACAATAATCAGGGCTGCTGAAATCAGGGCTTCTGCTAATCTTGCTGCTACTAAA
S-040530	2980	"C"G"G"G"A"C"C"G"C"CAGC"C"G"C"C"C"
I-3059	3117	ATGTCTGAGTGTGTTCTTGGACAATCAAAAAGAGTTGACTTTTGTGGAAAGGGTACCAC
S-040530	3040	"AGC"C"G"G"C"GAGC"G"G"G"C"C"C"TTTT"
I-3059	3177	CTTATGTCCTTCCCACAAGCAGCCCCGCATGGTGTGTCTTCTACATGTCACGTATGTG
S-040530	3100	"G"AG"C"G"C"C"C"G"G"G"C"G"C"C"
I-3059	3237	CCATCCCAGGAGAGGAACCTCACCACAGCGCCAGCAATTTGTCATGAAGGCAAAGCATAC
S-040530	3160	"TAG"C"C"C"C"C"C"C"G"G"C"
I-3059	3297	TTCCCTCGTGAAGGTGTTTTTGTGTTAATGGCACTTCTTGGTTTATTACACAGAGGAAC
S-040530	3220	"C"G"G"C"G"C"C"CAGC"C"C"C"C"
I-3059	3357	TTCTTTTCTCCACAAATAATTACTACAGACAATACATTTGTCTCAGGAAATGTGATGTC
S-040530	3280	"CAGC"C"G"C"C"C"C"C"G"C"C"C"
I-3059	3417	GTTATTGGCATCATTAACAACACAGTTTATGATCCTCTGCAACCTGAGCTTGACTCATTC
S-040530	3340	"G"C"C"TTTT"C"G"C"C"C"G"C"G"AGC"
I-3059	3477	AAAGAAGAGCTGGACAAGTACTTCAAAAATCATACATCACCAGATGTTGATCTTGGCGAC
S-040530	3400	"G"G"AAAA"G"C"C"C"C"C"G"C"G"TT
I-3059	3537	ATTTCAGGCATTAACGCTTCTGTCGTCACATTCAAAAAGAAATTGACCGCCTCAATGAG
S-040530	3460	"CAGC"C"C"G"G"C"G"G"G"C"A"A"G"C"A
I-3059	3597	GTCGCTAAAAATTTAAATGAATCACTCATTGACCTTCAAGAATTGGGAAAATATGAGCAA
S-040530	3520	"G"C"G"CC"G"C"GAGC"G"C"G"GC"C"G"C"
I-3059	3657	TATATTAAATGGCCTTGGTATGTTTGGCTCGGCTTATTGCTGGACTAATTGCCATCGTC
S-040530	3580	"C"C"G"C"G"C"G"G"C"C"C"G"C"
I-3059	3717	ATGGTTACAATCTTGCTTTGTTGCATGACTAGTTGTTGCAGTTGCCTCAAGGGTGCATGC
S-040530	3640	"G"C"C"G"C"C"C"C"TC"G"AC"C"
I-3059	3777	TCTTGTGGTTCTTGCTGCAAGTTTGTATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTG
S-040530	3700	AGC"CAGC"C"C"AGC"C"G"G"C"
I-3059	3837	AAATTACATTACACATAAACGAACCTTATGATTGTTTATGAGATTTTTTACTCTTGGAT
S-040530	3760	"GC"G"C"G"TCGA"
I-3059	3897	CAATTACTGCACAGCCAGTAAAAATTGACAATGCTTCTCCTGCAAGT
S-040530		

FIGURE 32.3



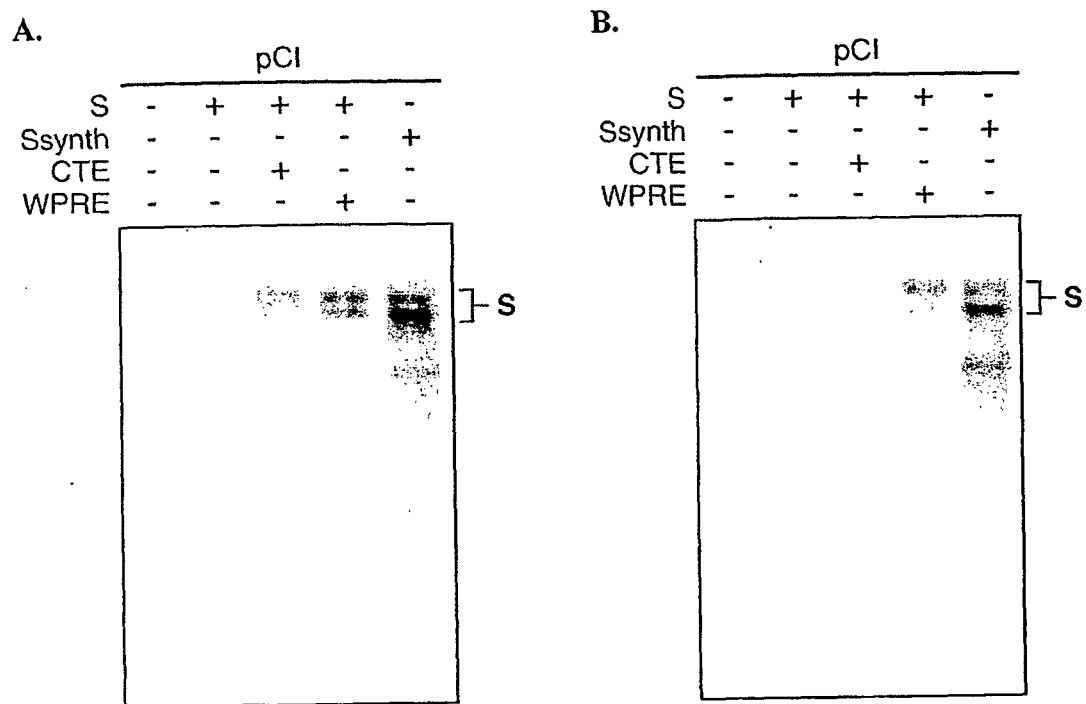
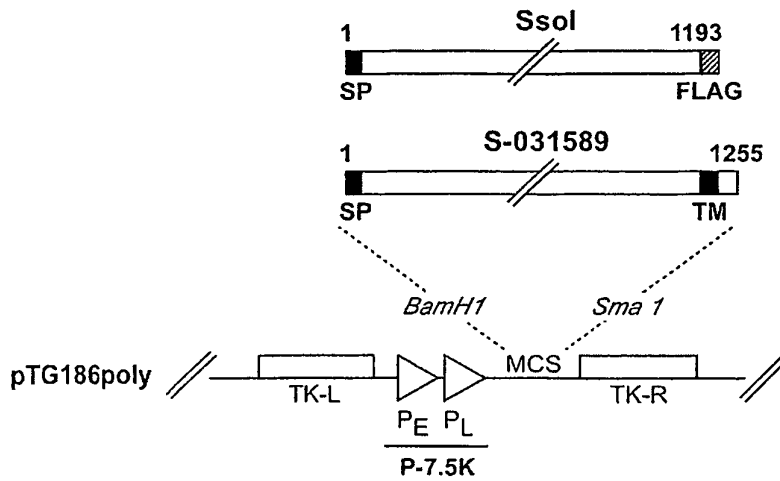
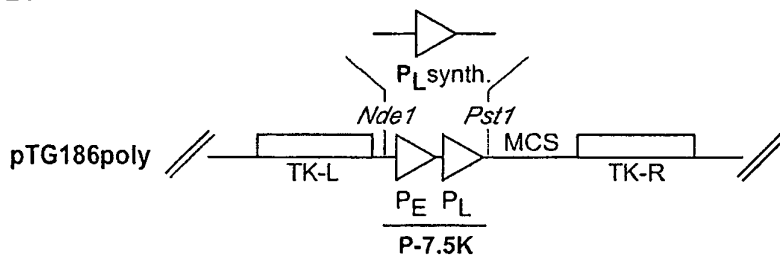


FIGURE 33

A.



B.



C.

$\frac{\text{CATATG}}{\text{NdeI}}$  AGC  $\frac{[\text{T}]_{20}\text{GGCATATAAATA}}{\text{promoteur 480}}$  GACTC  $\frac{\text{GGCGCGCC}}{\text{AseI}}$  AT  $\frac{\text{CTGCAG}}{\text{PstI}}$

FIGURE 34 A-C

D.

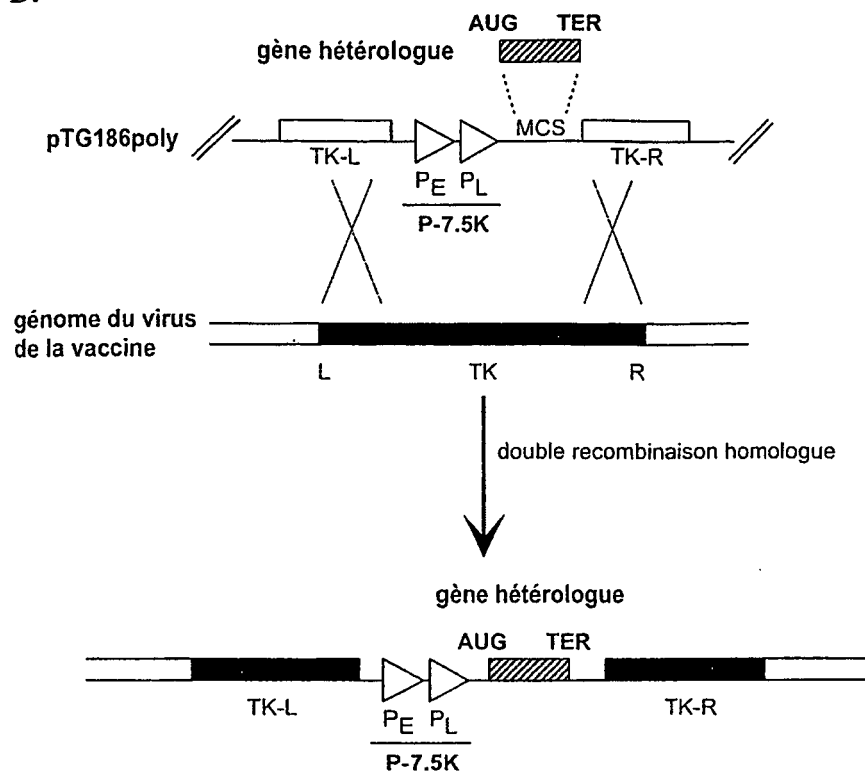


FIGURE 34 D

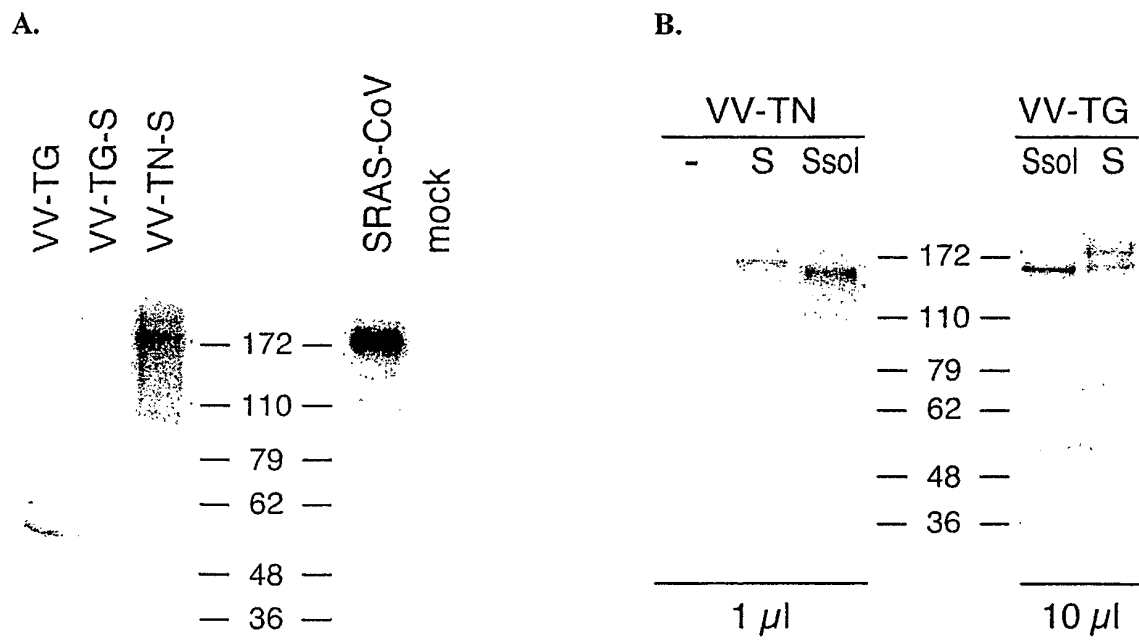
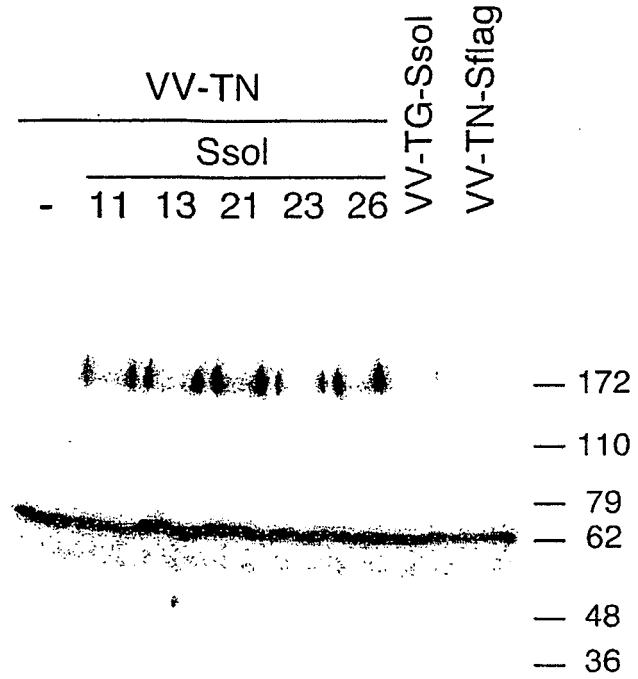


FIGURE 35

A.



B.

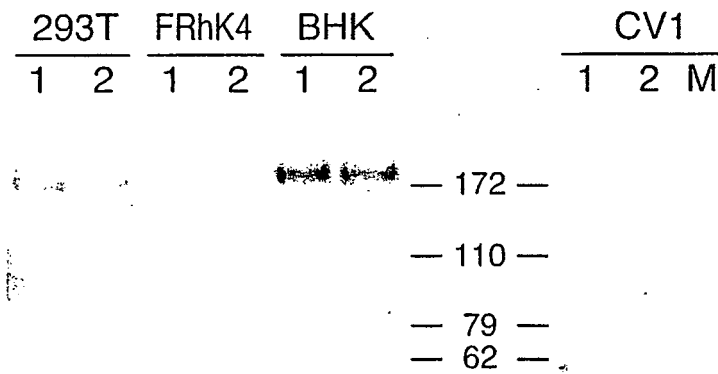


FIGURE 36

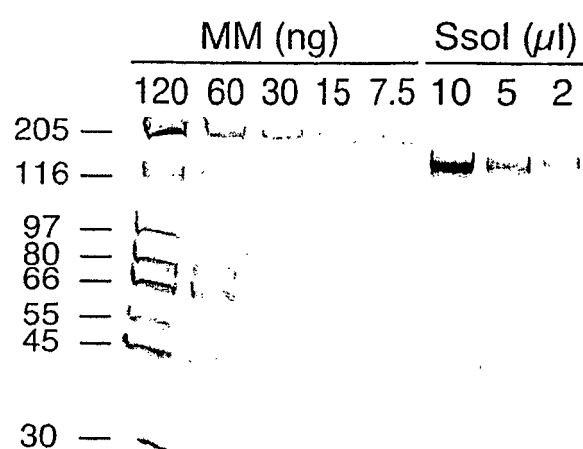


FIGURE 37

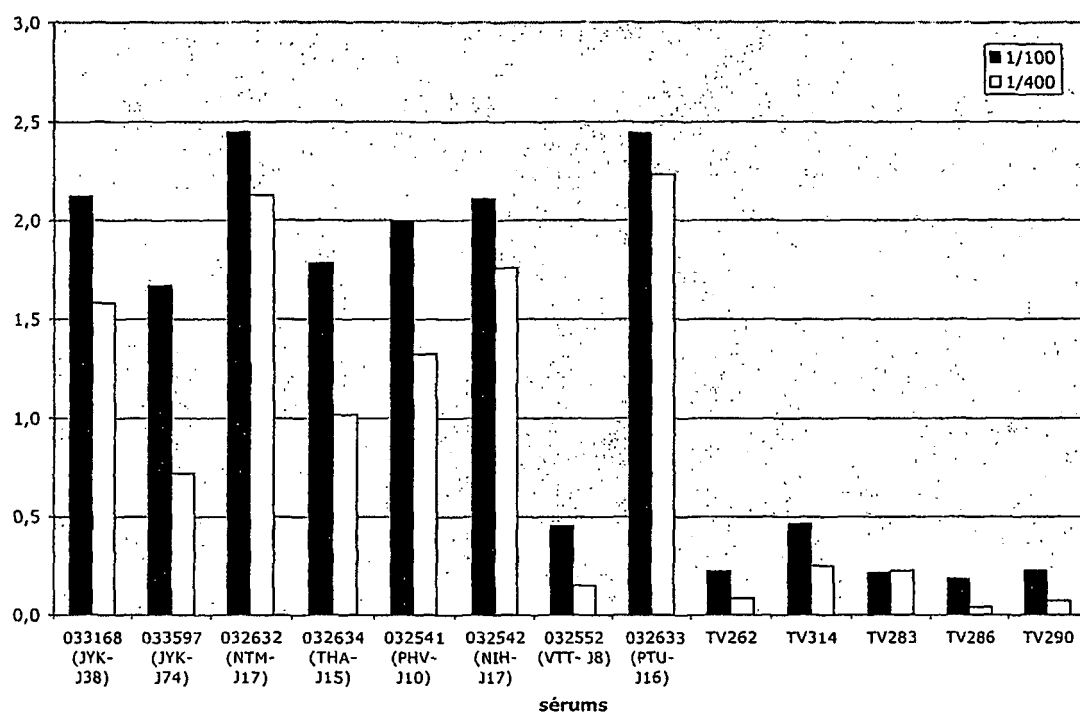
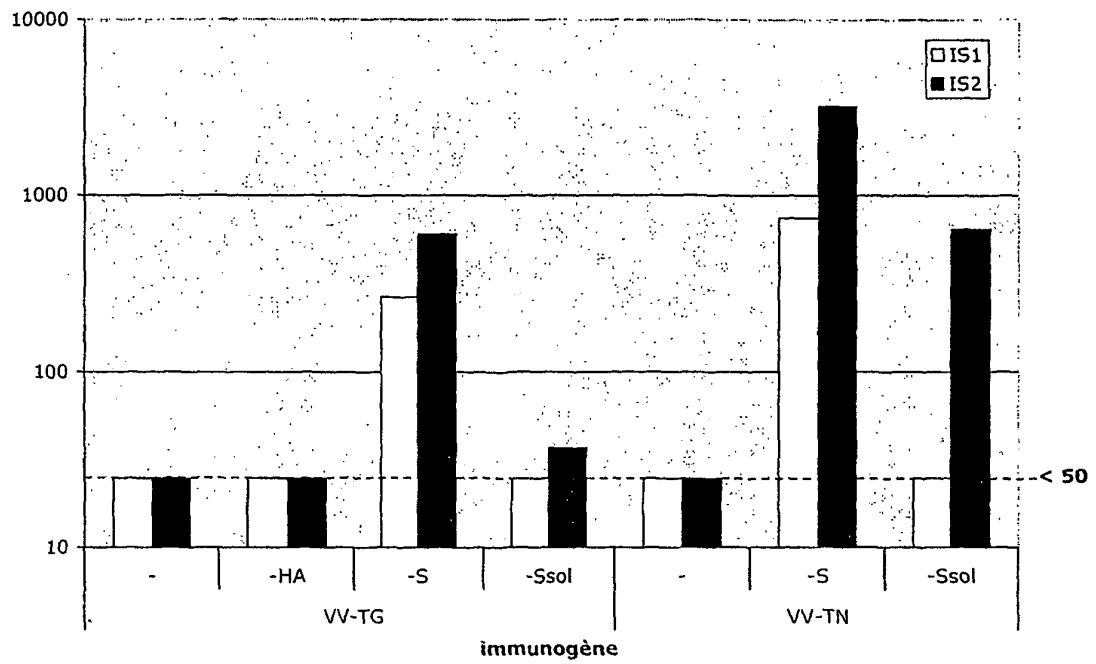


FIGURE 38

A.



B.

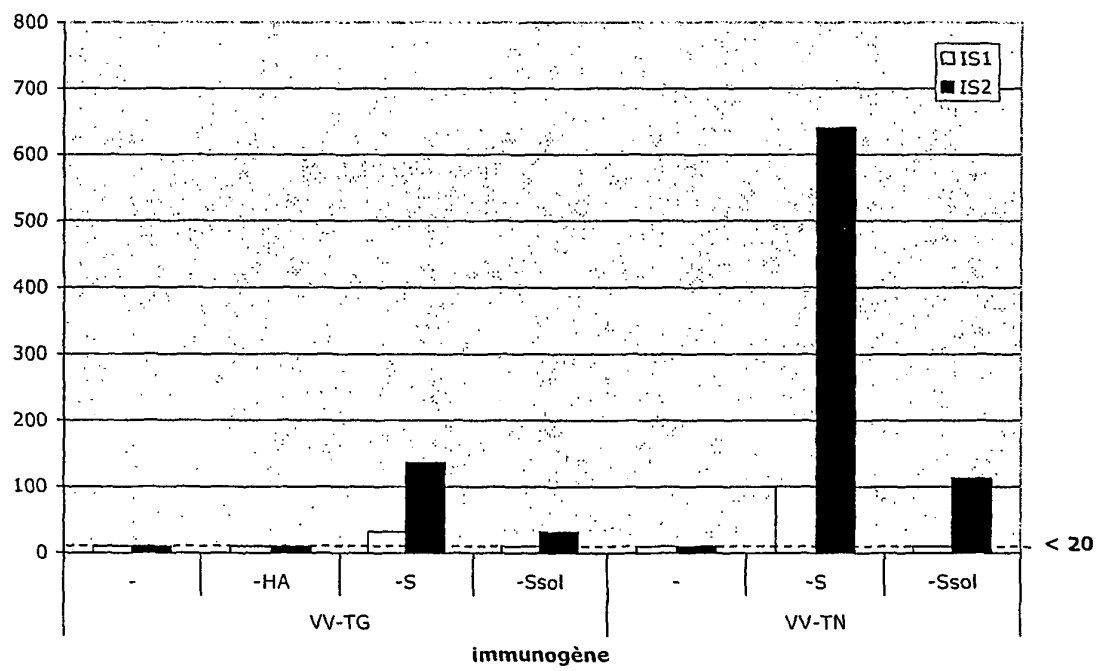
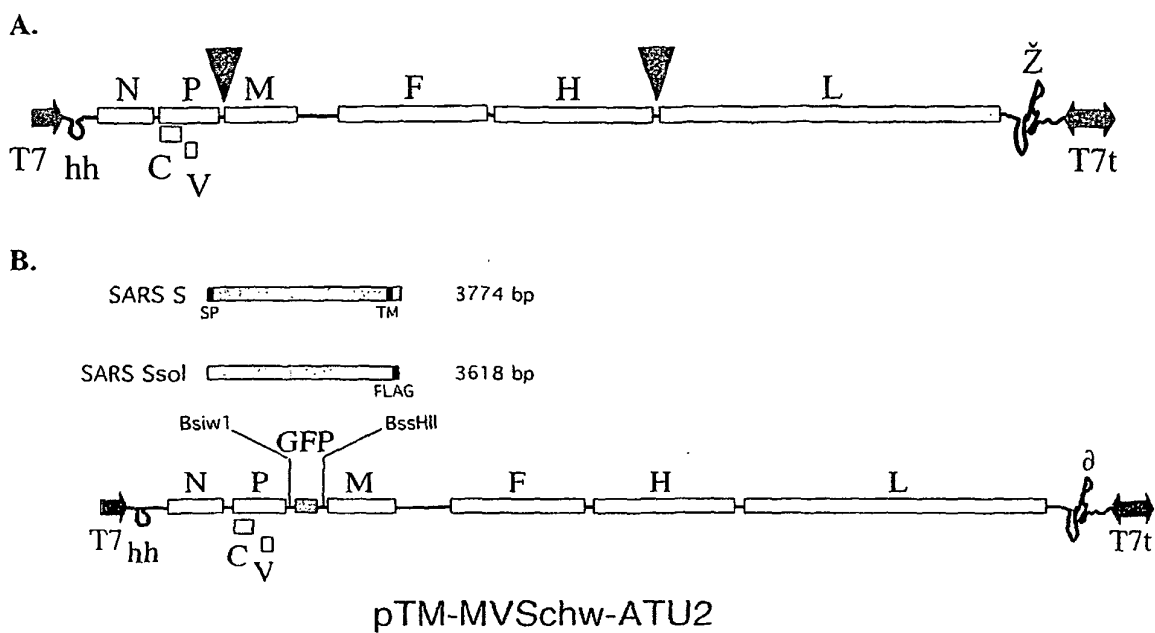


FIGURE 39





**FIGURE 40**

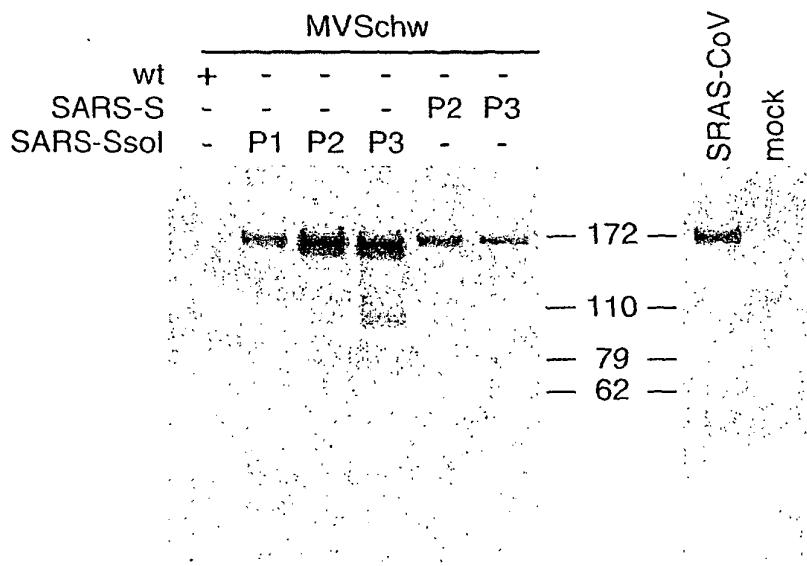


FIGURE 41

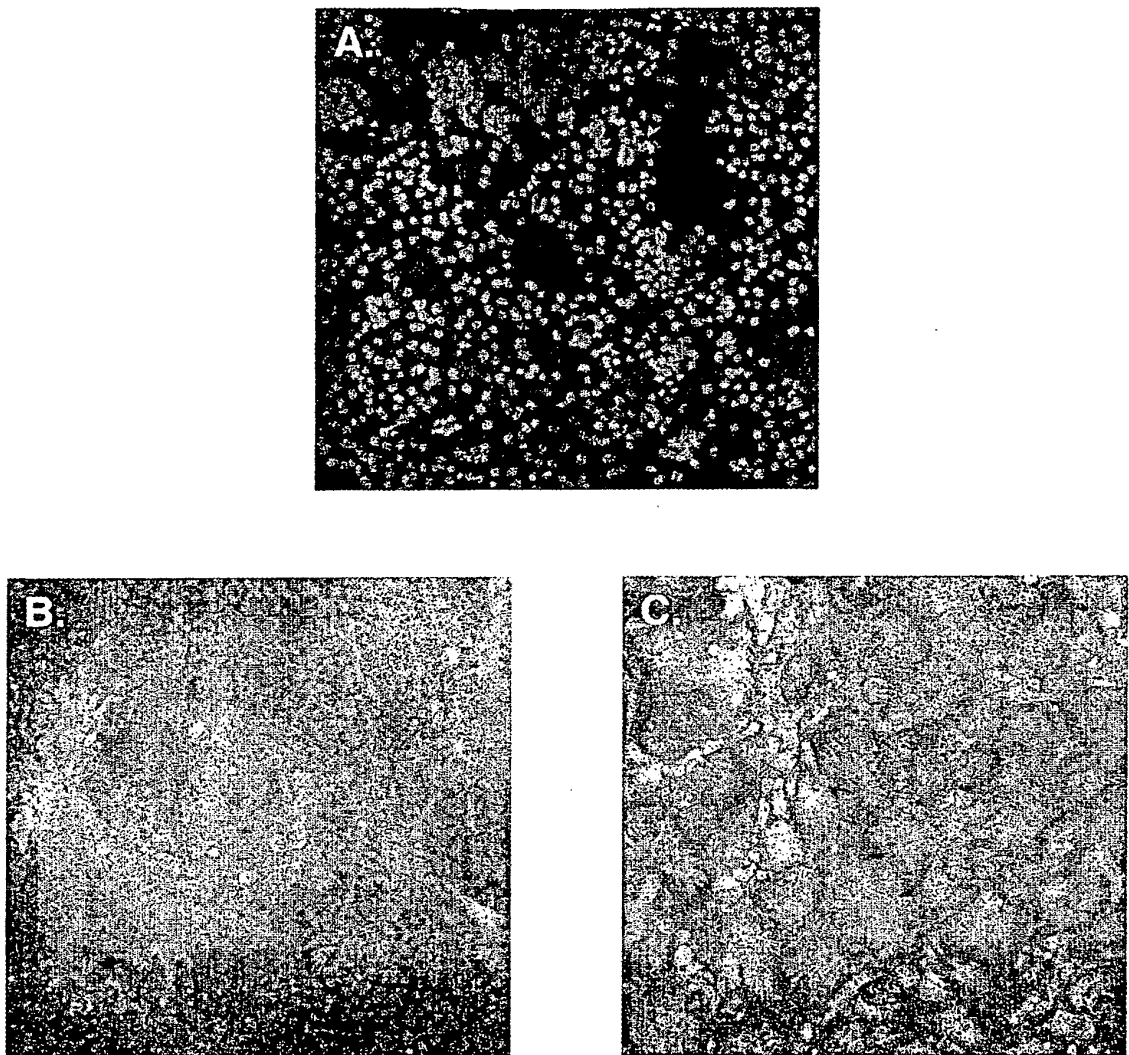


FIGURE 42

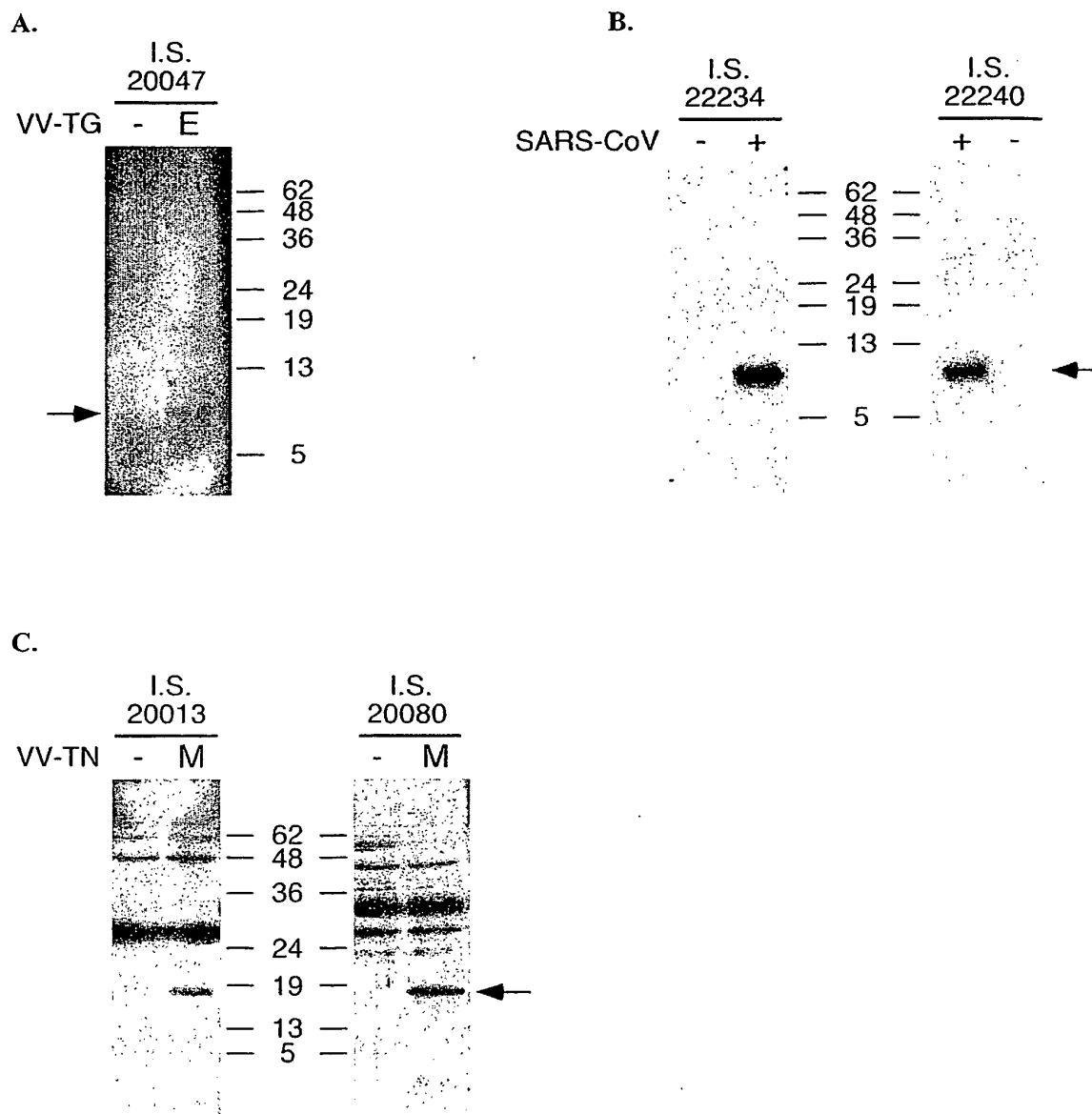


FIGURE 43

## RÉFÉRENCES CITÉES DANS LA DESCRIPTION

Cette liste de références citées par le demandeur vise uniquement à aider le lecteur et ne fait pas partie du document de brevet européen. Même si le plus grand soin a été accordé à sa conception, des erreurs ou des omissions ne peuvent être exclues et l'OEB décline toute responsabilité à cet égard.

### Documents brevets cités dans la description

- US 4816567 A [0097]
- WO 9845332 A [0098]
- EP 02291551 A [0402]
- EP 02291550 A [0402]
- FR 0314152 [0417]
- FR 0314151 [0417]

### Littérature non-brevet citée dans la description

- **T.G.KSIAZEK et al.** *New England Journal of Medicine*, 2003, vol. 348, 1319-1330 [0013]
- **C. DROSTEN et al.** *New England Journal of Medicine*, 2003, vol. 348, 1967-1976 [0013]
- **Peiris et al.** *Lancet*, 2003, vol. 361, 1319 [0013]
- **A. MARRA et al.** *Science*, 01 Mai 2003, vol. 300, 1399-1404 [0014]
- **A. ROTA et al.** *Science*, 2003, vol. 300, 1394-1399 [0014]
- **Holmes. J.C.I.**, 2003, vol. 111, 1605-1609 [0017]
- **Frederick M. AUSUBEL.** *Current Protocols in Molecular Biology*. Wiley and Son Inc., Library of Congress, 2000 [0094]
- **John E. Cologan.** *Current Protocols in Immunology*. Wiley and Son Inc. Library of Congress, 2000 [0094]
- **E. Howell ; D Lane.** *Antibodies : A Laboratory Manual*. Cold Spring Harbor Laboratory, 1988 [0094]
- **Köhler ; Milstein.** *Nature*, 1975, vol. 256, 495-497 [0097]
- **Winter ; Milstein.** *Nature*, 1991, vol. 349, 293-299 [0099]
- **Sizun et al.** *J. Virol. Met.*, 1998, vol. 72, 145-152 [0147]
- **Nielsen et al.** *Protein Engineering*, 1997, vol. 10, 1-6 [0147] [0337]
- **Sonnhammer et al.** *Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology*. AAAI Press, 1998, 175-182 [0147]
- **Sirven et al.** *Mol. Ther.*, 2001, vol. 3, 438-448 [0147] [0327] [0328]
- **Combredet.** *Journal of Virology*, 2003, vol. 77, 11546-11554 [0147] [0403]
- **Krogh et al.** *J. Mol. Biol.*, 2001, vol. 305 (3), 567-580 [0239]
- **von Heijne.** *J. Mol. Biol.*, 1992, vol. 225, 487-494 [0239]
- **Zhou ; Zhou.** *Protein Science*, 2003, vol. 12, 1547-1555 [0239]
- **Nalcane P.K. ; Kawaoi A.** Peroxydase-labeled antibody, a new method of conjugation. *The Journal of Histochemistry and Cytochemistry*, 1974, vol. 22 (23), 1084-1091 [0263]
- **Nakane et al.** *J. of Histo and cytochemistry*, 1974, vol. 22, 1084-1091 [0305]
- **Sirven.** *Mol. Ther.*, 2001, vol. 3, 438-448 [0321]
- **Zennou et al.** *Cell*, 2000, vol. 101, 173-185 [0327] [0328]
- **Zennou.** *Cell*, 2000, vol. 101, 173-185 [0334]
- **Yang.** *Nature*, 2004, vol. 428, 561-564 [0354]
- **Buchholz.** *PNAS*, 2004, vol. 101, 9804-9809 [0354]
- **Kieny.** *Biotechnology*, 1986, vol. 4, 790-795 [0377]
- **Kieny.** *Nature*, 1984, vol. 312, 163-166 [0379]
- **Hilleman.** *Vaccine*, 2002, vol. 20, 651-665 [0399]
- **Combredet.** *J. Virol.*, 2003, vol. 77, 11546-11554 [0399] [0402]
- **Calain ; Roux.** *J. Virol.*, 1993, vol. 67, 4822-4830 [0404]
- **Schneider et al.** *Virology*, 1997, vol. 227, 314-322 [0404]
- **Radecke.** *Embo J.*, 1995, vol. 14, 5773-5784 [0407]
- **Parks.** *J. Virol.*, 1999, vol. 73, 3560-3566 [0407]