



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2009년02월19일
 (11) 등록번호 10-0884214
 (24) 등록일자 2009년02월11일

(51) Int. Cl.
C12N 15/64 (2006.01) *C12N 15/66* (2006.01)
C12N 15/09 (2006.01)
 (21) 출원번호 10-2007-7007725
 (22) 출원일자 2007년04월04일
 심사청구일자 2007년04월05일
 번역문제출일자 2007년04월04일
 (65) 공개번호 10-2007-0053798
 (43) 공개일자 2007년05월25일
 (86) 국제출원번호 PCT/KR2004/002274
 국제출원일자 2004년09월07일
 (87) 국제공개번호 WO 2006/028302
 국제공개일자 2006년03월16일
 (56) 선행기술조사문헌
 Virus Research,93(2):199-209
 Virus Research,93(1); 25-30
 전체 청구항 수 : 총 19 항

(73) 특허권자
주식회사 마크로젠
 서울 금천구 가산동 60-24 월드메르디앙벤처센터 10층
 (72) 발명자
김연수
 서울특별시 서초구 서초4동 유원아파트 101-1607
이수경
 서울 강남구 압구정1동 현대아파트 84-303
김종필
 대전광역시 서구 월평3동 하나로아파트 107-1205
 (74) 대리인
이현실, 장성구

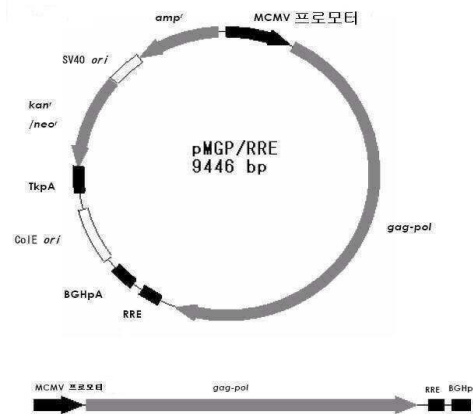
심사관 : 김지윤

(54) CAEV-계 벡터 시스템

(57) 요약

본 발명은 비-분열 및 분열 세포로 핵산을 전달하는 데 유용한 염소 관절염 뇌염 바이러스-계 벡터 및 벡터 시스템에 관한 것이다. 상기 벡터 시스템을 사용하여 비-분열 및 분열 세포에 핵산을 전달하는 방법 또한 개시하고 있다.

대표도 - 도2a



특허청구의 범위

청구항 1

1) (가) CAEV 유전체의 1 내지 511번째 뉴클레오티드로 이루어진 5' 비번역 영역, 및 (나) 상기 5' 비번역 영역의 3' 말단에 연결된 CAEV *gag*-코딩 서열의 1 내지 X번째 뉴클레오티드(이때, X는 613 미만임)로 필수적으로 이루어진 CAEV 패키징 서열; 및

2) 상기 CAEV 패키징 서열에 작동가능하게 연결된, 폴리아데닐화 반응, RNA 수송, 역전사 및 통합(integration)으로부터 선택된 하나 이상에 요구되는 시스-작용 요소를 포함하는, 재조합 염소 관절염 뇌염 바이러스(CAEV)-계 전이 벡터.

청구항 2

제 1 항에 있어서,

상기 X가 327인 것을 특징으로 하는, 전이 벡터.

청구항 3

제 1 항에 있어서,

상기 *gag*-코딩 서열의 시작 코돈이 *gag* 단백질의 번역을 방지하도록 변이된 것을 특징으로 하는, 전이 벡터.

청구항 4

제 3 항에 있어서,

상기 시작 코돈이 TAG로 변이된 것을 특징으로 하는, 전이 벡터.

청구항 5

제 3 항에 있어서,

상기 시작 코돈 ATG의 21 염기쌍 하류에 위치하는 *gag*-코딩 서열의 ATG 코돈이 *gag* 단백질의 번역을 방지하도록 변이된 것을 특징으로 하는, 전이 벡터.

청구항 6

제 1 항에 있어서,

상기 시스-작용 요소가 RRE(*rev*-반응 요소) 영역, U3가 결실된 CAEV 3' LTR, 및 이중의 프로모터로 이루어진 군으로부터 선택된 하나 이상의 서열을 포함하는 것을 특징으로 하는, 전이 벡터.

청구항 7

제 6 항에 있어서,

상기 이중의 프로모터가 인간 거대세포바이러스 주 극초기 프로모터(HCMV MIEP)인 것을 특징으로 하는, 전이 벡터.

청구항 8

제 1 항에 있어서,

상기 벡터가 도 3c에 제시된 pCAH/SINd1의 구조를 갖는 것을 특징으로 하는, 전이 벡터.

청구항 9

제 8 항에 있어서,

상기 벡터가 서열번호: 68과 최소 70%의 상동성을 갖는 것을 특징으로 하는, 전이 벡터.

청구항 10

제 1 항에 있어서,

이종의 프로모터에 작동가능하게 연결된 이종의 유전자를 포함하는 전사 카세트를 추가로 포함하는 것을 특징으로 하는, 전이 벡터.

청구항 11

제 1 항의 전이 벡터 및 패키징 벡터 시스템을 포함하며, 이때 상기 패키징 벡터 시스템은 CAEV *gag-pol*-코딩 서열 및 RRE를 포함하는 제1 폴리뉴클레오티드, 및 바이러스 외피-코딩 서열을 포함하는 제2 폴리뉴클레오티드를 포함하는 것을 특징으로 하는, 재조합 CAEV-계 벡터 시스템.

청구항 12

제 11 항에 있어서,

상기 전이 벡터가 이종의 프로모터에 작동가능하게 연결된 이종의 유전자를 포함하는 전사 카세트를 추가로 포함하는 것을 특징으로 하는, 벡터 시스템.

청구항 13

제 11 항에 있어서,

상기 바이러스 외피-코딩 서열이 비-CAEV 외피-코딩 서열인 것을 특징으로 하는, 벡터 시스템.

청구항 14

제 13 항에 있어서,

상기 비-CAEV 외피-코딩 서열이 VSV-G(vesicular stomatitis virus G; 수포성 구내염 바이러스 G) 당단백질-코딩 서열 또는 GaLV(gibbon ape leukemia virus; 원숭이 백혈병 바이러스) 외피 단백질-코딩 서열인 것을 특징으로 하는, 벡터 시스템.

청구항 15

제 11 항에 있어서,

상기 벡터 시스템이 *rev*-코딩 서열을 포함하는 제3 폴리뉴클레오티드 서열을 추가로 포함하는 것을 특징으로 하는, 벡터 시스템.

청구항 16

제 11 항에 있어서,

상기 벡터 시스템이 *vif*-코딩 서열을 포함하는 제4 폴리뉴클레오티드를 추가로 포함하는 것을 특징으로 하는, 벡터 시스템.

청구항 17

제 11 항에 있어서,

상기 패키징 벡터 시스템이 CAEV 패키징 서열을 포함하지 않는 것을 특징으로 하는, 벡터 시스템.

청구항 18

제 11 항에 있어서,

상기 제1 폴리뉴클레오티드를 포함하는 제1 벡터 및 상기 제2 폴리뉴클레오티드를 포함하는 제2 벡터를 포함하는 것을 특징으로 하는, 벡터 시스템.

청구항 19

제 12 항의 재조합 CAEV-계 벡터 시스템으로 세포를 형질감염시켜 제조된 복제-결합 벡터 입자.

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

명세서

기술분야

<1> 본 발명은 폴리뉴클레오티드(polynucleotide) 전달에 유용한 렌티바이러스 벡터, 및 더 구체적으로 폴리뉴클레오티드를 비분열 및 분열 세포들로 전달하는데 유용한 염소 관절염 뇌염 바이러스-계 벡터(caprine arthritis encephalitis virus)에 관한 것이다.

배경기술

<2> 렌티바이러스는 분열 세포 뿐 아니라 비분열 세포도 감염시킬 수 있는 레트로바이러스의 서브그룹이다. 렌티바이러스 유래의 벡터들은 분열 및 비분열 세포의 유전체(genome)에 안정적으로 유전자를 통합(integration)하여 장기간 유전자 발현을 매개할 수 있기 때문에 외래 유전자를 표적 세포로 전달하기 위한 이상적인 도구이다 (Gilbert and Wong-Staal, 2001; Mitrophanous et al., 1999; Naldini et al., 1996; Sauter and Gasmi, 2001).

<3> 렌티바이러스는 비-영장류(예: 고양이 면역결핍 바이러스(FIV), 소 면역결핍 바이러스(BIV), 말 감염 바이러스

(EIAV), 염소 관절염 뇌염 바이러스(CAEV) 및 비스나 바이러스) 뿐 아니라, 영장류(예: 사람 및 원숭이 면역결핍 바이러스(HIV-1, HIV-2, SIV))를 포함하는 여러 척추동물 종들로부터 분리된 바 있다. 이들 중, HIV 및 SIV가 현재 가장 잘 알려져 있다. 그러나, 박테리아 전염성 및 병원성 형태로 재조합될 수 있는 가능성으로 인해, 이러한 시스템을 인간에서 사용하는 것은 심각한 안전성 문제를 일으킨다. 따라서, 유전자 치료에서는 비-영장류 렌티바이러스를 사용하는 것이 바람직하다.

- <4> 비-영장류 렌티바이러스 박테리아들 중에는, FIV(Curran 및 Nolan, 2002) 및 EIAV[US 2001/0044149] 유래 박테리아들이 가장 잘 연구되었으며, 기타 비-영장류 박테리아들에 대해서는 거의 진전된 바가 없다.
- <5> CAEV는, 모든 렌티바이러스들이 그렇듯이, 최종적으로 분화된 세포 및 비분열 세포 뿐 아니라 분열 세포들에서도 감염 및 복제가 가능하다. CAEV의 생물학적 몇가지 특징들은 본 바이러스가 유전자 전이/치료 박테리아로 개발 시키기에 흥미로운 후보임을 보여준다. 첫째, CAEV의 일반 숙주는 염소류이며, CAEV에 의해 인간이 감염된 경우는 보고된 바 없다. 둘째, CAEV 유전체는 렌티바이러스들 중 HIV-1과 계통분류학적으로 가장 거리가 있다. 셋째, CAEV의 유전체 구성은 다른 렌티바이러스들과 비교해 상대적으로 간단하다. CAEV 유전체는 세 개의 구조 유전자(*gag*, *pol*, *env*) 및 세 개의 조절/약세서리 유전자(*vif*, *tat* 및 *rev*)를 포함한다.
- <6> 그러나, 이러한 장점들에도 불구하고, CAEV-계 전달 시스템을 개발시키기 위한 노력들은 성공하지 못했으며, 그 결과 CAEV-계 유전자 전달 시스템의 사용을 비현실적으로 만드는 안전하지 못하고 비효율적인 재조합 바이러스 박테리아 생산 시스템만을 얻었다.
- <7> 1998년에, L. 셀리-라칼 등(L. Mselli-Lakhali *et al.*)은 제1세대 CAEV-계 박테리아 시스템을 보고하였으나, 상기 시스템의 바이러스 역가(즉, 10-187 TU/ml)는 유효한 수준 이하였다. 상기 저자들은 이러한 비효율성이 세포질 내로의 유전체 RNA 축적의 부족 및 박테리아 RNA의 낮은 패키징 효율 때문이라고 설명하였다. 상기 연구의 다른 문제점으로는 패키징 시스템으로서 전염성 야생형 바이러스("헬퍼 바이러스")를 사용하는 점인데, 이로 인해 인간에 적용하기에는 거의 실용 가치가 없다.
- <8> 따라서, 광범위한 분열 및 비분열 세포로의 유전자 전이를 매개할 수 있는 안전하고 효과적인 CAEV-계 렌티바이러스 박테리아 시스템이 요구되고 있다.

<9> 발명의 개요

- <10> 본 발명은, 넓게는 외래 폴리뉴클레오티드를 표적 세포로 전달하는데 유용한 CAEV-계 렌티바이러스 박테리아 입자(particles)의 생산에 관한 것이다. 이러한 박테리아 입자는 항-바이러스, 항-종양 및/또는 유전자 치료에 사용될 수 있다.
- <11> 본 발명은 상술된 CAEV-계 박테리아 생산 시스템에서 사용되는 전이 박테리아를 제공하며, 상기 전이 박테리아는,
- <12> (a) (i) CAEV 5' LTR과 CAEV *gag*-코딩 서열 사이의 비번역 영역, 및
- <13> (ii) 상기 비번역 영역의 3' 말단에 연결된 CAEV *gag*-코딩 서열의 1 내지 X번째 뉴클레오티드(이때, X는 613 미만)
- <14> 로 필수적으로 이루어진, CAEV 패키징 서열; 및
- <15> (b) 상기 패키징 서열과 작동가능하게 연관된, 폴리아데닐화 반응, RNA 수송, 역전사 및 통합(integration)에 요구되는 *시스*-작용 요소(*cis*-acting elements)
- <16> 를 포함한다.
- <17> 본 발명의 일 실시양태에서, X는 60, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575 및 600으로 이루어진 군으로부터 선택된다.
- <18> 본 발명의 다른 실시양태에서, X는
- <19> (a) X는 25 초과 600 미만,
- <20> (b) X는 25 초과 500 미만,
- <21> (c) X는 25 초과 400 미만,
- <22> (d) X는 25 초과 300 미만,

- <23> (e) X는 25 초과 200 미만,
- <24> (f) X는 50 초과 600 미만,
- <25> (g) X는 50 초과 500 미만,
- <26> (h) X는 50 초과 400 미만,
- <27> (i) X는 50 초과 300 미만,
- <28> (j) X는 50 초과 200 미만,
- <29> (k) X는 75 초과 600 미만,
- <30> (l) X는 75 초과 500 미만,
- <31> (m) X는 75 초과 400 미만,
- <32> (n) X는 75 초과 300 미만,
- <33> (o) X는 75 초과 200 미만,
- <34> (p) X는 100 초과 600 미만,
- <35> (q) X는 100 초과 500 미만,
- <36> (r) X는 100 초과 400 미만,
- <37> (s) X는 100 초과 300 미만,
- <38> (t) X는 100 초과 200 미만,
- <39> (u) X는 125 초과 600 미만,
- <40> (v) X는 125 초과 500 미만,
- <41> (w) X는 125 초과 400 미만,
- <42> (x) X는 125 초과 300 미만,
- <43> (y) X는 125 초과 200 미만,
- <44> (z) X는 150 초과 600 미만,
- <45> (aa) X는 150 초과 500 미만,
- <46> (bb) X는 150 초과 400 미만,
- <47> (cc) X는 150 초과 300 미만,
- <48> (dd) X는 150 초과 200 미만,
- <49> (ee) X는 200 초과 600 미만,
- <50> (ff) X는 200 초과 500 미만,
- <51> (gg) X는 200 초과 400 미만,
- <52> (hh) X는 200 초과 300 미만,
- <53> (ii) X는 200 초과 200 미만,
- <54> (jj) X는 250 초과 600 미만,
- <55> (kk) X는 250 초과 500 미만,
- <56> (ll) X는 250 초과 400 미만, 및
- <57> (mm) X는 250 초과 300 미만
- <58> 으로 이루어진 군으로부터 선택된다.

- <59> 다른 실시양태에서, X는 40 초과 613 미만이다.
- <60> 다른 실시양태에서, X는 57 초과 613 미만이다.
- <61> 또 다른 실시양태에서, X는 약 327이다.
- <62> 본 발명의 일 실시양태에서, *gag*-코딩 서열의 시작 코돈은 *gag* 단백질의 번역을 방지하도록 변이된다. 또 다른 실시양태에서, 시작 코돈은 TAG로 변이된다.
- <63> 본 발명의 전이 벡터의 다른 실시양태에서, *gag*-코딩 서열의 ATG 코돈은 시작 코돈 ATG의 하류 X번째 염기쌍에 위치하며, 이때 상기 시작 코돈은 *gag* 단백질의 번역을 방지하도록 변이되고, 이때 상기 X는 30 미만이다. 또 다른 실시양태에서 X는 약 21이다.
- <64> 본 발명의 전이 벡터는 RRE 영역을 추가로 포함할 수 있다.
- <65> 본 발명의 다른 실시양태에서, 전이 벡터는 U3 영역이 삭제된 CAEV 3' LTR을 포함한다.
- <66> 본 발명의 전이 벡터는 이종의 프로모터(heterologous promoter)를 추가로 포함할 수 있다. 본 발명의 일 실시양태에서, 이종의 프로모터는 인간 거대세포바이러스 주 극초기 프로모터(human cytomegalovirus major immediate early promoter; HCMV MIEP)이다. 또 다른 실시양태에서, 전이 벡터는 pCAH/SINd1(서열번호: 68)이다.
- <67> 본 발명의 전이 벡터는 이종의 프로모터(예: 인간 거대세포바이러스 주 극초기 프로모터 HCMV MIEP, 또는 쥐 거대세포바이러스 주 극초기 프로모터 MCMV MIEP)에 작동가능하게 연결된 이종의 관심 폴리뉴클레오티드를 포함하는 전사 카세트를 추가로 포함할 수 있다. 이러한 전이 벡터는 바이러스 입자내로 관심 폴리뉴클레오티드가 결합하는 것을 가능하게 함으로써, 내부에 폴리뉴클레오티드를 포함하는 감염된 숙주 세포의 수를 증대시키기 위한 수단을 제공한다.
- <68> 본 발명은 또한 외래 폴리뉴클레오티드를 포유류 세포에 전달하는데 유용한 CAEV-계, 복제-결합 벡터 입자를 생산하는 CAEV-계 렌티바이러스 벡터 시스템을 제공한다. 상기 벡터 입자는 포유류 세포를 감염 및 형질도입시킬 수 있다. 벡터 시스템은, 상술된 전이 벡터; 및 GAEV *gag-pol*-코딩 서열 및 RRE를 포함하는 제1 폴리뉴클레오티드, 및 바이러스 외피 코딩 서열을 포함하는 제2 폴리뉴클레오티드를 포함하는, 패키징 벡터 시스템을 포함한다.
- <69> 일 실시양태에서, 상기 제2 폴리뉴클레오티드는 비-CAEV *env*-코딩 서열을 포함한다. 일 실시양태에서, 상기 제2 폴리뉴클레오티드는 VSV-G- 또는 GaLV-코딩 서열을 포함한다.
- <70> 다른 실시양태에서, CAEV 벡터 시스템은 *rev*-코딩 서열을 포함하는 제3 폴리뉴클레오티드 서열을 포함한다.
- <71> 다른 실시양태에서, CAEV 벡터 시스템은 *vif*-코딩 서열을 포함하는 제4 폴리뉴클레오티드 서열을 포함한다.
- <72> 또 다른 실시양태에서, 상술된 각 CAEV 벡터 시스템의 제1 폴리뉴클레오티드는 CAEV *gag-pol*-코딩 서열에 작동가능하게 연결된 이종의 조절 서열을 추가로 포함한다.
- <73> 또 다른 실시양태에서, 상술된 CAEV 벡터 시스템의 제2 폴리뉴클레오티드는 상기 바이러스 외피-코딩 서열에 작동가능하게 연결된 이종의 조절 서열을 추가로 포함한다.
- <74> 또 다른 실시양태에서, 제3 폴리뉴클레오티드는 *rev*-코딩 서열에 작동가능하게 연결된 이종의 조절 서열을 추가로 포함한다.
- <75> 또 다른 실시양태에서, 제4 폴리뉴클레오티드는 *vif*-코딩 서열에 작동가능하게 연결된 이종의 조절 서열을 추가로 포함한다.
- <76> 본 발명의 일 실시양태에서, CAEV 벡터 시스템은 적격(competent) CAEV 패키징 서열이 없는 패키징 벡터 시스템을 포함한다. 또 다른 실시양태에서, 패키징 벡터 시스템은 CAEV 유전체의 스플라이싱 제공 위치(splice donor site)와 *gag* 시작 코돈 사이의 5' 말단 영역이 결여되어 있다.
- <77> 일 실시양태에서, CAEV 벡터 시스템은 제1 폴리뉴클레오티드를 포함하는 제1 벡터 및 제2 폴리뉴클레오티드를 포함하는 제2 벡터를 포함한다. 다른 실시양태에서, 벡터 시스템은 제1 폴리뉴클레오티드를 포함하는 제1 벡터, 제2 폴리뉴클레오티드를 포함하는 제2 벡터, 및 제3 폴리뉴클레오티드를 포함하는 제3 벡터를 포함한다. 다른 실시양태에서, 벡터 시스템은 제1 폴리뉴클레오티드를 포함하는 제1 벡터, 제2 폴리뉴클레오티드를 포함하

는 제2 벡터, 제3 폴리뉴클레오티드를 포함하는 제3 벡터, 및 제4 폴리뉴클레오티드를 포함하는 제4 벡터를 포함한다. 상기 제3 벡터는 pHYK/rev(서열번호: 75)일 수 있고, 상기 제4 벡터는 pHYK/vif(서열번호: 76)일 수 있다.

- <78> 또 다른 실시양태에서, 벡터 시스템은 제1 폴리뉴클레오티드, 제3 폴리뉴클레오티드 및 제4 폴리뉴클레오티드를 포함하는 제1 벡터, 및 제2 폴리뉴클레오티드를 포함하는 제2 벡터를 포함한다.
- <79> 일 실시양태에서, CAEV 벡터 시스템의 제1 벡터는 CAEV *gag*-코딩 서열 및 이종의 프로모터에 작동가능하게 연결된 RRE를 포함한다. 상기 프로모터는 MCMV MIEP일 수 있다. 또 다른 실시양태에서, CAEV 벡터 시스템은 제1 벡터 pMGP/RRE(서열번호: 77)를 포함한다.
- <80> 일 실시양태에서, CAEV 벡터 시스템의 제2 벡터는 이종의 프로모터에 작동가능하게 연결된 VSV-G-코딩 서열이다. 상기 프로모터는 HCMV MIEP일 수 있다. 상기 제2 벡터는 베타 글로빈 인트론(beta globin intron)을 추가로 포함할 수 있다. 또 다른 실시양태에서, CAEV 벡터 시스템은 제2 벡터 pHGVSV-G(서열번호: 74)를 포함한다.
- <81> 일 실시양태에서, CAEV 벡터 시스템의 제2 벡터는 이종의 프로모터에 작동가능하게 연결된 GaLV *env*-코딩 서열이다. 상기 프로모터는 MCMV MIEP일 수 있다. 상기 제2 벡터는 진핵성 연장 인자-1 알파 인트론(eukaryotic elongation factor-1 alpha intron)을 추가로 포함할 수 있다. 또 다른 실시양태에서, CAEV 벡터 시스템은 제2 벡터 pMYKEF-1/*env*(서열번호: 72)를 포함한다.
- <82> 본 발명의 다른 태양은 포유류 세포를 감염시키는데 유용한 CAEV-계 렌티바이러스 벡터 입자를 생산하는 방법이다. 상기 방법은 (a) CAEV-계 입자 생산에 적합한 조건하에 상술된 벡터 시스템으로 세포를 형질감염시키는 단계(이때, 상기 벡터 입자는 감염- 및 형질도입-적격이며, 복제-결합임), 및 (b) 상기 벡터 입자를 회수하는 단계를 포함한다.
- <83> 본 발명은 또한 CAEV-계 렌티바이러스 벡터 입자 및 선택적으로 담체를 포함하는 조성물을 제공하며, 이때 상기 벡터 입자는 상술된 방법에 의해 생산된다.
- <84> 본 발명은 또한 상술된 전이 벡터 또는 CAEV-계 렌티바이러스 벡터 시스템을 포함하는 키트를 제공한다.
- <85> 본 발명은 또한 CAEV *gag-pol*-코딩 서열 및 RRE, 및 선택적으로 바이러스 *env*-코딩 서열을 포함하는 패키징 세포를 제공한다. 상기 패키징 세포는 *rev*-코딩 및/또는 *vif*-코딩 서열을 추가로 포함할 수 있다. 상기 세포는 전이 벡터의 RNA 형태를 감염- 및 형질도입-적격이면서 복제-결합인 벡터 입자내로 패키징하는데 유용하다.
- <86> 일 실시양태에서, 벡터 시스템은 상술된 제1 폴리뉴클레오티드를 포함하는 세포를 포함한다. 상기 벡터 시스템은 상술된 제3 및/또는 제4 폴리뉴클레오티드를 추가로 포함할 수 있다.
- <87> 다른 실시양태에서, 벡터 시스템은 상술된 제1 폴리뉴클레오티드 및 제2 폴리뉴클레오티드를 포함하는 세포를 포함한다. 상기 벡터 시스템은 상술된 제3 및/또는 제4 폴리뉴클레오티드를 추가로 포함할 수 있다.
- <88> 다른 실시양태에서, 벡터 시스템은 CAEV *gag-pol*-코딩 서열 및 RRE를 포함하는 제1 벡터를 포함하는 세포를 포함한다. 상기 제1 벡터는 *rev*-코딩 및/또는 *vif*-코딩 서열을 추가로 포함할 수 있다. 한편, 상기 세포는 CAEV *gag-pol*-코딩 서열 및 RRE를 포함하는 제1 벡터, *rev*-코딩 서열을 포함하는 제2 벡터, 및/또는 *vif*-코딩 서열을 포함하는 제3 벡터를 포함할 수 있다.
- <89> 일부 실시양태에서, 벡터 시스템은 CAEV *gag-pol*-코딩 서열 및 RRE를 포함하는 제1 벡터, 및 바이러스 *env*-코딩 서열을 포함하는 제2 벡터를 포함하는 세포를 포함한다. 상기 제1 벡터는 *rev*-코딩 및/또는 *vif*-코딩 서열을 추가로 포함할 수 있다. 다른 한편으로, 상기 세포는 CAEV *gag-pol*-코딩 서열 및 RRE를 포함하는 제1 벡터, 바이러스 *env*-코딩 서열을 포함하는 제2 벡터, 및 선택적인 *rev*-코딩 서열을 포함하는 제3 벡터 및/또는 *vif*-코딩 서열을 포함하는 제4 벡터를 포함할 수 있다.
- <90> 본 발명의 다른 태양은 폴리뉴클레오티드 또는 폴리펩타이드를 포유류 세포로 전달하거나 상기 폴리펩타이드를 코딩하는 폴리뉴클레오티드를 복제하는 방법이며, 상기 방법은 상기 폴리뉴클레오티드를 포유류 세포의 유전체내로 통합할 수 있는 조건 및 선택적으로 상기 폴리뉴클레오티드에 의해 코딩된 상기 폴리펩타이드가 증식할 수 있는 조건하에, 상기 포유류 세포를 상술된 벡터 입자와 접촉시키는 단계를 포함한다. 상기 포유류 세포는 분열 세포, 비-분열 세포 또는 CD34+ 줄기 세포일 수 있다. 폴리뉴클레오티드 또는 폴리펩타이드를 포유류 세포로 전달하는 방법 또는 상기 폴리펩타이드를 코딩하는 폴리뉴클레오티드 분자를 복제하는 방법은, 세포를 벡터

입자와 접촉시키기 전에 포유동물로부터 세포를 분리하는 단계를 추가로 포함할 수 있다. 상기 방법은 세포를 벡터 입자와 접촉시킨 후에 배양액에서 상기 세포를 증식시키는 단계를 추가로 포함할 수 있다. 상기 방법은 접촉된 세포를 증식시키기 전 또는 후에 세포를 포유동물에 재도입하는 단계를 추가로 포함할 수 있다.

- <91> 본 발명은 또한, 전달된 폴리뉴클레오티드에 의해 코딩된 폴리펩타이드가 척추동물에서의 생물학적 반응을 검출 또는 유도하기에 충분한 양으로 척추동물에서 발현되도록 하기 위해, 이종의 관심 폴리뉴클레오티드를 포함하는 CAEV-계 렌티바이러스 벡터 입자를 척추동물에 투여하는 단계(이때, 상기 벡터 입자는 상술된 방법에 의해 생산됨)를 포함하는, 폴리펩타이드를 척추동물에 전달하는 방법을 제공한다.
- <92> 본 발명은 또한, (a) CAEV 5' LTR과 CAEV *gag*-코딩 서열 사이의 비번역 영역, 및 (b) 상기 비번역 영역의 3' 말단에 연결된 CAEV *gag*-코딩 서열의 1 내지 X번째 뉴클레오티드(이때, X는 613 미만임)로 필수적으로 이루어진, CAEV 패키징 서열을 포함하는 벡터를 제공한다.
- <93> 본 발명자들은, 본원에서 제시된 바와 같이, 상기 CAEV-계 렌티바이러스 벡터 입자의 생산이 렌티바이러스 벡터 설계에 있어서 기존 CAEV-계 벡터 입자들에 비해 향상된 효율성 및 안전성을 나타내는 것을 발견하였다. 상기 향상된 효율성은 효율적인 캡시드화를 가능하게 하는 효율적인 패키징 서열을 제공하는, 5' LTR과 *gag* 시작 코돈 사이의 비번역 영역 및 *gag*-코딩 영역의 최적의 길이를 발견함으로써 달성되며, 이는 바이러스 역가를 증대시킨다. 바이러스 역가는 또한 패키징 플라스미드의 설계시 강력한 이종의 프로모터를 사용하여 향상된다. 상기 향상된 안전성은 *tat*-독립적인 전이 벡터 및 플라스미드-계 패키징 시스템의 제작을 통해 달성된다.

발명의 상세한 설명

- <117> 본 발명은, 특히, CAEV-계 렌티바이러스 벡터 시스템 및 상기 벡터를 사용하여 관심 폴리펩타이드를 분열 및 비분열 세포로 전달하는 방법에 관한 것이다.
- <118> CAEV 유전체
- <119> 야생형 CAEV 바이러스는, 이중-가닥 DNA 중간 물질을 통해 복제되고 핵단백질 코어(nucleoprotein core)를 포함하는 구형의 외피 비리온(virion)으로 패키징되는, 2량체 RNA 유전체를 갖는다(단일-가닥, 양성 극성). 상기 유전체는 구조 및 효소 단백질인 *Gag*, *Pol* 및 *Env*를 코딩하는 세 개의 유전자, 및 통합된 바이러스 유전체 각 말단에서의 긴 말단 반복(long terminal repeat; LTR)을 포함한다. 또한, 상기 유전체는 세 개의 조절 단백질인 *vif*, *tat* 및 *rev*를 코딩한다.
- <120> 상기 *gag* 유전자는 내부 구조 단백질을 코딩하고, *pol* 유전자는 바이러스 복제 효소를 코딩하며, *env* 유전자는 바이러스가 세포 표면에 부착하는데 매개하는 외피 당단백질(glycoprotein)을 코딩한다. *Vif* 단백질은 바이러스 감염성(infectivity)에 연관되어 있고, *Tat* 단백질은 5' LTR의 전사활성화(transactivation)에 연관되어 있다. *Rev* 단백질 및 그 표적 서열 RRE(*Rev* 반응 요소)는 바이러스 RNA의 안정성, 바이러스 RNA 스플라이싱의 조절, 및 대형 RNA(비-스플라이싱되거나 단일-스플라이싱된)을 핵으로부터 세포질로 수송하는데 연관되어 있다. 프로바이러스 LTR 서열은 U3(구조 단백질들의 하류에 위치하는 특이(unique) 서열 요소), R(각 유전체 말단의 짧은 반복서열), 및 U5(상기 R 서열 바로 다음의 특이 서열 요소) 영역들을 포함한다. 5'LTR의 U3 영역은 바이러스 프로모터 및 인핸서(enhancer)를 포함한다. 유전체의 3' 말단은 3'LTR에서의 폴리아데닐화 신호를 포함한다.
- <121> CAEV의 야생형 유전체는 또한 프로바이러스 통합을 위한 LTR 말단의 att(부착 위치) 등의 여러 시스-작용 요소; 5'LTR에서 통합된 프로바이러스의 전사 개시를 조절하는 프로모터 요소; 5'LTR의 하류에 위치하는 PBS(프라이머 결합 위치); 5'-스플라이싱 제공 위치; 패키징 서열(본원에서는 호환가능하게 패키징 위치 또는 패키징 신호를 의미한다.); 3'LTR 부근에 위치하는 ppt(폴리퓨린 트랙) 위치; 및 3'LTR에서의 폴리아데닐화 신호를 포함한다.
- <122> 본원에서 사용된 용어 "시스"는 핵산의 동일 염색체 또는 선형 부분 상에 존재하는 유전자들을 의미한다. 따라서, 용어 "시스-결합"은 핵산의 선형 서열 상에서 발견된 결합을 의미한다. 용어 "시스-작용"은 핵산의 동일 염색체 또는 선형 부분 상에 존재하는 유전자에 대한 조절 유전자의 제어 효과를 의미한다. 예를 들면, 하류 mRNA의 합성에 영향을 주는 프로모터는 시스-작용 조절 요소이다.
- <123> CAEV의 두 분리주에 대한 전장 유전체 서열이 공지되어 있으며, 상기 서열들은 생명공학 정보센터(National Center for Biotechnology Information; NCBI) 데이터베이스에 NC_001463(서열번호: 1) 및 AF322109(서열번호: 2)(Saltarelli et al., 1990, 및 Gjerset, B.J. et al., 비공개)로서 기탁되어 있다. 청구된 본 발명의 핵산은 CAEV의 특정 분리주에 한정되지는 않으나, 그 유전체 서열의 공지된 기능을 유지하는 서열일 수 있다. 예를

들면, 유전자 서열에서의 자연적인 변이들이 바이러스 복제 중에 일어남으로써 유사한 기능을 갖는 단백질을 코딩하는 유사 핵산 서열들이 발생된다는 것이 당 분야에 공지되어 있다.

<124> NC_001463(서열번호: 1) 및 AF322109(서열번호: 2) 유전체 서열들의 서열 정렬(sequence alignment)을 표 1에 나타내었다. 표 1에서 볼 수 있는 바와 같이, 상기 서열들간에 상당한 핵산 상동성이 있으나, 핵산 수준에서의 차이가 명백하다. 표 2(서열번호: 3 내지 6)에 기재된 CAEV *gag* 영역의 변이가능성은 특히 중요하다. NC_001463 5'LTR, *pol*, *rev* 및 *vif* 유전자들, 및 그에 상응하는 AF322109로부터의 유전자들의 서열 정렬은 표 3-6(서열번호: 7-14)에서 확인될 수 있다. 또한, CAEV 유전체의 많은 부분 서열들이 공지되고 기탁되어 있다. 예를 들면, 기탁 번호 AY081139, AY101347, AY101348, AY047362, AF402668, AF402667, AF402666, AF402665, AF402664, AJ305042, AJ305041, 및 AJ305040는 모두 CAEV의 브라질 분리주로부터의 *gag* 유전자 서열을 제공한다. 기탁 번호 AF015181, L78453, L78451, L78450, L78447 및 L78446은 또한 다양한 CAEV 분리주들로부터의 *gag* 유전자 서열을 포함한다. 기탁번호 X64828 및 M63106은 다양한 CAEV 분리주들로부터의 *rev* 유전자 서열을 포함한다. 기탁번호 AF015182, AJ305053, K03327, L78448, L78452 및 U35814는 다양한 CAEV 분리주들로부터의 *pol* 유전자를 포함한다. NC_001463의 *gag* 유전자(서열번호: 15, 17)와, AF015181(서열번호: 16, 17), AF402664(서열 번호: 20, 26), AF402665(서열 번호: 21, 27), AF402666(서열 번호: 22, 28), AF402667(서열 번호: 23, 29) 및 AF402668(서열 번호: 24, 30)로부터의 *gag* 유전자들간의 서열 정렬은 표 8에서 확인된다. NC_001463의 *gag* 유전자(서열 번호: 31, 35)와, AJ305040(서열 번호: 32, 36), AJ305041(서열 번호: 33, 37) 및 AJ305042(서열 번호: 34, 38)로부터의 *gag* 유전자들 간의 서열 정렬은 표 9에서 확인된다. NC_001463의 *gag* 유전자(서열 번호: 39, 41)와 AY047362(서열 번호: 40, 42)로부터의 *gag* 유전자 간의 서열 정렬은 표 10에서 확인된다. NC_001463 *gag* 유전자(서열 번호: 43, 45)와 AY081139(서열 번호: 44, 46)로부터의 *gag* 유전자 간의 서열 정렬은 표 11에서 확인된다. NC_001463(서열 번호: 47, 50)의 *gag* 유전자와, AY101347(서열 번호: 48, 51) 및 AY101348(서열 번호: 49, 52)로부터의 *gag* 유전자들 간의 서열 정렬은 표 12에서 확인된다. NC_001463 *gag* 유전자(서열 번호: 53, 59)와, L78446(서열 번호: 54, 60), L78447(서열 번호: 55, 61), L78450(서열 번호: 56, 62), L78451(서열 번호: 57, 63) 및 L78453(서열 번호: 58, 64)으로부터의 *gag* 유전자들 간의 서열 정렬은 표 13에서 확인된다.

<125> 상기 정렬들은 벡터 NTI(VectorNTI, Informax, USA)를 사용하여 하기 파라미터를 사용하여 수행되었다:

<126> 이중정렬(pairwise alignment)의 경우: 공백(gap) 개시 벌점: 15

<127> 공백 연장 벌점: 6.6

<128> 다중정렬(multiple alignment)의 경우: 공백 개시 벌점: 15

<129> 공백 연장 벌점: 6.6

<130> 공백 분리 벌점 범위 : 8

<131> 표 14에는 상기 제시된 *gag* 유전자 서열들의 서열 정렬에 대한 상동성 백분율 값들이 요약되어 있다. 표 15에는 NC_001463(서열 번호: 1) 및 AF322109(서열 번호: 2)의 전체 유전체 정렬, 및 *gag*, 5 LTR, *pol*, *rev* 및 *vif* 영역들의 정렬들의 상동성 백분율이 요약되어 있다. 다양한 CAEV 분리주들로부터의 많은 부분 서열들 뿐 아니라 두 CAEV 분리주들의 유전체 서열이 공지되어 있으며, 일치되는 서열들을 용이하게 식별할 수 있으므로, 다양한 CAEV 서열들을 사용하여 청구된 본 발명을 실시하기 위해 과도한 실험이 요구되지 않는다.

<132> 본 발명의 CAEV 벡터

<133> 본 발명의 벡터는 넓은 계통분류학적 범위의 숙주 세포에서 숙주 세포의 핵과 무관한 폴리뉴클레오티드 또는 유전자들을 복제하고 발현하기 위한 수단을 제공한다. 이러한 이중 핵산의 벡터-매개된 숙주세포로의 결합은 숙주 세포의 형질감염 또는 감염을 의미하며, 이때 감염은 바이러스 입자의 사용을 의미하고, 형질감염은 핵산 분자 그대로의 사용을 의미한다.

<134> 용어 "유전자"는 폴리펩티드 또는 전구체의 생산에 필요한 조절(control) 및 코딩 서열들을 포함하는 DNA 서열을 의미한다. 본원에서 상호호환적으로 사용된 용어 "폴리뉴클레오티드" 또는 "핵산 분자"는 두 개 이상과 같이 임의의 길이를 갖는 뉴클레오티드 폴리머를 의미하며, DNA 및 RNA 둘 다를 포함한다. 뉴클레오티드는 데옥시리보뉴클레오티드, 리보뉴클레오티드, 뉴클레오티드 유사물(analog)(변형된 인산기, 염기 또는 당 등), 또는 DNA 중합효소 또는 RNA 중합효소 등의 적절한 효소에 의해 폴리머로 결합될 수 있는 임의의 치환기일 수 있다. 폴리펩타이드는 전장 코딩 서열 또는 폴리펩타이드의 목적하는 활성이 유지되는 정도의 길이를 갖는 코딩 서열

의 임의의 부분에 의해 코딩될 수 있다.

- <135> 용어 "야생형"은 자연발생원(naturally occurring source)으로부터 분리되어 그 유전자 또는 유전자 산물의 특징을 갖는 유전자 또는 유전자 산물을 의미한다. 야생형 유전자는 개체군들 중에서 가장 빈번하게 관찰되어 임의로 그 유전자의 "일반적인" 또는 "야생의" 형태로서 고안된 것이다. 반면에, 용어 "변이된" 또는 "변이체"는 상기 야생형 유전자 또는 유전자 산물과 비교할 때, 서열 또는 기능적 특징에서의 변형(즉, 대체된 특징)을 나타내는 유전자 또는 유전자 산물을 의미한다. 자연-발생적 변이체는 분리될 수 있으며, 야생형 유전자 또는 유전자 산물과 비교할 때, 변화된 특징을 갖는다는 사실에 의해 식별될 수 있다.
- <136> 본원 명세서 및 청구항에서 사용되는 "a", "an", "the" 등의 단일형 형태들은 문맥상 명확하게 기재하고 있지 않는 한 복수의 인용물들을 포함한다. 따라서, 예를 들면, "폴리뉴클레오티드"의 의미에는 폴리뉴클레오티드들도 포함되며, "줄기 세포"의 의미에는 복수의 세포들도 포함된다.
- <137> 본원에 사용된 용어 "레트로바이러스"는 그들의 복제 주기 동안 역전사효소를 이용하는 RNA 바이러스들을 의미한다. 레트로바이러스 유전체 RNA는 역전사 효소에 의해 이중-가닥 DNA로 전환된다. 바이러스의 이러한 이중-가닥 DNA 형태는 감염된 세포의 염색체내로 통합될 수 있으며; 일단 통합되면, 이것을 "프로바이러스"라고 일컫는다. 프로바이러스는 RNA 중합효소 II에 대한 주형(template)으로 제공되며 신규 바이러스 입자들을 생산하는데 필요한 구조 단백질 및 효소들을 코딩하는 RNA 분자들의 발현을 지배한다.
- <138> 본원에 사용된 용어 "렌티바이러스"는 질병을 서서히 진행시키는 레트로바이러스의 한 군(또는 속)을 의미한다. 이러한 군에 포함되는 바이러스들로는 인간 면역결핍 바이러스(HIV); 양에서 뇌염(비스나) 또는 폐렴(매디)을 일으키는 비스나-매디, 염소 관절염 뇌염 바이러스(CAEV); 말 감염성 빈혈 바이러스(EIAV); 고양이 면역결핍 바이러스(FIV); 소 면역결핍 바이러스(BIV); 및 원숭이 면역결핍 바이러스(SIV)가 있다. 이러한 바이러스들에 의해 발병되는 질병들은 긴 잠복기와 장기화된 경로를 특징으로 한다. 일반적으로, 상기 바이러스들은 단핵세포 및 대식세포들을 잠재적으로 감염시키고, 이들로부터 다른 세포들로 퍼져나간다.
- <139> 본원에 사용된 용어 "벡터"는 전이 폴리뉴클레오티드(예: DNA) 조각들을 하나의 세포로부터 다른 세포로 전이하는 핵산 분자들을 의미한다. 용어 "전달물질(vehicle)"은 때때로 "벡터"와 상호호환적으로 사용된다. 전달물질 또는 벡터의 임의의 형태는 이러한 정의 내에 포함된다. 예를 들면, 벡터로는 이에 한정되는 것은 아니나, 바이러스 입자, 플라스미드, 트랜스포존(transposon) 등이 포함된다.
- <140> 본 발명의 벡터의 제작을 위한 기본 기술들은 통상의 기술자들에게 이미 공지되어 있으며, [Sambrook et al., Molecular Cloning: A Laboratory Manual 2nd Ed. (Cold Spring Harbor, N.Y., 1989)]와 같은 인용문헌에서 확인될 수 있다. 다양한 방법들이 DNA 절편을 결합하기 위해 활용가능하며, 이의 선택은 DNA 절편의 말단 성질에 따라 통상의 기술자에 의해 용이하게 이루어진다.
- <141> 본 발명의 적합한 폴리아데닐화 서열로는 이들로 한정되는 것은 아니나, 소 성장 호르몬(BGH) 폴리아데닐화 신호(Pfarr et al., 1986), SV40 초기 영역 폴리아데닐화 위치(Hall et al., 1983) 및 SV40 후기 영역 폴리아데닐화 위치(Carswell and Alwine, 1989), β -글로빈 폴리A, 및 허피스 심플렉스 바이러스 티미딘 카이네이즈 폴리A 등이 포함된다.
- <142> 본 발명의 프로모터는 포유류 또는 바이러스 기점에 대한 프로모터를 포함할 수 있으며, 세포에서 말단에 위치하는 서열(즉, 프로모터 서열의 5' 말단에 연결된 서열)의 전사를 지배하기에 충분할 것이다. 프로모터 영역으로는 또한 전사의 증대 또는 감소에 대한 조절 요소들이 포함될 수 있다. 본 발명에 적합한 프로모터로는 이에 한정되는 것은 아니나, 인간 또는 쥐 거대세포바이러스 극초기 프로모터(HCMV MIEP 또는 MCMV MIEP), 연장 인자 1 알파(ef-1 α), 및 로우스 육종 바이러스 긴 말단 반복 프로모터(pRSV) 등이 있다. 인트론 서열들은 또한 프로모터와 결합될 수 있다. 인트론 서열들로는 이에 한정되는 것은 아니나, ef-1 α 인트론 및 β -글로빈 인트론 등이 있다. 또한, 유도발현 시스템(inducible expression systems)이 사용될 수 있다. 유도 시스템의 예로는 이에 한정되는 것은 아니나, 엑시존(ecdysone)-유도 포유류 발현 시스템(Invitrogen, CA, USA) 및 Tet-On 및 Tet-Off 유전자 발현 시스템(Clontech, CA, USA) 등이 있다. 세포 또는 조직 특이적 프로모터들은 특정 세포 개체군내 유전자 서열의 표적 발현을 위해 이용될 수 있다.
- <143> 프로모터 또는 터미네이터(terminator) 서열들의 상류, 및 코딩 영역의 하류로부터의 인핸서(enhancer) 서열들은 발현을 촉진하기 위해 본 발명의 벡터에 선택적으로 포함될 수 있다. 본 발명의 벡터는 또한 세포가 벡터의 핵산에 의해 발현된 단백질을 효율적 및 효과적으로 처리하도록, 인트론 서열, 이동 서열(localization sequence) 또는 신호 서열과 같은 추가의 핵산 서열들을 포함할 수 있다. 인트론 서열의 예로는 β -글로빈 인

트론(Kim et al., 2002) 및 인간 EF-1 α 인트론(Kim et al., 2002) 등이 포함된다. 이같은 추가의 서열들은 벡터에 삽입되어, 전사를 목적으로 하는 경우에는 프로모터 서열에 작동가능하게 연결되거나, 번역 및 처리를 목적으로 하는 경우에는 개시 및 처리 서열에 추가적으로 연결된다. 한편, 삽입된 서열들은 벡터 내 어떤 곳에도 위치할 수 있다.

- <144> 용어 "작동가능하게 연결된"은, 유전자 서열의 전사가 작동가능하게 연결된 프로모터 서열에 의해 지배되고, 유전자 서열의 번역이 작동 가능하게 연결된 번역 조절 서열에 의해 지배되고, 또한 유전자 서열의 번역-후 처리가 작동가능하게 연결된 처리 서열에 의해 지배되는, 유전자 서열과 프로모터 또는 다른 조절 또는 처리 서열간의 연결을 의미한다.
- <145> 용어 "SIN 벡터"는 3'LTR에서의 절단된(truncated) U3 영역을 갖는 자가-불활성(self-inactivating) 벡터를 의미한다. 역전사 기간 동안, 절단된 U3는 5'LTR에서 복제되고, 그 결과 전사 용량(capacity)의 손실 및 내부 프로모터에 대한 저해 효과가 나타난다.
- <146> 전이 벡터의 패키징 서열은 (i) CAEV 5' LTR과 CAEV *gag*-코딩 서열 간의 비번역 영역, 및 (ii) 상기 비번역 영역의 3' 말단에 연결된 CAEV *gag*-코딩 서열의 1 내지 X번째 뉴클레오티드(이때, X는 613 미만임)로 필수적으로 이루어진다. 일 실시양태에서, X는 60, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575 및 600으로 이루어진 군으로부터 선택된다.
- <147> 본 발명의 다른 실시양태에서, X는 다음으로 이루어진 군으로부터 선택된다:
- <148> (a) X는 25 초과 600 미만,
- <149> (b) X는 25 초과 500 미만,
- <150> (c) X는 25 초과 400 미만,
- <151> (d) X는 25 초과 300 미만,
- <152> (e) X는 25 초과 200 미만,
- <153> (f) X는 50 초과 600 미만,
- <154> (g) X는 50 초과 500 미만,
- <155> (h) X는 50 초과 400 미만,
- <156> (i) X는 50 초과 300 미만,
- <157> (j) X는 50 초과 200 미만,
- <158> (k) X는 75 초과 600 미만,
- <159> (l) X는 75 초과 500 미만,
- <160> (m) X는 75 초과 400 미만,
- <161> (n) X는 75 초과 300 미만,
- <162> (o) X는 75 초과 200 미만,
- <163> (p) X는 100 초과 600 미만,
- <164> (q) X는 100 초과 500 미만,
- <165> (r) X는 100 초과 400 미만,
- <166> (s) X는 100 초과 300 미만,
- <167> (t) X는 100 초과 200 미만,
- <168> (u) X는 125 초과 600 미만,
- <169> (v) X는 125 초과 500 미만,
- <170> (w) X는 125 초과 400 미만,

- <171> (x) X는 125 초과 300 미만,
- <172> (y) X는 125 초과 200 미만,
- <173> (z) X는 150 초과 600 미만,
- <174> (aa) X는 150 초과 500 미만,
- <175> (bb) X는 150 초과 400 미만,
- <176> (cc) X는 150 초과 300 미만,
- <177> (dd) X는 150 초과 200 미만,
- <178> (ee) X는 200 초과 600 미만,
- <179> (ff) X는 200 초과 500 미만,
- <180> (gg) X는 200 초과 400 미만,
- <181> (hh) X는 200 초과 300 미만,
- <182> (ii) X는 200 초과 200 미만,
- <183> (jj) X는 250 초과 600 미만,
- <184> (kk) X는 250 초과 500 미만,
- <185> (ll) X는 250 초과 400 미만, 및
- <186> (mm) X는 250 초과 300 미만.
- <187> 다른 실시양태에서, X는 40 초과 613 미만이다. 또 다른 실시양태에서, X는 약 327이다. 전이 벡터의 일 실시양태에서, *gag* 번역을 개시하는 코돈은 변이되거나(예: ATG가 TAG, TTG, CTG 또는 ATT로 치환) 삭제된다. 용어 "코돈"은, 특정 아미노산의 성장 폴리펩타이드 채로의 결합에 대한 지령을 의미하는, DNA 또는 전령(messenger) RNA 분자에서의 세 개 뉴클레오티드로 이루어진 하나의 서열을 의미한다. 전이 벡터는 이종의 프로모터 및 하나 이상의 시스-작용 서열들을 추가로 포함한다.
- <188> 본원에 사용된 용어 "패키징 신호" 또는 "패키징 서열"은 바이러스 RNA가 바이러스 캡시드 또는 입자로 캡시드 화하는데 요구되는, CAEV 유전체의 5' LTR에 인접하여 위치하는 서열을 의미한다. 여러 레트로바이러스 벡터는 바이러스 유전체의 캡시드화에 요구되는 최소한의 패키징 신호(또한 싸이(psi[ψ]) 서열이라고도 함)를 사용한다. 따라서, 본원에 사용된 용어 "패키징 서열", "패키징 신호", "싸이", 및 기호 " ψ "은 바이러스 입자 형성 기간 동안에 CAEV RNA 가닥들의 캡시드화에 요구되는 비-코딩 서열을 의미한다.
- <189> 본 발명의 다른 실시양태에서, 전이 벡터는 또한 전사 카세트(transcription cassette)를 포함한다. 본원에 사용된 용어 "전사 카세트"는 이종의 프로모터에 작동가능하게 연결된, 유전학적 요소들의 특정 그룹, 일반적으로 이종의 프로모터에 작동가능하게 연결되고, 관심 폴리펩타이드를 발현하는 폴리뉴클레오티드를 포함하는 핵산의 절편(fragment) 또는 조각(segment)을 의미한다. 상기 카세트는 단일 단위로서 제거되거나 벡터 또는 플라스미드에 삽입될 수 있다.
- <190> 본 발명의 전이 벡터에 대한 실례가 도 3c에 나타나 있다. 도 3c는 플라스미드 pCAH/SINd1(서열 번호: 68)을 나타낸다. pCAH/SINd1(서열 번호: 68)은 HCMV MIEP 프로모터, CAEV 5'LTR에서의 R-U5 서열 영역, PBS 위치를 포함하는 잔여 비번역 서열, ATG→TAG 이중 점변이를 갖는 *gag* 유전자의 327 bp 절편, RRE 영역 및 U3-삭제된 CAEV 3'LTR 영역을 포함하는, 4,238 bp의 플라스미드이다. 상기 벡터는 또한 Col E1 복제 기점(bp 2535-3118) 및 카나마이신 내성 유전자 영역(bp 3370-4182)을 포함한다. 전이 벡터의 다른 실례들이 도 3a 내지 3h에 나타나 있다.
- <191> 본 발명은 상술된 전이 벡터 및 패키징 벡터 시스템을 포함하는 CAEV 벡터 시스템을 제공한다. 상기 패키징 벡터 시스템은 제1 및 제2 폴리뉴클레오티드 벡터 서열을 포함한다. 제1 폴리뉴클레오티드 서열은 CAEV *gag-pol* 및 RRE-코딩 서열을 포함하며, 제2 폴리뉴클레오티드는 바이러스 외피 코딩 서열을 포함한다. 일 실시양태에서, 제2 폴리뉴클레오티드는 비-CAEV 외피를 코딩한다.
- <192> 본원에서 사용된 문구 "구조 유전자"는 바이러스 유전체의 캡시드화(예: 패키징)에 요구되는 단백질을 코딩하는

폴리뉴클레오티드 서열을 의미하며, *gag*, *pol* 및 *env* 등이 이에 속한다.

- <193> 본 발명의 제1 패키징 벡터의 실례가 도 2a에 나타나 있다. 도 2a는 플라스미드 pMGP/RRE(서열 번호: 77)를 나타낸다. 상기 플라스미드는 9,446 염기쌍(bp)을 가지며, MCMV MIEP 영역, CAEV *gag-pol* 코딩 영역, RRE 영역, 및 소 성장 호르몬(BGH) 폴리아데닐화 신호를 포함한다. 상기 벡터는 또한 네오마이신 내성 유전자 코딩 영역, SV40 복제 기점, Col E1 복제 기점, 및 암피실린 내성 유전자 영역을 포함한다.
- <194> 본 발명의 바이러스 벡터가 감염시킬 수 있는 세포들의 숙주 범위는 다른 근친관계의 바이러스로부터의 외피 유전자를 이용함으로써 대체될 수 있다. 즉, 특정 바이러스의 외피 단백질이 갖는 다른 바이러스의 캡시드화에 관여하는 능력의 장점을 취하여 본원 발명의 CAEV 벡터의 숙주 범위를 넓히는 것이 가능하다. 레트로바이러스-유래 *env* 유전자의 예로는 이에 한정되는 것은 아니나 다음과 같다: 소수포성 구내염 바이러스의 G-단백질(VSV-G), 긴팔 원숭이 백혈병 바이러스(GaLV), 라우스 육종 바이러스(RSV), 몰로니 쥐 백혈병 바이러스(MoMuLV), 마우스 유방 종양 바이러스(MMTV), 및 인간 면역결핍성 바이러스(HIV). 이들 바이러스 외피 단백질은 모두 다른 바이러스의 유전체 및 매트릭스 구성원들과 함께 의사형 비리온을 효과적으로 형성한다. 본원에서 사용된 용어 "의사형(pseudotype)"은 어떤 바이러스의 핵산과 다른 바이러스의 외피 단백질을 포함하는 바이러스 입자를 의미한다. 일반적으로, VSV-G나 GaLV 의사형 벡터는 매우 넓은 숙주 범위를 가지며, 여전히 높은 수준의 감염성을 나타내는 반면, 초원심분리에 의해 높은 농도의 역가로 펠렛화 될 수 있다(Burns et al., 1993).
- <195> 본 발명의 제2 패키징 벡터의 다른 실례가 도 6a 및 6b에 나타나 있다. 도 6a는 플라스미드 pHGVSV-G(서열 번호: 74)를 나타낸다. pHGVSV-G(서열 번호: 74)는 HCMV MIEP, β -글로빈 인트론 영역, VSV-G 코딩 영역, BGH 폴리아데닐화 신호, Col E1 복제 기점, 네오마이신 내성 유전자 코딩 영역, 암피실린 내성 유전자 코딩 영역, 및 F1 복제기점을 포함하는 7,623 bp의 플라스미드이다. 도 6b는 플라스미드 pMYKEF1/*env*(서열 번호: 72)를 나타낸다. 상기 플라스미드는 MCMV MIEP, 인간 EF1- α 인트론 영역, GaLV *env* 코딩 영역, BGH 폴리아데닐화 신호, Col E1 복제 기점, 네오마이신 내성 유전자 코딩 영역, 및 암피실린 내성 유전자 코딩 영역을 포함하는 7,579 bp를 포함한다.
- <196> 본 발명의 다른 실시양태에서, 패키징 벡터는 Rev를 코딩하는 제3 폴리뉴클레오티드를 포함한다. 감염된 세포에서, Rev는 바이러스 전사물에서의 Rev-반응 요소(RRE)와 결합하여 복제 후기에 바이러스 구조 단백질의 특징적인 단일-스플라이싱되거나 비스플라이싱된 전사물들의 전사를 일으킨다. 따라서, Rev는 바이러스 유전자 발현의 일시적인 조절을 매개한다. 포유류 세포의 스플라이싱 기작은 mRNA가 핵내 합성 위치로부터 세포질로 이동하는 것과 짝을 이루기 때문에, Rev는 또한 RRE를 포함하는 바이러스 전사물의 수송에도 영향을 미친다.
- <197> 본 발명의 제3 패키징 벡터의 실례가 도 5에 제시되어 있다. 도 5는 플라스미드 pHYK/*rev*(서열 번호: 75)를 나타낸다. pHYK/*rev*(서열 번호: 75)는 HCMV MIEP, *rev* 유전자 코딩 영역, BGH 폴리아데닐화 신호, Col E1 복제 기점, 네오마이신 내성 유전자 코딩 영역, 및 암피실린 내성 유전자 코딩 영역을 포함하는 5,419 bp의 플라스미드이다.
- <198> 본 발명의 또 다른 실시양태에서, 패키징 벡터는 Vif를 코딩하는 제4 폴리뉴클레오티드를 포함한다. Vif의 결합은 선택된 패키징 세포주에 따라 비리온의 감염 및 패키징에 필수적일 수 있다.
- <199> 본 발명의 제4 패키징 벡터의 실례가 도 4에 제시되어 있다. pHYK/*vif*(서열 번호: 76)는 HCMV MIEP, *vif* 유전자 코딩 영역, BGH 폴리아데닐화 신호, Col E1 복제 기점, 네오마이신 내성 유전자 코딩 영역, 및 암피실린 내성 유전자 코딩 영역을 포함하는 5,729 bp의 플라스미드이다.
- <200> 레트로바이러스 벡터 DNA가 세포 내로 형질감염될 때, 이는 염색체 DNA로 통합되거나 통합되지 않을 수도 있고, 이후 전사됨으로써 ψ 서열을 포함하는 전장 레트로바이러스 벡터 RNA를 생산한다. 이러한 조건 하에, 오직 벡터 RNA만 바이러스 캡시드 구조물로 패키징된다. 이러한 완성된, 그러나 복제-결합의, 바이러스 입자들은 상대적으로 높은 효율성을 가지고 레트로바이러스 벡터를 표적세포로 전달하는데 사용될 수 있다.
- <201> 본원에서 사용된 용어 "복제-결합"은 감염성 비리온이 생산되지 않도록, 완전하고 효과적으로 복제할 수 없는 바이러스를 의미한다(예: 복제-결합 렌티바이러스 프로제니(progeny)). 용어 "복제-적격"은 바이러스 복제에 의해 감염성 비리온이 생산될 수 있도록 하는 복제 능력이 있는, 야생형 바이러스 또는 변이 바이러스를 의미한다(예: 복제-적격 렌티바이러스 프로제니).
- <202> 패키징은 또한 비-유도성(non-inducible)일 뿐 아니라 유도성(inducible)일 수 있다. 유도 패키징 세포 및 패키징 세포주에서, CAEV 입자들은 최소 하나 이상의 유도물질(inducer)에 반응하여 생산된다. 유도 세포주를 사용한 바람직한 실시양태에서, 유도물질은 Tat이다. 비-유도 패키징 세포주 및 패키징 세포에서는, 렌티바이러스

스 입자 생산을 일으키기 위해 유도성 물질을 필요로 하지 않는다.

- <203> CAEV 백터 서열
- <204> 본 발명의 기능적으로 동등한 서열들은 또한 각각의 고유 서열(native sequence)과 실질적으로 동일한 기능을 유지하는 CAEV 유전체의 다양한 절편들이 포함된다. 이러한 절편들은 특정 관심 유전자의 요소들의 최소 약 10, 15개의 연속적인 뉴클레오티드, 최소 약 20개의 연속적인 뉴클레오티드, 최소 약 24, 50, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 340, 360, 380, 또는 그 이상 전체길이의 연속적인 뉴클레오티드를 포함할 수 있다. 이러한 절편들은 고유의 바이러스 유전체를 절단하는 제한 효소의 사용에 의해; 바이러스 유전체의 고유의 뉴클레오티드 서열로부터 뉴클레오티드서열을 합성함으로써; 또는 PCR 기술의 사용을 통하여 얻어질 수 있다(특히, (Mullis 및 Faloona, 1987) 및 (Erllich, 1989) 참조). 또한, 위치-유도 돌연변이(site-directed mutagenesis)의 결과에서처럼, 다양한 백터 구성원들의 변이체들이 본 발명의 방법에 포함된다. 좀 더 자세히 설명하면, 방법들은 기능적 동등성을 결정하는 당분야에 통상적인 것이다.
- <205> "변이체"로는 실질적으로 유사한 서열들이 포함된다. 따라서, 뉴클레오티드 서열 또는 아미노산 서열의 경우, 변이체로는 바이러스 백터 시스템의 다양한 구성원들과 기능적으로 동등한 서열들이 포함된다. 변이 뉴클레오티드 서열에는 또한 예를 들면, 고유 서열의 기능을 여전히 유지하는 위치 유도 돌연변이에 의해 발생된 합성적으로 유도된 뉴클레오티드 서열들이 포함된다. 일반적으로, 본 발명의 뉴클레오티드 서열 변이체 또는 아미노산 서열 변이체들은 그 각각의 고유 뉴클레오티드 서열에 대해 최소 70%, 일반적으로 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 상동성을 가질 것이다.
- <206> 본 발명의 변이체로는 본원에 개시된 백터(서열 번호:67-79)들의 서열과 최소 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 상동성을 갖는 서열들을 포함하거나, 이들로 필수적으로 이루어지거나, 또는 이들로 이루어진, 폴리뉴클레오티드(예: 백터)가 포함된다.
- <207> 통상의 기술자는 개시된 핵산 구조물의 많은 보존적 변이(conservative variation)들이 기능적으로 동일한 구조물을 생산한다는 것을 인식할 것이다. 특정 핵산 서열의 보존적 변이는 동일하거나 본질적으로 동일한 아미노산 서열을 코딩하는 핵산을 의미하거나, 또는 핵산이 아미노산 서열을 코딩하지 않을 경우 본질적으로 동일한 서열을 의미한다. 유전자 코드의 축퇴성으로 인해, 많은 기능적으로 동일한 핵산들은 임의이 주어진 폴리펩타이드를 코딩한다. 예를 들면, 유전자 코드의 축퇴성으로 인해, "침묵 치환(silent substitutions)"(즉, 코딩된 폴리펩타이드의 변이를 야기하지 않는 핵산 서열의 치환)은 아미노산을 코딩하는 모든 핵산 서열의 특징을 내재한다. 유사하게, 패키징 또는 패키징 가능한 구조물의 아미노산 서열의 하나 또는 소수의 아미노산에서 "보존적 아미노산 치환"은 매우 유사한 특징을 갖는 다른 아미노산으로 치환되며, 또한 개시된 구조물과 매우 유사하여 용이하게 확인된다. 예를 들면, 코돈 CGU, CGC, CGA, CGG, AGA 및 AGG는 모두 아미노산 아르기닌을 코딩한다. 따라서, 아르기닌이 코돈에 의해 특정되는 모든 위치에서, 상기 코돈은 코딩된 폴리펩타이드를 변화시키지 않고 제시된 임의의 대응 코돈으로 대체될 수 있다. 이러한 핵산 변이가 "보존적으로 변형된 변이"의 한 종류인, "침묵 변이"이다. 폴리펩타이드를 코딩하는 본원의 모든 핵산 서열은 또한 모든 가능한 침묵 변이를 제시한다. 통상의 기술자는 핵산에서의 각 코돈(통상적으로 메티오닌에 대한 유일한 코돈인 AUG는 제외)이 기본 기술에 의해 기능적으로 동일한 분자를 생산하도록 변형될 수 있다는 것을 인식할 것이다. 따라서, 폴리펩타이드를 코딩하는 핵산의 각 "침묵 변이"는 임의의 제시된 서열에 내재되어 있다. 또한, 통상의 기술자는 코딩된 서열에서 단일 아미노산 또는 일부 아미노산(일반적으로 5% 미만, 더 일반적으로 1% 미만)을 대체하거나, 첨가하거나 또는 삭제하는 개개의 치환, 삭제 또는 첨가가 아미노산을 화학적으로 유사한 아미노산으로 치환시키는 "보존적 변형 변이(conservatively modified variations)"인 것을 인식할 것이다. 기능적으로 유사한 아미노산들을 제공하는 보존적 치환표가 이미 당 분야에 공지되어 있다. 다음의 각 6 군은 서로 보존적 치환인 아미노산들을 포함한다:
 - <208> 1) 알라닌(A), 세린(S), 트레오닌(T);
 - <209> 2) 아스파르트산(D), 글루탐산(E);
 - <210> 3) 아스파라긴(N), 글루타민(Q);
 - <211> 4) 아르기닌(R), 리신(K);
 - <212> 5) 이소루신(I), 루신(L), 메티오닌(M), 발린(V); 및
 - <213> 6) 페닐알라닌(F), 타이로신(Y), 트립토판(W).

- <214> 또한, 문헌 (Creighton(1984) Proteins W. H. Freeman and Company)을 참조하라. 끝으로, 비-기능적 서열과 같이, 핵산 분자의 활성을 변화시키지 않는 서열의 첨가는 기본 핵산의 보존적 변형이다. 각 개시된 서열의 이러한 보존적 치환 변이는 본 발명의 특징이다.
- <215> 본 발명의 벡터 시스템에서 사용되는 다양한 전장 또는 성숙(mature) 폴리펩타이드에 대한 아미노산 서열과 관련하여, 변이체로는 고유의 폴리펩타이드에서 N-말단 및/또는 C-말단에서의 하나 이상의 아미노산의 삭제(절단이라고도 함) 또는 첨가; 고유의 폴리 펩타이드에서 하나 이상의 위치에서의 하나 이상의 아미노산의 삭제 또는 첨가; 또는 고유의 폴리펩타이드에서 하나 이상의 위치에서의 하나 이상의 아미노산의 치환에 의해, 고유의 폴리펩타이드로부터 유래된 폴리펩타이드들이 포함된다. 이러한 변이체들은, 예를 들면, 유전자 다형성 (polymorphism) 또는 인간의 조각으로부터 야기될 수 있다. 이러한 조각 방법은 일반적으로 당분야에 공지되어 있다.
- <216> 통상의 기술자는 주어진 핵산 구조물 내에서 치환을 일으키는 여러 방법들을 인식할 것이다. 이러한 공지된 방법들에는 위치-유도 돌연변이, 축퇴성 올리고뉴클레오티드를 이용하는 PCR 증폭, 핵산을 포함하는 세포의 변이 시약 또는 방사선으로의 노출, (예를 들어, 대형 핵산을 만드는 연결(ligation) 및/또는 클로닝에 관련된) 목적하는 올리고뉴클레오티드의 화학적 합성 및 기타 공지된 기술들이 포함된다(Gillam 및 Smith, 1979), (Roberts, Cheetham, 및 Rees, 1987), 및 Sambrook, Innis, Ausbel, Berger, Needham VanDevanter 및 Mullis(모두 상기 문헌) 참조).
- <217> 고유의 핵산 또는 고유의 폴리펩타이드의 변이체는 고유의 서열 또는 고유의 폴리펩타이드와 실질적인 상동성을 갖는다. 변이체는 1 내지 10개의 아미노산 잔기와 같이 적은, 6-10개와 같은, 5개와 같이 적은, 4, 3, 2 또는 1개의 아미노산 잔기와 같이 적은 수에 의해 달라질 수 있다. 핵산 서열의 변이체는 1 내지 30개의 핵산과 같이 적은, 6 내지 20개와 같은, 5개와 같이 적은, 4, 3, 2 또는 1개의 핵산 잔기와 같이 적은 수에 의해 달라질 수 있다.
- <218> "서열 상동성"은 변이체의 뉴클레오티드 서열 또는 아미노산 서열의 특성의 연속적인 절편을 인용 서열의 뉴클레오티드 서열 또는 아미노산 서열에 대해 정렬하고 비교할 때, 변이체 서열과 인용 서열내에 동일한 뉴클레오티드 또는 아미노산 잔기가 확인되는 것을 의미한다. 서열들간의 서열 정렬(sequence alignment) 및 상동성 결정 방법은 당분야에 공지되어 있다. 두 뉴클레오티드 서열들의 최적의 정렬과 관련하여, 변이 뉴클레오티드 서열의 연속적인 절편은 인용 뉴클레오티드 서열에 대해 추가의 뉴클레오티드 또는 삭제된 뉴클레오티드를 가질 수 있다. 마찬가지로, 두 아미노산 서열의 최적의 정렬을 위해서, 변이 아미노산 서열의 연속적인 절편은 인용 아미노산 서열에 대해 추가의 아미노산 잔기 또는 삭제된 아미노산 잔기를 가질 수 있다. 인용 뉴클레오티드 서열 또는 인용 아미노산 서열과의 비교를 위해 사용되는 연속적인 절편은 최소 20개의 연속적인 뉴클레오티드 또는 아미노산 잔기를 포함할 것이며, 30, 40, 50, 100, 또는 그 이상의 뉴클레오티드 또는 아미노산 잔기를 포함할 수 있다. 변이체의 뉴클레오티드 서열 또는 아미노산 서열에서 공백을 포함하는 것과 관련하여, 증가된 서열 상동성에 대한 보정은 공백 벌점을 부과함으로써 이루어질 수 있다.
- <219> 두 서열들 간의 상동성 백분율의 결정은 수학적 알고리즘을 사용하여 수행될 수 있다. 예를 들면, 아미노산 서열의 상동성 백분율은 12의 공백 개시 벌점 및 2의 공백 연장 벌점을 갖는 아핀 6 공백 검색(affine 6 gap search)인 블로섬 매트릭스 62(BLOSUM matrix 62)를 사용하는 스미스-워터만 상동성 검색 알고리즘(Smith-Waterman homology search algorithm)을 사용하여 결정될 수 있다. 한편, 뉴클레오티드 서열의 상동성 백분율은 25의 공백 개시 벌점 및 5의 공백 연장 벌점을 사용하는 스미스-워터만 상동성 검색 알고리즘을 사용하여 결정될 수 있다. 이러한 서열 상동성 결정법은, 예를 들면, 타임로직 버전 G(TimeLogic Version G)로부터의 디사이퍼 하드웨어 가속장치(DeCypher Hardware Accelerator)을 사용하여 수행될 수 있다. 상기 스미스-워터만 상동성 검색 알고리즘은 스미스 및 워터만에 의해 교시되고 있으며, 본원에서 인용되었다. 한편, 디폴트 파라미터(default parameters)를 사용하는 정렬 프로그램 GCG 갭(GCG Gap, Wisconsin Genetic Computing Group, Suite Version 10.1)이 사용될 수 있다. GCG 갭 프로그램은 니들만 및 운츠 알고리즘(Needleman and Wunch algorithm)을 적용하며, 3의 공백 개시 벌점 및 1의 공백 연장 벌점을 가지고 뉴클레오티드 서열 정렬을 위해 사용될 수 있다. 이외 바람직하나 이에 한정되는 것은 아닌 두 서열의 비교에 이용되는 수학적 알고리즘의 예는 카를린 및 알트슐(Karlin and Altschul, 1993)에 의해 변형된 카를린 및 알트슐의 알고리즘(Karlin and Altschul, 1990)이 있다. 이러한 알고리즘은 알트슐 등(Altschul et al., 1990)의 N블라스트(NBLAST) 및 X블라스트(XBLAST) 프로그램에 반영되었다. 블라스트(BLAST) 뉴클레오티드 검색법은 N블라스트 프로그램, 스코어(score)=100, 워드길이(wordlength)=12의 조건으로 수행되어 충분한 서열 상동성을 갖는 뉴클레오티드 서열들을 얻을 수 있다. 블라스트 단백질 검색법은 X블라스트 프로그램, 스코어=50, 워드길이=3의 조건으로 수행되어 충

분한 서열 상동성을 갖는 아미노산들을 얻을 수 있다. 비교의 목적으로 공백이 있는 정렬(gapped alignments)을 얻기 위해서는, 알트슐 등(Altschul et al., 1997)에 제시된 바에 따라 공백 블라스트(Gapped BLAST)가 이용될 수 있다. 한편, 분자들간의 원연 관계(distant relationships)를 검색하는 반복 검색(iterated search)을 수행하기 위해 PSI-블라스트(PSI-Blast)가 사용될 수 있다(상기 알트슐 등(Altschul et al., 1997) 참조). 블라스트, 공백 플라스트 및 PSI-블라스트 프로그램을 이용하는 경우, 각 프로그램(예: X블라스트 및 N블라스트)의 디폴트 파라미터가 사용될 수 있다(<http://www.ncbi.nlm.nih.gov> 참조). 서열 비교에 이용되는 수학적 알고리즘의 다른 비한정적인 예로는 마이어스 및 밀러(Myers and Miller (1988) CABIOS 4:11-17)의 알고리즘이 있다. 이러한 알고리즘은 GCG 서열 정렬 소프트웨어 패키지(GCG sequence alignment software package)의 일부인 얼라인 프로그램(ALIGN program, version 2.0)에 반영되어 있다. 아미노산 서열 비교를 위해 얼라인 프로그램을 이용하는 경우, PAM120 분자량 잔기 표(PAM120 weight residue table), 12의 공백 길이 별점, 및 4의 공백 별점에 사용될 수 있다. 아미노산 서열의 상동성 백분율은 또한 벡터NTI(VectorNTI, Informax, USA)를 사용하여 결정될 수 있다.

<220> 통상의 기술자는 제공된 서열 및 일반적으로 CAEV에 관련하여 당분야에 공지된 것을 근거로 본 발명의 목적하는 핵산을 선별할 수 있다. 렌티바이러스의 생명주기, 유전체 구조, 발생 조절 및 관련 분자 생물학은 지난 10년간의 집중적인 연구의 초점이 되어왔다. 많은 렌티바이러스 유전체에서 여러 변이들의 특정 효과가 공지되어 있다. 또한, 일부 CAEV 균주들의 핵산 서열 다양성이 공지되어 있다. 아울러, 통상의 단백질 및 핵산의 특성과 관련된 통상의 지식으로 인해 통상의 기술자는 본원의 서열 목록에 개시된 핵산 및 폴리펩타이드와 유사하거나 동등한 활성을 갖는 적절한 서열을 선택할 수 있다.

<221> 끝으로, 대부분의 핵산에 대한 변형은 목적하는 특성에 대한 적합한 분석에서의 일련의 스크리닝 기술에 의해 확인된다. 예를 들면, 코딩된 폴리펩타이드의 면역학적 특성에서의 변형은 적절한 면역학적 분석에 의해 검색될 수 있다. 상보적인 핵산에 대한 핵산 혼성화(hybridization), 코딩된 단백질의 산화환원 반응 또는 열안정성, 소수성, 단백질분해에 대한 감수성, 또는 응집 경향 등과 같은 다른 성질에 대한 변형도 기본 기술들에 따라 모두 분석된다.

<222> 관심 폴리뉴클레오티드

<223> 통상의 기술자가 인식하는 바와 같이, 삽입된 관심 폴리펩타이드의 뉴클레오티드 서열은 임의의 뉴클레오티드 서열일 수 있다. 예를 들면, 폴리뉴클레오티드 서열은 리포터(reporter) 유전자 또는 선별 마커(selectable marker) 유전자 서열일 수 있다. 본원에서 사용된 리포터 유전자 서열은 발현시 그 존재 또는 활성이 감지될 수 있는 단백질을 생산하게 하는 임의의 유전자 서열이다. 적합한 리포터 유전자의 대표적인 예로는 갈락토카이네이스(galactokinase), β-갈락토시데이즈(β-galactosidase), 클로람페니콜 아세틸트랜스퍼레이스(chloramphenicol acetyltransferase), β-락타메이스(β-lactamase), 녹색 형광 단백질(green fluorescent protein), 강화 녹색 형광 단백질(enhanced green fluorescent protein) 등이 있다. 한편, 리포터 유전자는 그 발현이 세포 생리에 영향을 주는 유전자 산물을 생산해내는 임의의 유전자 서열일 수 있다. 본 발명의 폴리뉴클레오티드 서열은 이미 하나 이상의 프로모터, 개시 서열, 또는 처리 서열을 갖는 하나 이상의 유전자 서열들을 포함할 수 있다.

<224> 선별 마커 유전자 서열은 그 존재로 인해 이를 포함하는 세포를 선택적으로 증식시키는 단백질을 발현할 수 있는 임의의 유전자 서열이다. 선별 마커 유전자의 예로는 항생제(예: 퓨로마이신, 하이그로마이신, 네오마이신, 테오신 등)에 대한 숙주 내성을 부여하거나, 아미노산 유사체에 대한 숙주 내성을 부여하거나, 기타 가혹(impermissible) 배양 조건 하에 추가의 탄소원 상에서 박테리아의 성장을 가능하게 할 수 있는 유전자 서열들이 포함된다.

<225> 리포터 또는 선별 마커 유전자 서열은 일반 세포에서 박테리아의 인식 또는 선별을 가능하게 하기에 충분하다. 본 발명의 일 실시양태에서, 리포터 유전자 서열은 일반적으로 포유류 세포에 없으며, 따라서 그 존재로 이러한 세포들에서의 박테리아의 존재를 명확하게 입증할 수 있는 효소 또는 기타 단백질을 코딩할 수 있다.

<226> 본 발명의 전이 박테리아는 추가적으로 이종의 핵산 또는 폴리뉴클레오티드를 바이러스 입자로 결합하도록 하여, 이로써 내부에 이종 핵산을 포함하는 감염된 숙주세포의 수를 증대시키기 위한 수단을 제공할 수 있다. 이종의 폴리뉴클레오티드의 결합은 바이러스 입자내 이종 핵산의 복제, 및 이후의 이종 단백질의 생산을 촉진시킨다. 본원에서 이종의 단백질은 단백질의 전부 또는 부분이 숙주 세포에 의해 발현되지 않는 단백질 또는 그의 절편으로 정의된다. 핵산 또는 유전자 서열은 유전자를 세포내로 전달하기 위해 사용된 바이러스 박테리아의 야생형 내에 본래부터 존재하는 것이 아니라면, 이종이라고 일컫는다. 본원에서 사용된 용어 이종의 핵산 서열 또는 폴

리뉴클레오티드 서열은 핵산 분자(바람직하게는 DNA)를 의미한다. 폴리뉴클레오티드 서열 또는 이중의 폴리뉴클레오티드 서열은 또한 적절한 생물학적 활성 단백질 또는 폴리펩타이드, 면역성 또는 항원성 단백질 또는 폴리펩타이드, 또는 치료 활성 단백질 또는 폴리펩타이드와 같은 목적하는 산물의 코딩 서열을 포함할 수 있다. 폴리펩타이드는 숙주 세포내 내인성 단백질의 결핍 또는 비존재(nonexistent) 발현을 보충할 수 있다. 이러한 유전자 서열들은 DNA, cDNA, 합성 DNA, RNA 또는 이들의 조합을 포함하는 다양한 출처로부터 유래될 수 있다. 이러한 유전자 서열들은 또한 자연적으로 발생하는 인트론을 포함하거나 포함하지 않을 수 있는 유전체 DNA를 포함할 수 있다. 또한, 이러한 유전체 DNA는 프로모터 서열 또는 폴리아데닐화 서열과 관련하여 수득될 수 있다. 본 발명의 유전자 서열은 바람직하게 cDNA이다. 유전체 또는 cDNA는 여러 방법에 의해 얻어질 수 있다. 유전체 DNA는 당분야에 공지된 방법에 의해 적절한 세포로부터 추출되고 분리될 수 있다. 한편, mRNA는 세포로부터 분리되어 역전사 또는 다른 방법을 통해 cDNA를 제조하기 위해 사용될 수 있다. 한편, 폴리뉴클레오티드 서열은 안티센스(antisense) RNA 서열과 같이 RNA 서열에 상보적인 서열을 포함할 수 있으며, 이러한 안티센스 서열은 개체내에 투여되어 개체 세포 내 상보적인 폴리뉴클레오티드의 발현을 저해할 수 있다.

- <227> 이중 유전자의 발현은 항체반응을 달성하는 면역성 또는 항원성 단백질을 또는 폴리펩타이드를 제공할 수 있다. 이렇게 발생된 항체들은 동물의 혈액, 혈청 또는 복수(ascites)와 같은 체액으로부터 회수될 수 있다.
- <228> 이중의 유전자는 또한 전사될 수 있는 임의의 관심 핵산일 수 있다. 일반적으로 외래 유전자는 폴리펩타이드를 코딩한다. 바람직하게 폴리펩타이드는 특정 치료적 장점을 가진다. 폴리펩타이드는 숙주 세포 내 내인성 단백질의 결핍 또는 비존재 발현을 보충할 수 있다. 폴리펩타이드는 숙주세포에 키메라 신호전달 수용체(chimeric signaling receptor)(미국 특허 제5,359,046호 참조)와 같은 새로운 성질들을 부여할 수 있다. 통상의 기술자는 본원에서 교시되고 당분야에 공지된 외래 유전자 시행 기술들의 적합성을 결정할 수 있다. 예를 들면, 상기 기술자는 외래 유전자가 캡시드화에 적절한 크기를 갖는지와 외래 유전자 산물이 적절히 발현되는지를 인식할 것이다.
- <229> 본 발명에서 사용될 수 있는 특정의 이중 단백질은 본 발명에 필수적인 것은 아니다.
- <230> 본 발명에서 사용될 수 있는 이러한 이중 단백질의 구체적인 예로는 디스트로핀(dystrophin)(Hoffman, Brown, and Kunkel, 1987), 응고 인자 VIII (Wion et al., 1985), 낭성섬유증 막전도 조절 단백질(cystic fibrosis transmembrane regulator protein; CFTR)(Anderson et al., 1991; Crawford, 1991), 오르니틴 트랜스카바밀레이즈(Ornithine Transcarbamylase; OTC)(Murakami et al., 1988), 및 α 1-안티트립신(α 1-antitrypsin)(Fagerhol and Cox, 1981) 등이 있다.
- <231> 여러 이중 단백질을 코딩하는 유전자들이 당분야에 공지되어 있으며, 유전체 또는 cDNA 라이브러리(libraries)로부터 클로닝될 수 있다[Sambrook et al, supra]. 이러한 유전자들의 예로는 디스트로핀 유전자(Lee et al., 1991), 응고 인자 VIII 유전자(Toole et al., 1984), CFTR 유전자(Rommens et al., 1989; Riordan, 1989), OTC 유전자(Horwich et al., 1984), 및 α 1-안티트립신 유전자(Lemarchand et al., 1992) 등이 있다.
- <232> 또한, 동맥경화증과 같은 혈관 증식성 질환의 치료를 위한 Rb(Chang et al., 1995), 및 암(Wills et al., 1994; Clayman, 1995) 및 HIV 질환(Bridges and Sarver, 1995)의 치료를 위한 p53과 같은 이중 단백질을 코딩하는 유전자들이 본 발명에 사용될 수 있다.
- <233> 벡터는 항상 기능적인 이중의 유전자 산물을 코딩할 필요가 없으며, 즉, 진행성 효소의 저해 물질로서 작용하는 부분 유전자 산물을 또한 코딩할 수 있다(Warne, Vician, and Downward, 1993; Wang, 1991).
- <234> 본 발명의 방법에 의한 분자 도입에 의해 세포내 유전자 조절 분자의 발현을 조정하는 것이 바람직하다. 용어 "조정하다(modulate)"는 유전자가 과발현되는 경우 그 발현을 억제하거나, 저발현되는 경우 그 발현을 증대시키는 것을 의미한다. 세포 증식성 질환이 유전자의 발현에 관련되는 경우, 번역 수준에서의 유전자 발현을 저해하는 핵산 서열이 사용될 수 있다. 예를 들면, 안티센스 핵산 또는 삼중 항원(triplex agent)으로 mRNA를 차폐시키거나 리보자임으로 이를 절단함으로써 특정 mRNA의 전사 또는 번역을 차단하기 위해, 안티센스 핵산, 리보자임 또는 삼중 항원 등이 이용될 수 있다.
- <235> 안티센스 핵산은 특정 mRNA 분자의 최소 일부분에 대해 상보적인 DNA 또는 RNA 분자들이다. 세포 내에서, 안티센스 핵산은 대응 mRNA와 혼성화 결합하여 이중-가닥 분자를 형성한다. 세포는 이중-가닥 mRNA를 번역하지 못하므로, 안티센스 핵산은 mRNA의 번역을 저해한다. 약 15개 또는 그 이상의 뉴클레오티드를 갖는 안티센스 올리고머는 용이하게 합성되면서도 표적세포내로 도입될 때 큰 분자들에 비해 문제를 덜 발생시키므로 바람직하다. 유전자의 생체의 번역을 저해하는 안티센스 방법의 용도는 당 분야에 공지되어 있다((Marcus-

Sekura, 1988).

- <236> 안티센스 핵산은 알츠하이머 질환(Alzheimer's disease)에서 축적되는 아밀로이드 전구체 단백질과 같이, 변이 단백질 또는 과도한 활성의 유전자 산물의 발현을 차단하기 위해 사용될 수 있다. 이러한 방법들은 또한 헌팅턴 질환(Huntington's disease), 유전성 파킨슨 질환 및 기타 질환들의 치료에 유용하다. 안티센스 핵산은 또한 독성과 관련된 단백질의 발현을 저해하는 데에도 유용하다.
- <237> 올리고뉴클레오티드는 올리고머가 이중-나선(double-helical) DNA 주위에 결합됨으로 인한 삼중화 방법(triplex strategy)으로 공지된 기작을 통해, 전사를 지연시키기 위해 사용할 수 있다. 따라서, 삼중 화합물들은 선택된 유전자의 특정 위치를 인식하도록 고안될 수 있다(Maher, Wold, and Dervan, 1991; Helene, 1991).
- <238> 리보자임은 DNA 제한 엔도뉴클리에이즈(endonuclease)와 유사한 방법으로 다른 단일-가닥 RNA를 특이적으로 절단할 수 있는 능력은 갖는 RNA 분자이다. 이러한 RNA를 코딩하는 핵산 서열들의 변형을 통해 RNA 분자에서 특정 핵산 서열을 인식하고 절단하는 분자들을 가공하는 것이 가능하다(Cech, 1988). 이러한 접근의 주된 장점은 특정 서열을 갖는 mRNA만을 불활성화시킨다는 것이다.
- <239> 생물학적 반응 변형 물질(biological response modifier)을 코딩하는 핵산을 전이 시키는 것이 바람직할 수 있다. 이러한 범주에는, 예를 들면, 인터루킨(interleukin) 1 내지 12의 "인터루킨"으로 분류되는 여러 사이토카인을 코딩하는 핵산 등의 면역강화제가 포함된다. 상기 범주에는 또한, 동일 기작에 따라 필수적으로 작용하지는 않으나, 인터페론(interferons), 특히 감마 인터페론(γ -IFN), 종양괴사인자(TNF) 및 과립구-대식세포 콜로니 자극 인자(GM-CSF) 등이 포함된다. 선천적 효소 결핍 또는 면역 결핍을 치료하기 위해 이러한 핵산을 골수 세포 또는 대식세포로 전달하는 것이 바람직하다. 성장인자, 독성 펩타이드, 리간드, 수용체 또는 기타 생리학적 중요 단백질을 코딩하는 핵산들이 또한 특정 비-분열 세포들에 도입될 수 있다.
- <240> 따라서, 본 발명의 재조합 CAEV 벡터 시스템은 항-HIV 분자로 HIV-감염된 세포(예: T-세포 또는 대식세포)를 치료하는데 사용될 수 있다. 또한, 예를 들면, 양성 섬유증의 치료를 위해 양성섬유증 전이막 전도 조절 물질(CFTR)에 대한 유전자를 갖는 본 발명의 재조합 렌티바이러스로 호흡기 상피세포를 감염시킬 수 있다.
- <241> 따라서, 본 발명의 재조합 CAEV 벡터 시스템은 여러 인간의 질병을 치료하는데 사용될 수 있다. CAEV 벡터 시스템이 적용 가능한 인간 질환의 대표적인 예로는, 이에 한정되는 것은 아니나, 알츠하이머 질환, 파킨슨 질환, 근위축성 측삭 경화증(amyotrophic lateral sclerosis disease), 헌팅턴 질환, 베타-용혈성빈혈 (beta-thalassemia), 색소성 망막염(retinitis pigmentosa), 점액성다당 질환(mucopolysaccharide disease), 백질이 양성 질환(leukodystrophy diseases), 반성유전성 SCID(X-linked SCID), 페닐케톤뇨증(phenylketonuria), 티로신혈증(tyrosinemia), A 및 B형 혈우병(hemophilia A and B), 윌슨 질환(Wilsons diseases), LDL 수용체 결핍(LDL receptor deficiency), 인간 면역 결핍(Human Immunodeficiency), 및 듀센형 근이영양증(Duchennes dystrophy) 등이 포함된다.
- <242> CAEV 벡터 입자
- <243> 본 발명의 방법에서, 감염성 및 복제-결핍 CAEV 벡터 입자들은 당분야에 공지된 기술과 접목된 본원에 개시된 방법에 따라 제조될 수 있다. 상기 방법에는 본 발명의 벡터 발현 시스템으로 렌티바이러스-허용 세포를 형질 감염시키는 단계; 형질감염된 세포에서 CAEV-유래 입자들을 생산하는 단계; 및 상기 세포로부터 바이러스 입자를 회수하는 단계들을 포함한다.
- <244> 용어 "형질감염(transfection)"은 외래 DNA를 진핵 세포로 도입하는 것을 의미한다. 형질감염은 이에 한정되지는 않으나 칼슘 포스페이트-DNA 공침전법, DEAE-텍스트란-매개 형질감염법, 폴리브렌-매개 형질감염법, 전기충격법, 미세주사법, 리포좀 융합법, 리포펙타민 및 원형질체 융합법 등의 당분야에 공지된 여러 방법에 의해 수행될 수 있다.
- <245> 본원에서 사용된 용어 "형질도입(transduction)"은 형질감염에 의해서라기 보다는 감염(infection)을 수단으로 하여 바이러스 또는 레트로바이러스 벡터 입자를 사용하는 유전자 전달을 의미한다. 어떤 실시양태에서, 레트로바이러스 벡터는 형질도입된다. 따라서, "형질도입된 유전자"는 렌티바이러스 또는 벡터 감염 및 프로바이러스 통합을 통해 세포로 도입된 유전자이다. 특정 실시양태에서, CAEV 바이러스 벡터 입자들은 "표적 세포" 또는 숙주 세포로 유전자를 형질도입한다.
- <246> 세포내 감염성 바이러스 입자의 생산을 촉진하는 단계는 또한 기본 세포 배양 성장 기법과 같이 통상적인 기술을 사용하여 수행될 수 있다.

- <247> 감염성 바이러스 입자를 회수하는 단계는 통상의 기술을 사용하여 수행될 수 있다. 예를 들면, 감염성 입자들은 공지된 바와 같이, 세포 배양액의 상등액을 회수하여 회수될 수 있다. 선택적으로, 회수된 바이러스 입자들은 필요에 따라 정제될 수 있다. 적절한 정제 기법은 통상의 기술자에게 공지되어 있다.
- <248> 통상의 기술자의 필요에 따라, CAEV 모액(stock solution)을 본 발명의 벡터 및 방법을 사용하여 제조할 수 있다. 바이러스 모액의 제조방법은 당분야에 공지되어 있으며, 예를 들면 [(Soneoka et al., 1995) 및 (Landau and Littman, 1992)]에 기재되어 있다. 본 발명의 모액을 생산하는 방법에서, 렌티바이러스-허용 세포들은 본 발명의 벡터 시스템으로 형질감염된다. 상기 세포들은 적절한 세포 배양 조건 하에 배양되고, 상술된 세포 배양액으로부터 CAEV 입자들을 회수한다. 적절한 허용 세포주로는 이에 한정되는 것은 아니나, 인간 세포주 293, 293T, 및 HeLa 원숭이 세포주 Vero, 및 염소 세포주 GSM 및 Ch1Es 등이 포함된다. 본 발명의 벡터는 또한 안정(stable) 패키징 세포(즉, 그 자체로서는 감염성 바이러스 입자를 만들어낼 수 없으면서 CAEV 구조 단백질들을 안정되게 발현하는 세포) 및 바이러스 생산 세포(virus producing cells; VPC)를 제조하는데 유용하다. 레트로바이러스 단백질을 발현하는 패키징 세포를 제조하는 방법은 당분야에 공지되어 있으며, 예를 들면, 미국 특허 제 4,650,764호(Temin et al.) 등에 이러한 방법이 실시되어 있으며, 이는 그대로 본원에 접목되어 있다. 본 발명의 범위 내에서, 패키징 세포는 본 발명에 제시된 최소 하나의 CAEV 패키징 벡터로부터의 CAEV 핵산 서열을 포함하는 렌티바이러스-허용 숙주 세포를 포함할 것이며, 상기 핵산 서열은 패키징-신호 결합이므로 세포가 복제-적격 감염성 바이러스를 생산할 수는 없으나, 최소 하나 이상의 CAEV 구조 단백질을 생산할 수 있도록 한다.
- <249> 패키징 세포는 CAEV-허용 숙주 세포(예: 인간 배아 신장 293 또는 293T 세포)를 상기 제시된 바와 같이 공지된 방법에 따라 적절한 CAEV 핵산 서열로 형질감염시켜 만들 수 있다. 따라서, 그 결과 얻어진 패키징 세포는 최소 하나의 CAEV 구조 단백질을 발현하고 생산할 수 있다. 그러나, 패키징 세포는 여전히 재조합 CAEV 바이러스를 생산할 수 없다. 패키징 세포는 그 후 다른 핵산 서열, 즉, 관심 이종 유전자 및 적절한 패키징 신호를 포함하는 전이 벡터로 형질감염될 수 있다. 일단 추가의 서열 또는 서열들로 형질감염되면, 상기 패키징 세포는 이종의 유전자를 포함하는 CAEV 바이러스 스탁(stocks)을 제공하는데 사용될 수 있으며, 이때 바이러스는 그 자체로서 복제-부적격이다. 따라서, 그 결과 얻어진 바이러스 생산 세포(VPC)는 이종의 관심 유전자를 포함하는 감염성 바이러스 입자들을 생산할 수 있다.
- <250> 유전자 전이 및 치료
- <251> 단일 유전자에서의 변경으로 인한 여러 인간 유전 질환은 유전자 치료의 주요 후보들이다. 본원에서 사용된 용어 "유전자 치료" 또는 "유전자 전이"는 임상 치료를 목적으로 세포내로 유전자를 삽입하는 것으로 정의된다. 유전자 치료의 많은 적용방법, 특히 줄기세포 유전자 삽입을 통한 방법들이 공지되어 있으며, 광범위하게 검토되고 있다. 생체외 또는 생체내 치료 유전자 전이를 위한 표적 세포로는 이에 한정되는 것은 아니나, 조혈 줄기 세포, 림프구, 혈관내피세포(vascular endothelial cell), 호흡기 상피 세포(respiratory epithelial cell), 케라틴 형성 세포(keratinocyte), 골근 세포(skeletal and muscle cells), 간세포(liver cell), 신경세포(neuron cell), 및 암 세포(cancer cell) 등이 포함된다.
- <252> 본 발명의 유전자 전이 기술은 또한 펩타이드 처리 공정의 확인 및 다양한 단백질의 기능 도메인의 확인에 사용될 수 있다. 처리 및 세포 운명에서의 세포-특이적 차이를 연구하기 위해, 단백질에 대한 클로닝된 cDNA 또는 유전체 서열을 생체 외 또는 생체 내의 서로 다른 표적 세포로 도입할 수 있다. 강력한 프로모터의 조절 하에 코딩 서열을 위치시킴으로써, 상당량의 목적 단백질을 만들 수 있다. 또한, 단백질 처리, 세포내 저장 또는 생물학적 활성에 관여하는 특정 잔기는 코딩 서열의 구분 잔기들의 변이적 변화에 의해 결정될 수 있다.
- <253> 본 발명의 유전자 전이 기술은 또한 단백질의 발현을 조절하고 그 세포수준의 현상을 조절하는 능력을 확인하는 수단을 제공하는데 적용될 수 있다. 분화에서의 그 역할과 같은 단백질의 특정 기능은 조직 배양에서 연구될 수 있는 반면, 다른 것들은 관련 특성에서의 변화를 관찰하기 위해 발달과정의 서로 다른 시기에 생체내 시스템으로의 재도입이 필요할 것이다.
- <254> 유전자 전이는 특정 유전자의 발현을 조절하는 핵산 서열 및 세포 인자들을 연구하기 위한 수단을 제공한다. 이러한 연구의 일례로 리포터 유전자에 조절 요소들을 융합시켜 리포터 유전자의 발현을 분석하는 연구가 있다.
- <255> 유전자 전이는 또한 질환 상태에 대해 이해하고 치료를 제공하는데 있어서 실질적으로 중요한 용도를 갖는다. 결합 유전자들이 공지되고 클로닝된 여러 유전 질환이 있다. 일부의 경우, 이러한 클로닝된 유전자의 기능은 공지되어 있다. 일반적으로, 상기 질환의 상태는 두 부류로 나눌 수 있다: 일반적으로 열성으로 유전되는, 일반적으로 효소에서, 결핍상태, 및 우성으로 유전되는, 최소한 때때로 조절 또는 구조 단백질에 관련된, 불균

형 상태. 결핍 상태 질환의 경우, 유전자 전이는 안티센스 돌연변이를 사용하여 상기 질환에 대한 동물 모델을 만들어낼 뿐 아니라, 대체 요법을 위해 발병 조직에 정상 유전자를 전달하기 위해 사용될 수 있다. 불균형 질환 상태의 경우, 유전자 전이는 모델 시스템에서 질환 상태를 만들어내기 위해서 사용될 수 있고, 이 모델 시스템은 질환 상태를 역전시키기 위한 연구에 사용될 수 있다. 따라서, 본 발명의 방법은 유전적 질환의 치료를 가능하게 한다. 본원에서 사용된 바와 같이, 질환 상태는 질환을 일으키고 점차 악화시키는 결핍 또는 불균형을 부분적으로 또는 전체적으로 치료함으로써 치료된다. 변이를 일으키거나 결함을 고치는 핵산 서열의 위치-특이적 통합도 역시 사용할 수 있다.

<256> 본 발명의 방법은 신경 세포, 아교 세포(glial cell), 섬유아세포(fibroblast) 또는 간엽세포(mesenchymal cell)의 이식(transplantation), 또는 본 발명의 재조합 렌티바이러스로 감염된 세포의 생체의 이식, 중추 신경계 또는 심실강(ventricular cavities)으로의 생체내 감염, 또는 숙주 뇌 표면의 경막하(subdurally) 감염과 관련된 "그래프팅(grafting)"에 유용할 수 있다. 이러한 그래프팅 방법은 통상의 기술자들에게 공지되어 있으며, 문헌 [Neural Grafting in the Mammalian CNS, Bjorklund & Stenevi, eds. (1985)]에 제시되어 있다.

<257> 단백질 산물의 결핍으로 인한 질병의 경우, 유전자 전이는 안티센스 변이를 사용하여 질병에 대한 동물 모델을 만들 뿐 아니라, 대체 요법을 위해 환부 조직으로 정상 유전자를 도입하는데 사용될 수 있다. 예를 들면, 근육, 비장 또는 간 세포의 감염을 위해 인자 VIII 또는 IX 코딩 핵산을 CAEV 입자내로 도입시킬 수 있다.

<258> 유전자 치료, 특히 줄기 세포 유전자 삽입을 통한 유전자 치료의 많은 적용방법이 공지되어 있고 광범위하게 검토되고 있다. 본원에서 사용된 용어 "줄기 세포"로는, 이에 한정되는 것은 아니나, 조혈 줄기 세포, 신경 줄기 세포, 간엽(특히 근육의) 줄기세포, 및 간 줄기 세포 등이 포함된다. 줄기세포는 생체내 조직들을 재증식(repopulation)할 수 있다. 조혈 줄기 세포는 발달 초기의 인간 조혈 세포로부터 유래된 전구 세포(progenitor cells)이다.

<259> 조혈 줄기 세포를 이용한 유전자 치료는 또한 일반적으로 결함 단백질의 생산 또는 이상(abnormal) 수준의 유전자 발현을 야기하는 림프계(lymphoid) 또는 골수계(myeloid) 세포들에서의 유전자 이상(abnormality)을 치료하는데 유용하다.

<260> 이러한 여러 질환에 대해, 결함 유전자의 정상 복사체(copy) 또는 기능적 상동체(homolog)의 도입, 및 소량일지라도 소실 유전자(missing gene)의 생산은 유익한 효과를 갖는다. 동시에, 유전자 산물의 과발현이 해로운 효과를 가져오리라고는 예상되지 않는다. 조혈 줄기 세포로의 유전자 전이가 매우 유용한 질환의 비한정적 예는 다음과 같다. 이러한 질환에는 일반적으로 골수 질환, 적혈구 세포 결함, 대사 질환 등이 포함된다. 조혈 줄기 세포 유전자 치료는 글로빈 유전자 또는 응고 인자 유전자(예: 인자 IX 및 인자 X 유전자)에 결함이 있는 α - 및 β -탈라세미아 빈혈, 겸상 적혈구 빈혈(sickle cell anemia) 및 A형 및 B형 혈우병(hemophilia A and B)과 같은 혈액 세포들의 유전자 질환 치료에 유익하다. 다른 바람직한 예로는 T 및 B 림프세포에 독성인 특정 부산물을 제거하도록 돕는 아데노신 디아미네이즈(adenosine deaminase; ADA) 효소가 결여되어 환자가 감염에 대해 무방비상태가 되는, 중증 합병성 면역결핍증 장애(severe combined immunodeficiency disease; SCIDS)의 치료가 있다. 이러한 환자들은 과거에 행해졌던 환자의 림프세포 대신, ADA 유전자를 그들의 조혈 줄기 세포로 도입하여 유전자 치료를 받기 위한 이상적인 후보들이다. 다른 질병들로는 호중구(neutrophils)가 결함 사이토크롬 b(cytochrome b)를 발현하는 만성 육아종증(chronic granulomatosis) 및 대식세포에서의 비정상 글루코세레브로시데이즈(glucocerebrosidase) 유전자 생산으로 인한 고세병(Gaucher disease) 등이 있다.

<261> 또한, 파킨슨 질환(Parkinsons disease) 등의 신경 퇴행성 질환은 GDNF (glial cell line-derived neurotrophic factor) 유전자를 선조체(striatum) 및 흑색질(substantia)로 도입시키는 유전자 치료의 좋은 표적이다(Kordower et al., 2000).

<262> 암의 다양한 형태를 치료하는 방법 또한 유전자 치료에 포함된다. CAEV 백터는 예를 들면, 암 세포를 특이적으로 사멸시키는데 효과적인 독소(toxin) 또는 세포 사멸(apoptosis) 유도 물질을 코딩하는 유전자를 포함할 수 있다. 종양 세포 특이적 사멸은 단지 종양세포만이 자살 유전자(suicide gene)를 발현한다는 조건 하에 암 조혈 세포로 자살 유전자를 도입하여 수행될 수 있다. 자살 유전자 산물은 일반적으로 비독성 약제를 독성 유도체로 전환시켜 세포에 치사 감수성을 부여한다. 예를 들면, 효소 사이토신 디아미네이즈는 비독성 물질 5'-플루오로사이토신을 독성 유도체인 5-플루오로우라실로 전환시킨다(Mullen, Kilstrup, and Blaese, 1992). 종양-특이적 림프 세포는, 예를 들면, 항-종양 활성을 갖는 유전자 산물을 종양 위치에 위치특이적으로 전달하여 이러한 유전자 산물의 전신 전달과 관련된 독성을 피하도록 유전학적으로 변형될 수 있다. 유전자 치료의 접근법은 또한 화학치료(chemotherapy)의 독성 효과에 대해 내성인 골수 세포들에 적용될 수 있다.

- <263> 유전자 치료는 또한 HIV 및 HTLV-1 감염과 같은 바이러스 감염을 예방하거나 치료하는데 사용될 수 있다. 예를 들면, 조혈 줄기 세포는 HIV 감염에 대해 내성을 부여하도록 유전학적으로 변형될 수 있다. 한 접근법은 특별히 안티센스 RNA를 사용하거나 존재하는 바이러스 조절 경로를 파괴하여 바이러스 유전자 발현을 저해한다. 레트로바이러스 RNA에 상보적인 안티센스 RNA는 HIV(Rhodes and James, 1991) 및 HTLV-1(von Ruden and Gilboa, 1989)을 포함하는 여러 레트로바이러스의 복제를 저해하는 것으로 나타났다(To, Booth, and Neiman, 1986).
- <264> 조혈 줄기 세포에서의 유전자 치료가 사용될 수 있는 다른 영역으로는 자가면역 질환(autoimmune disease)의 완화가 있다. 치료 유전자는 예를 들면, 자가반응성(autoreactive) 세포의 사멸 및 제거를 야기하는 정상 세포사멸 신호를 재구성할 수 있는, B 또는 T 세포 신호 분자를 코딩할 수 있다.
- <265> 진단, 연구 또는 유전자 치료를 위한 생체 외 세포 형질전환법(예: 형질전환된 세포의 숙주세포로의 재-융합을 통한)이 통상의 기술자에게 공지되어 있다. 본 발명의 일 실시양태에서, 세포는 대상 유기체로부터 분리되고, 관심 폴리펩타이드를 포함하는 본 발명의 벡터로 형질감염되며, 대상 유기체(예: 환자)로 다시 재융합된다.
- <266> 생체외 형질전환에 적합한 다양한 세포형이 통상의 기술자에게 공지되어 있다. 특히 바람직한 세포는 상술된 줄기 세포(예: Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, third edition Wiley-Liss, New York, 및 여기에 인용된 환자로부터의 세포를 분리하고 배양하는 방법에 관한 논의 참조)이다. 형질전환된 세포는 당분야에 공지된 방법(Kuchler (1977) Biochemical Methods in Cell Culture and Virology, Kuchler, R. J., Dowden, Hutchinson and Ross, Inc., 및 Atlas (1993) CRC Handbook of Microbiological Media (Parks ed) CRC press, Boca Raton, Fl. 참조)에 의해 배양된다. 포유류 세포 시스템은 종종 세포 단일층으로 형성될 것이나, 포유류 세포 부유물(suspension) 또한 사용될 수 있다. 한편, 세포는 세포은행(예: 혈액은행)에 저장된 것으로부터 유래될 수 있다. 포유류 세포주의 바람직한 예로는 HEC-1-B 세포주, VERO 및 Hela 세포, 중국 햄스터 난소(Chinese hamster ovary; CHO) 세포주, W138, BHK, Cos-7 또는 MDCK 세포주(예: Freshney, supra 참조) 등이 포함된다.
- <267> T 세포 또는 B 세포는 또한 일부 생체의 유전자 전이 공정에서 사용될 수 있다. T 및 B 세포를 분리해내는 여러 기법이 공지되어 있다. 표면 마커의 발현은 이러한 세포의 분리 및 정제를 용이하게 한다.
- <268> 요약하면, 본 발명의 바이러스 벡터는 분열 또는 비분열 세포를 안정되게 형질도입시키고 이종의 유전자를 안정되게 발현하는데 사용될 수 있다. 이러한 벡터 시스템을 사용하면, 상기 세포의 생리작용에 영향을 줄 수 있는 단백질 코딩 유전자를 분열 또는 비분열 세포로 도입하는 것이 가능하다. 따라서, 본 발명의 벡터는 질병 또는 세포 생리 작용의 실험적 변형을 위한 유전자 치료에 유용하게 사용될 수 있다.
- <269> 키트(Kits)
- <270> 본 발명의 또 다른 목적은 본원에 기술된 방법에 사용하기 위한 벡터를 포함하는 키트 또는 약제 전달 시스템을 제공하는 것이다. 모든 표적 레트로바이러스 입자의 투여에 요구되는 필수 물질 및 시약이 키트에 포함될 수 있다(예: 패키징 세포 구조물 또는 세포주). 키트의 구성성분은 다양한 제형으로 제공된다. 하나 이상의 CAEV 입자들은 하나 이상의 약제(예: 화학치료제)를 가지고 단일 약학적으로 허용가능한 조성물 또는 분리된 약학적으로 허용가능한 조성물로 제형화될 수 있다.
- <271> 이러한 키트 또는 약제 전달 시스템의 구성원은 또한 건조 또는 감압 동결건조된 형태로 제공될 수 있다. 시약 또는 구성원이 건조 형태로 제공되었을 경우, 일반적으로 다른 용기 수단에 제공될 수 있는 적절한 용매의 첨가로 인해 재구성된다. 본 발명의 키트는 또한 표적 CAEV 입자에 대한 복용 및/또는 투여 정보와 관련된 지시서를 포함한다. 본 발명의 키트 또는 약제 전달 시스템으로는 또한, 바이얼(vial)이 유지되도록 일반적으로 주사 또는 블로우 몰딩된(blow-molded) 가소성 용기(container)와 같은 시관용 밀집 바이얼 등의 수단이 포함될 것이다. 상기 용기의 수와 형태에 상관 없이, 키트는 또한 최종 구조체 조성물의 대상 체내 주사/투여 또는 위치이동을 돕기 위한 기구를 포함하거나 또는 이로 패키징될 수 있다. 이러한 기구는 어플리케이터(applicator), 흡입기, 시린지(syringe), 피펫(pipette), 포셉(forceps), 측정 수저(measured spoon), 안구-점적기(eye-dropper) 또는 이러한 임의의 의학적으로 승인된 전달 물질일 수 있다.
- <272> 이하 실시예는 본 발명을 다양한 측면에서 설명하나, 본 발명의 방법을 한정하는 것은 아니다.
- <273> **실시예**
- <274> 하기 실시예들은 본 발명의 특정 실시양태 및 측면을 제시하고 있으며, 본 발명의 범위를 한정하는 것은 아니다.

- <275> 하기 실시예들은 본 발명의 재조합 CAEV-계 렌티바이러스 벡터 시스템이 공지된 HIV-1 계 렌티바이러스 시스템에서와 같이 발현에 있어서 효과적이라는 발견을 입증한다. 상기 실시예들은 본 발명의 CAEV-계 벡터 입자 생산 시스템의 유전체 RNA 전사, 캡시드화, 번역, 역전사, 및 통합 수준이, 오랜 기간 매우 효과적인 유전자 전이 시스템으로 인정받고 있는 HIV-1-계 렌티바이러스 벡터 시스템(Naldini et al., 1996)과 비교할만 하다는 것을 보여준다.
- <276> 본 발명은 최소 세 개의 플라스미드 공-형질감염 방법을 근거로 하면서, *gag-pol* 및 *env* 유전자, 및 선택적으로 *rev* 유전자의 발현을 요구하는, 높은 역가의 CAEV-계 벡터 시스템의 구조에 대한 첫 번째 보고이다.
- <277>
- <278> 물질 및 방법
- <279> 플라스미드 구조
- <280> 모체 플라스미드(parent plasmids). 본 발명의 CAEV 벡터가 유래된 모체 플라스미드는 마리 수잔 박사(Dr. Marie Suzan, Institut National de la Sante et de la Recherche Medicale INSERM, France)에 의해 제공된 플라스미드 pWTE-BM 및 플라스미드 pCAEV-LTR이다. pWTE-BM 플라스미드는 *env*, *rev*, 및 U3 영역을 포함하는 0.4 kb *Hind* III 절편을 제외한 전장 유전체 CAEV cDNA 및 1337 염기쌍의 비필수 절편(stuffer fragment)을 포함한다. 플라스미드 pCAEV-LTR는 pWTE-BM에서 제외된 0.4 kb *Hind* III 절편을 포함한다(Saltarelli et al., 1990; Saltarelli, 1993). 두 벡터 모두 야생형 바이러스를 생산할 수 없다.
- <281> CAEV *gag-pol* 발현 벡터(pMGP/RRE(서열번호: 77) 및 pMGP/REV/RRE). pMGP/RRE (서열번호: 77) 플라스미드는 pWTE-BM 유래 *gag-pol* 발현 플라스미드이다(도 2a에 제시됨). pMGP/RRE(서열번호: 77) 플라스미드는 강력하고 이종인 MCMV 주 극초기 프로모터(MCMV MIEP), *gag-pol* 유전자, 및 *rev* 반응 요소(RRE)를 포함한다. pMGP/RRE (서열번호: 77) 플라스미드는 또한 항생물질 선별 마커로서 네오마이신 내성 유전자를 코딩한다. 상기 플라스미드 구조에서, pWTE-BM 유래 *gag-pol* 유전자 절편(CAEV 유전체의 512 내지 5046번째 뉴클레오티드)은 몇몇 PCR 및 서브클로닝 단계에 대한 표준방법을 사용하여 pGL2-Basic(Promega, WI, USA) 클로닝 벡터로 서브클로닝 되었다. MCMV MIEP 절편은 플라스미드 pMYK(Kim et al., 2002)로부터 절단되어 *gag* 유전자의 상류에 삽입되었고, RRE 영역(CAEV 유전체의 7824 내지 8183번째 뉴클레오티드 또는 7849 내지 8150번째 뉴클레오티드)은 *pol* 유전자의 하류에 삽입되었다. pMGP/REV/RRE는 CAEV *rev* 유전자를 포함하는 다른 *gag-pol* 발현 플라스미드이다(도 2b에 제시됨). 또한, CAEV의 주 스플라이싱 제공 위치(CAEV 유전체의 330 내지 346번째 뉴클레오티드)가 MCMV 프로모터의 하류에 삽입되었다.
- <282> 전이 벡터(pCAH/SINd 시리즈). pCAH/SINd 시리즈에 속하는 플라스미드들(도 3a-3h에 제시됨) (서열번호: 67-71, 73, 78, 및 79)은 본 발명의 전이 벡터 설계를 위한 최적의 패키징 서열을 확인하기 위해 제작되었다. 상기 시리즈내 각 플라스미드는 서로 다른 길이의 5' 비번역 영역 및 *gag*-코딩 영역의 개시부분을 포함하도록 설계되어 이러한 영역에서의 다양한 길이의 효과를 나란히 비교할 수 있게 하였다. 특히 안전성 문제를 처리하기 위해, 이러한 플라스미드는 3'LTR이 삭제된 U3 영역을 갖는 SIN(self-inactivation) 벡터로서 설계되었다. 트랜스-작용 인자인 *tat*가 없는 전이벡터로부터 벡터 RNA가 높은 수준으로 발현되도록, 5'LTR의 U3 영역을 HCMV MIEP로 치환하였다. 또한, 폴리아데닐화, RNA 수송, 역전사, 및 통합에 요구되는 모든 공지된 시스-작용 서열 요소가 상기 전이 벡터 시리즈에 포함되었다.
- <283> pCAH/SINd 시리즈의 플라스미드들(서열 번호: 67-71, 73, 78, 79)은 다음과 같은 방법으로 제작되었다. pCAH/SINd(PBS-결핍 음성 대조군 벡터)(서열번호: 73)(도 3a)는 단지 5'LTR 내 5' 비번역 서열(R 및 U5 영역)(CAEV 유전체의 1 내지 163번째 뉴클레오티드)만을 포함하도록 설계되었다. pCAH/SINd0(서열번호: 67)(도 3c)는 5' 비번역 영역 전부(CAEV 유전체의 1 내지 511번째 뉴클레오티드)를 포함하도록 설계되었다. pCAH/SINd1(서열번호: 68)(도 3c)는 5' 비번역 영역 전부 및 점변이를 갖는 *gag* 유전자의 327 bp 절편(CAEV 유전체의 1 내지 839번째 뉴클레오티드)을 포함하도록 설계되었다. pCAH/SINd2(서열번호: 69)(도 3d)는 5' 비번역 영역 전부 및 점변이를 갖는 *gag* 유전자의 612 bp 절편(CAEV 유전체의 1 내지 1124번째 뉴클레오티드)을 포함하도록 설계되었다. 플라스미드 pCAH/SINd3(서열번호: 70)(도 3e)는 5' 비번역 영역 전부 및 점변이를 갖는 *gag* 유전자의 908 bp 절편(CAEV 유전체의 1 내지 1420번째 뉴클레오티드)을 포함하도록 설계되었다. 플라스미드 pCAH/SINd4(서열번호: 71)(도 3f)는 5' 비번역 영역 전부 및 점변이를 갖는 *gag* 유전자의 1,198 bp 절편(CAEV 유전체의 1 내지 1710번째 뉴클레오티드)을 포함하도록 설계되었다. pCAH/SINd1/hlacZ(서열번호: 78)(도 3g)는 HCMV MIEP 및 *lacZ* 유전자로 이루어진 발현 카세트를 pCAH/SINd1(서열번호: 68) 내로 삽입시켜 제작되었

다. 플라스미드 pCAH/SINd60/hlacZ(서열번호: 78)(도 3h)는 점변이를 갖는 *gag* 유전자의 처음 60 bp 절편을 포함하는 *gag* 유전자의 길이(CAEV의 1 내지 569번째 뉴클레오티드)에 대한 것을 제외하고는, pCAH/SINd1(서열번호: 68)에서와 동일한 설계를 갖는다.

- <284> CAEV *vif* 발현 벡터(*pHYK/vif*)(서열번호: 76). 신속하고 효율적인 바이러스 복제를 위해 요구되는 것으로 알려진 *vif* 유전자(CAEV 유전체의 5006 내지 5695번째 뉴클레오티드)는 진핵세포 발현 벡터 *pHYK*(Kim et al., 2002)(도 4)에 클로닝되었다.
- <285> CAEV *rev* 발현 벡터(*pHYK/rev*)(서열번호: 75). RRE와의 상호작용으로 바이러스 유전자 발현을 전사 후 수준에서 조절하는 *rev* 유전자는 두 개의 엑손(제1 엑손은 CAEV 유전체의 6,012 내지 6,123번째 뉴클레오티드, 그리고 제2 엑손은 CAEV 유전체의 8514 내지 8803번째 뉴클레오티드)으로 이루어진다. Rev/RRE 시스템은 비-스플라이싱된 RNA의 핵 배출을 촉진하며 렌티바이러스 복제에 필수적인 것으로 알려져 있다. *rev* 유전자의 전장 cDNA는 RT-PCR에 의해 합성되었으며 *pHYK* 벡터로 서브클로닝되었다(도 5).
- <286> 바이러스 외피 유전자 발현 벡터. 본원에서 사용된 외피 유전자 발현 벡터 시스템은 플라스미드 *pHGVSV-G*(서열번호: 74) 및 플라스미드 *pMYKEF1/env*(서열번호: 72)(도 6a 및 6b)이다. 플라스미드 *pHGVSV-G*(서열번호: 74)는 소수포성 구내염 바이러스-G(VSV-G) 당단백질을 발현하도록 설계되었으며, 프로모터로서 β -글로빈 인트론을 갖는 HCMV MIEP를 포함한다. *pMYKEF1-env*(서열번호: 72)는 긴팔원숭이 백혈병 바이러스(GaLV) 외피 단백질을 발현하도록 고안되었으며 프로모터로서 진핵세포 연장 인자-1 α 를 갖는 MCMV MIEP를 포함한다.
- <287> *MuLV*- 및 *HIV-1*-계 플라스미드. 대조군 벡터 시스템으로서, *pMFG/lac/Zpuro* 및 *pHR/lacZ* 벡터들이 본 발명에 사용되었으며, 이들은 각각 쥐 백혈병 바이러스(*MuLV*)(Kim et al., 1997) 및 1형 인간 면역결핍 바이러스(*HIV-1*)(Naldini et al., 1996)로부터 유래된 *lacZ*-함유 레트로바이러스 벡터였다. *MuLV* 및 *HIV-1* 벡터 시스템의 패키징 플라스미드로는, *pEQPAM3*(Persons et al., 1998) 및 *pCMV Δ R8-2*가 각각 사용되었다. *HIV-1* 패키징 플라스미드 *pCMV Δ R8-2*는, *HIV-1 vpu* 유전자를 코딩하고 *env* 유전자에서 1.3-kb *Bgl*III 절편이 삭제된 것을 제외하고, *pCMV Δ R9*(Naldini et al., 1996)와 동일하다.
- <288>
- <289> 벡터 입자 생산
- <290> 의사형 CAEV-계 렌티바이러스 벡터 입자들은, 형질감염 하루 전에 6-웰 배양 디쉬(dish) 당 5×10^5 세포 농도로 분주된 293T 세포들에, 세 개 이상의 플라스미드의 리포솜 매개 일시적 형질전환을 수행하여 생산하였다. 세 개의 플라스미드 공형질감염은 *gag-pol* 발현 플라스미드, 전이 벡터 플라스미드, 및 *env*-코딩 플라스미드의 1:1:1 몰비로 수행하였다. 네 개의 플라스미드 공형질감염은 *gag-pol* 발현 플라스미드, 전이 벡터 플라스미드, *env*-코딩 플라스미드, 및 *rev*-발현 플라스미드의 3:3:3:1 몰비로 수행하였다. 다섯 개의 플라스미드 공형질감염은 *gag-pol* 발현 플라스미드, 전이 벡터 플라스미드, *env*-코딩 플라스미드, *rev*-발현 플라스미드 및 *vif*-발현 플라스미드의 3:3:3:1:1 몰비로 수행하였다. 바이러스 벡터 입자를 포함하는 배양 상등액은 48 시간 후에 회수되고, 0.45 μ M 막 여과로 여과한 후(Nalgene, NY, USA), 바로 사용하거나 -70°C 딥-프리저(deep-freezer)에 보관하였다.
- <291> 생체 내 형질도입
- <292> 형질도입은 바이러스 벡터 입자를 8 μ g/ml 폴리브렌(polybrene) 존재하에 293T 세포에 4시간 동안 가한 후 신선한 배지를 첨가하여 수행되었다. 48시간 후, 상기 세포를 1% 폼알데하이드 및 0.2% 글루타알데하이드로 이루어진 용액으로 고정시키고 300 μ g의 5-브로모-4-클로로-3-인돌릴 β -D-갈락토사이드(X-Gal, Promega, WI, USA), 4 mM 포타슘 페로시아나이드(potassium ferrocyanide), 4 mM 포타슘 페리시아나이드(potassium ferricyanide) 및 2 mM MgCl₂를 포함하는 용액으로 37°C에서 12시간 동안 염색한 후, 베타-gal 발현(Beta-Gal expression)을 분석하였다. 역가는 ml 당 LacZ-형성 단위(LacZ-forming units)(LFU/ml)로서 청색 포커스(foci)의 수를 측정하여 결정될 수 있다.
- <293> RT-PCR 분석
- <294> 트리졸 LS 시약(TRIZOL LS Reagent, GIBCO BRL, CA, USA)을 사용하는 방법에 의해 배양된 세포 또는 세포 상등액으로부터 총 RNA를 추출하였다. 상기 총 RNA는 RNA 분해효소가 없는 DNA 분해효소 I(RNase free-DNase I, 1 unit/mg의 DNA를 37°C에서 20분간 처리, Promega, WI, USA)으로 처리하여 DNA 오염물을 제거하였다. DNA 분해효소 I의 반응은 상기 DNA 분해효소와 함께 제공되는 RQ1 DNA 분해효소 중지 용액(RQ1 DNase stop solutio

n)을 가하여 증지시켰으며, RNA는 알앤이지 미니 키트(RNeasy mini kit, Qiagen, Germany)를 사용하는 방법을 통해 정제되었다. 정제된 RNA를 역전사(RT) 반응(37°C에서 90분)을 통해 cDNA로 역전사하였다. 구체적으로, RT 반응은 MuLV 역전사효소, 올리고-dT 프라이머 또는 C-말단 특이적 프라이머, 및 dNTPs 혼합물 존재하에 수행되었다. PCR 증폭은 주형 DNA의 반-정량 분석(semi-quantitative analysis)을 위해 특이적 프라이머를 가지고 수행되었다. 구체적으로, PCR 산물 DNA는 열안정성 Ex Taq 중합효소, 서열 특이적 DNA 프라이머, 및 dNTPs 혼합물 존재하에 cDNA 또는 염색체 DNA로부터 합성되었다.

<295> 서던 블롯 분석(Southern Blot Analysis)

<296> 유전체 DNA는 디앤이지 조직 키트(DNeasy Tissue Kit, Qiagen, Germany)를 사용하여 의사형 HIV-1 또는 CAEV 벡터 입자 중 하나로 형질도입된 세포, 및 비-형질도입된 대조군 세포(mock-transduced control cells)로부터 제조되었다. HIV-1 벡터로 형질도입된 세포로부터의 유전체 DNA 10 µg을 *Bam*H I 및 *Kpn* I으로 절단하였다. CAEV 벡터 형질도입된 세포로부터의 유전체 DNA 각 10 µg을 *Eco*R I 및 *Ssp* I으로 절단하였다. 절단된 유전체 DNA들을 0.7% 아가로스 겔 상의 전기영동에 의해 분리하고 양이온 전하된 나일론막(Roche, Germany)으로 전이시켰다. Dig-표지된 프로브를 *lacZ* 유전자 특이적인 프라이머(전방향 프라이머: CTGGCGTAATAGCGAAGAGG(서열번호: 65), 역방향 프라이머: AACTCGCCGCACATCTGAAC(서열번호: 66))들을 이용한 PCR에 의해 제조하였으며, 서던 혼성화는 Dig 적용 매뉴얼(Dig application manual, Roche, Germany)에 따라 수행되었다.

<297> 세포의 성장 정체(Growth Arrest) 및 성장 정체된 세포의 FACS 분석

<298> 293T 세포들을 아피디콜린(aphidicolin, Sigma, USA) 처리(25 µg/ml)로 성장정체시켰으며, 그 후 CAEV 바이러스 벡터 입자들로 형질도입시켰다. 양성 대조군 또는 음성 대조군으로서, 세포들을 HIV-1 벡터 또는 MuLV 레트로바이러스 벡터 중 하나로 나란히 형질도입시켰다. 형질도입후 2일에, 세포들의 베타-gal 활성을 알아보기 위해 X-gal로 염색시켰다. 아피디콜린 처리된 세포에서, 아피디 콜린은 감염 전 또는 후에 존재하였다.

<299> 세포의 성장 정체를 FACS 분석에 의해 확인하였다. 아피디콜린 처리된 또는 처리되지 않은 대조군 세포들을 PBS로 세척하고, 70% 에탄올로 -20°C에서 밤새 고정시켰으며, 프로피듐 아이오다이드(100 µg/ml)(Sigma, USA) 및 RNase A(100 µg/ml)(Qiagen, Germany)로 실온에서 1 시간 처리하였다. 세포들을 FACS 분석에 의해 분석하였으며, 세포 주기의 G1, S 및 G2/M 단계에서의 총 살아있는 세포들의 백분율을 산출하였다(Becton Dickinson, Sanjose, CA).

<300> 실시예 1: CAEV-계 렌티바이러스 벡터 입자의 생산

<301> 복제 결합 렌티바이러스 벡터 입자들은 CAEV *gag-pol* 발현 플라스미드, CAEV *env*-발현 플라스미드 및 전이 벡터 플라스미드의 최소 3-플라스미드 시스템을 이용한 인간 293T 세포의 일시적 공-형질감염에 의해 생산되었다. 4-플라스미드 시스템에서는, CAEV *rev* 발현 플라스미드가 추가되고, 5-플라스미드 시스템에서는, CAEV *vif* 발현 플라스미드가 추가된다. 효율적인 패키징을 위해, 전이 벡터는 *gag* 단백질의 발현을 방지하기 위해 시작 ATG 코돈 및 하류에 위치한 ATG 코돈에 변이(ATG에서 TAG로)가 도입된, *gag*-코딩 서열의 처음 부분을 포함하도록 고안되었다. RRE는 패키징 효율을 증대시키기 위해 포함되었으며, 4- 및 5-플라스미드 시스템에서의 *rev*는 CAEV mRNA 배출을 위해 벡터로부터 발현되었다. 전이 벡터 플라스미드에서의 내부 HCMV-MIEP 프로모터로 작동되는 β-갈락토시데이즈 유전자는 리포터 유전자로서 삽입되었다. 5'LTR의 U3 영역은 강력한 바이러스 프로모터인, HCMV-MIEP로 대체되어 벡터 유전체가 *tat* 독립적이 되도록 하였다.

<302> 전이 벡터 RNA 전사 수준. 전이 벡터로부터의 유전체 RNA의 전사 수준은 패키징 세포로부터의 재조합 바이러스 벡터의 높은 역가 생산에 매개하는 중요한 요소 중 하나이다. 본 발명에서, HCMV 인핸서/프로모터 요소는 전이 벡터 RNA의 안전하고 효율적인 전사를 위한 HCMV/CAEV 하이브리드 LTR 프로모터 시스템을 제작하는데 사용되었다. 하이브리드 LTR 프로모터를 포함하는 pCAH/SINd(서열번호: 67-71, 73, 78, 및 79) 시리즈의 전이 벡터 플라스미드들의 전사 수준을 시험하기 위해, 각 전이벡터 플라스미드를 패키징 플라스미드(pMGP/RRE (서열번호: 77), pHYK/rev (서열번호: 75), pHYK/vif (서열번호: 76), pHGVSV-G (서열번호: 74) 또는 pMYKEF1/env (서열번호: 72))와 함께, 리포솜-매개 형질감염에 의해 인간 T 세포로 도입시켰다. 48시간 배양 후, 형질감염된 세포로부터 총 RNA를 정제하여, 벡터 RNA 전사물을 측정하기 위한 역전사 중합효소 연쇄 반응(RT-PCR) 분석을 수행하였다. CAEV 전이 벡터에 대한 PCR 프라이머 세트(RRE primer set)는 RRE 영역 부분을 코딩하는 348-bp PCR 산물을 합성하도록 설계되었다. HIV-1 전이 벡터인 pHRLacZ(Naldini et al., 1996)에 대한 다른 PCR 프라이머 세트(*lacZ* primer set)는 *lacZ* 유전자를 코딩하는 645 bp PCR 산물을 합성하도록 설계되었다. 도 7에 나타낸 바와 같이, 본 발명의 CAEV 전이 벡터는 HIV-1-계 렌티바이러스 전이 벡터와 비교할만한 수준의 RNA 전사물을

생산하였다.

- <303> **백터 입자의 형성 및 방출.** 성숙 및 감염성 바이러스 백터 입자의 형성 및 방출을 시험하기 위해, pMGP/RRE (서열번호: 77) *gag-pol* 발현 플라스미드, pHGVSV-G(서열번호: 74) *env* 발현 플라스미드, pHYK/rev(서열번호: 75) *rev* 발현 플라스미드, pHYK/vif(서열번호: 76) *vif* 발현 플라스미드, 및 pCAH/SINd60/hlacZ(서열번호: 78) 전이 백터 플라스미드를 인간 293T 세포(DuBridge et al., 1987)로 리포솜-매개 공-형질감염시켜 CAEV 백터 입자를 생산하였다. 형질감염 후 48시간에, 감염을 위해 배양 상등액을 형질감염된 세포로부터 회수하여 8 µg/ml 폴리브렌 존재하에 신선한 인간 293T 세포에 적용시켰다. 그 결과는 본 발명의 5개의 플라스미드 시스템이 MuLV-계 렌터 바이러스 백터 시스템(pEQPAM3, pMFG/lacZ/puro, pHGVSV-G (서열번호: 74))(Ory, Neugeboren, and Mulligan, 1996; Persons et al., 1998)의 것과 비교할만한 바이러스 백터 입자 역가를 생산할 수 있음을 보여주었다(도 8에 제시).
- <304> 실시예 2: 백터 입자 생산에 대한 *Rev* 및 *Vif* 발현 효과
- <305> 백터 입자 생산에 대한 CAEV *rev* 및 *vif* 조절 유전자 발현의 효과를 결정하기 위해, (1) *rev*- 및 *vif*-코딩 서열을 갖지 않는, 3-플라스미드 시스템(pCAH/SIN, pMGP/RRE(서열번호: 77), pHGVSV-G(서열번호: 74) 또는 pMYKEF1/*env*(서열번호: 72)), (2) *vif*-코딩 서열을 갖지 않는, 4-플라스미드 시스템(pCAH/SIN, pMGP/RR (서열번호: 77), pHGVSV-G(서열번호: 74) 또는 pMYKEF1/*env*(서열번호: 72), pHYK/rev(서열번호: 75)), 및 (3) *rev*- 및 *vif*-코딩 서열을 모두 포함하는, 5-플라스미드 시스템(pCAH/SIN, pMGP/RRE(서열번호: 77), pHGVSV-G(서열번호: 74) 또는 pMYKEF1/*env*(서열번호: 72), pHYK/rev(서열번호: 75), pHYK/vif(서열번호: 76))의 백터 입자 생산 시스템을 백터 입자 생산에서의 그 효율성에 대해 나란히 시험하였다. 각 시스템의 플라스미드들을 293T 세포로 형질감염시켰다. 형질감염 후 2일에, 전이 백터 RNA 및 비리온 RNA를 형질감염된 세포 및 형질감염된 세포의 배지로부터 각각 추출하여, 전이 백터 RNA 유전체를 검출하기 위해 *lacZ* 프라이머 세트와 함께 RT-PCR 주형으로 사용하였다.
- <306> 도 9에 나타난 바와 같이, 패키징 세포에서의 전이 백터 RNA의 발현 수준이 *rev* 또는 *vif* 유전자 발현에 비의존적이기는 하나(도 9에서의 레인 1, 2 및 3), *rev*가 없이 캡시드화된 전이 백터 RNA의 양(도 9의 레인 4)은 *rev* 존재하에서의 양(도 9의 레인 5)보다 훨씬 낮았다. 그러나, 놀랍게도, *vif* 존재하에 캡시드화된 RNA를 가지고 RT-PCR로 측정된 백터 입자의 역가(도 9의 레인 6)는 CAEV *vif*가 없을때의 양(도 9의 레인 5)보다 낮았다. 이러한 결과는 백터 입자 생산을 위해 CAEV *rev* 및 *vif*가 요구되지는 않으나, *rev*는 효율적인 백터 입자 생산을 위해 바람직하다는 것을 의미한다.
- <307> *vif* 발현과 관련된 본 발명의 결과는, *vif* 유전자가 염소 활막 세포에서 CAEV의 효율적 복제를 위해 필수적이며 바이러스 복제 주기의 말기(예: RNA 캡시드화, 숙주 세포로부터의 바이러스 입자 방출)에 영향을 미친다는, 하르마셰 등(Harmache et al., 1995; Harmache et al., 1996)에 의해 보고된 결과와 일치하지 않는다. 상기 불일치에 대한 적절한 설명으로는 재조합 CAEV 백터 입자의 생산에서 염소 세포 대신 인간 293T 세포를 사용했기 때문일 수 있다. 이러한 해석은 *vif* 및 바이러스-생산 세포가 바이러스 감염성에 대한 *vif* 작용을 조절할 수 있다는 세로드 등에 의해 제시된 가설(Seroude et al., 2002)을 지지한다.
- <308>
- <309> 실시예 3: 최적의 패키징 신호 서열의 확인
- <310> CAEV 전이 백터 RNA의 캡시드화를 위한 최적의 패키징 신호 서열을 확인하기 위해, CAEV *gag*-코딩 영역의 서로 다른 부분들 및 5'LTR과 *gag* 시작 코돈 사이의 비번역 영역을 포함하는 시리즈 플라스미드들을 다음과 같은 방법으로 이들의 백터 입자 생산 효율성에 대해 비교하였다. 인간 293T 세포를 pMGP/RRE(서열번호: 77) *gag-pol* 발현 플라스미드, pHGVSV-G(서열번호: 74) *env* 발현 플라스미드, pHYK/rev(서열번호: 75) *rev* 발현 플라스미드, pHYK/vif(서열번호: 76) *vif* 발현 플라스미드, 및 pCAH/SINd(서열번호: 67-71, 73, 78, 및 79) 전이 백터 시리즈 플라스미드로 공-형질감염시켰다. 음성 대조군으로서, CAEV 전이 백터 pCAM/lacZ(L)를 패키징 플라스미드 없이 형질감염시켰다. 형질감염 후 2일에, 비리온 RNA를 형질감염된 세포의 배양액으로부터 추출하여 CAEV 전이 백터 시리즈 RNA 유전체를 검출하기 위한 RRE 프라이머 세트 또는 HIV-1 전이 백터 RNA 유전체를 검출하기 위한 *lacZ* 프라이머 세트와 함께 RT-PCR 주형으로 사용하였다. 도 10에 나타난 바와 같이, *gag* 영역의 처음 327 bp 뿐 아니라 전장 5'LTR을 포함하는 pCAH/SINd1(서열번호: 68)로 형질감염된 바이러스 생산 293T 세포로부터 회수된 배양액에서 바이러스 RNA를 포함하는 바이러스 입자의 효율적인 방출을 의미하는 강력한 PCR 산물 신호를 확인하였다(도 10의 레인 3). 이러한 신호는 양성 대조군인 HIV-1 백터를 이용하여 얻어진 것과

비교할만한 것으로, 본 발명의 캡시드화된 CAEV 전이 벡터 RNA의 양이 HIV-1-계 전이 벡터의 것(도 10의 레인 8)과 비교할만하다는 것을 의미한다. *gag*-코딩 영역의 처음 612 bp 또는 그 이상을 갖는 CAEV 전이 벡터의 패키징 효율은 현저히 감소되었다(레인 4, 5 및 6). 전이 벡터가 *gag*-코딩 서열 없이 사용되었을 때에는 PCR 산물 신호가 검출되지 않았다(도 10의 레인 1 및 2). 음성 대조군은 전이벡터만으로 형질감염되었으며, 양성대조군인 HIV-1 벡터는 pCMVΔR8-2, pHR/lacZ 및 pHGVSV-G(서열번호: 74)와 함께 형질감염되었다(도 10의 레인 7 및 8).

<311> 결론적으로, 전이 벡터가 5'LTR과 *gag* 개시 코돈 간의 비번역 영역 전부 뿐 아니라 N-말단 *gag*-코딩 서열의 약 600 bp 미만을 포함할 때, 전이 벡터 RNA는 패키징 세포에서 효율적으로 캡시드화되었다. 이러한 결과는 패키징 서열 내 RNA의 2차 구조의 역할이 RNA 캡시드화에서의 1차 구조보다 더 중요하다는 것을 의미한다.

<312> 실시예 4: CAEV 벡터 비리온의 의사형화

<313> 제조합 CAEV 벡터 비리온이 VSV-G 당단백질 뿐 아니라 GaLV 당단백질로 의사형화될 수 있는지를 결정하기 위해, GaLV 발현 벡터 및 pMYKEF1/env(서열번호: 72), 또는 VSV-G 발현 벡터 및 pHGVSV-G(서열번호: 74)로 전이 벡터 플라스미드 및 패키징 플라스미드와 함께 인간 293T 세포를 공형질감염시켰다. 형질감염 후 42시간에, 형질감염된 세포로부터 방출된 의사형 비리온 입자들을 포함하는 배양 상등액을 회수하고, 0.45 μm 막 여과로 정화하여, 293T 인간 표적 세포를 감염시키기 위해 사용하였다. 감염 하루 후, 유전체 DNA를 유전체 DNA 분리 키트(Genomic DNA Isolation kit, Qiagen, HL, Germany)를 사용하여 정제한 후, PCR 실험을 수행하여 통합된 프로바이러스 cDNA를 검출하였다. 예상했던 바와 같이, CAEV 벡터(도 11의 레인 1)는 MuLV-(도 11의 레인 3) 및 HIV-1-계 벡터(도 11의 레인 4)에 비해 VSV-G 단백질에 의해 효율적으로 의사형화되었다. 또한, HIV-1 렌티바이러스 벡터 시스템과는 다르게, 본 발명의 CAEV 벡터는 GaLV 외피(도 11의 레인 2)로도 성공적으로 의사형화되었다. 이러한 CAEV 벡터의 GaLV 외피를 이용한 의사형화 능력은 임상 수준의 렌티바이러스 벡터 시스템의 개발에 큰 장점을 부여할 수 있다. MuLV(pEQPAM3, pMFG/lacZ/puro 및 pHGVSV-G(서열번호: 74)로 형질감염됨) 및 HIV-1(pCMVΔR8-2, pHR/lacZ 및 pHGVSV-G(서열번호: 74)로 형질감염됨) 벡터 대조군은 레인 3 및 4에 각각 제시되어 있다.

<314> 실시예 5: CAEV 패키징 세포주의 생산

<315> pMGP/RRE(서열번호: 77) 및 pHYK/rev(서열번호: 75) 벡터는 진핵세포에서의 선별을 위한 *neo^r* 유전자를 코딩한다. *gag-pol* 및 *rev* 발현 벡터로 공형질감염 후 효율적인 선별을 위해, *neo^r* 유전자를 박테리아 *gpt* 유전자와 같은 다른 항생제 내성 유전자로 치환시켜 다른 CAEV *gag-pol* 발현 벡터를 제작할 수 있다. 한편, *gag*, *pol* 및 *rev* 유전자를 코딩하는 패키징 플라스미드 시스템이 사용될 수 있다. CAEV 패키징 단백질을 발현하는 안정(stable) 293T 세포가 만들어질 수 있는지를 결정하기 위해, 선별 배지하에 항생제 내성 콜로니가 선별된다. 안정 293T 세포로부터의 제조합 CAEV 벡터의 생산은 CAEV 벡터 생산을 위한 안정 패키징 세포주 생산의 용이성을 보여준다.

<316> 실시예 6: 숙주 염색체내로의 CAEV-계 벡터 cDNA의 통합

<317> 형질도입 후 CAEV 벡터 cDNA의 통합을 시험하기 위해, pMGP/REV/RRE *gag-pol* 발현 플라스미드, pHGVSV-G(서열번호: 74) *env* 발현 플라스미드, 및 pCAH/SINd1/hlacZ(서열번호: 79) 전이 벡터 플라스미드를 인간 293T 세포로 리포솜-매개 공-형질감염시켜 CAEV 벡터 입자들을 생산하였다. 양성 대조군으로, pCMVΔR8.2 *gag-pol* 발현 플라스미드, pHGVSV-G(서열번호: 74) *env* 발현 플라스미드, 및 pHR/lacZ 전이 벡터를 293T 세포로 공형질감염시켜 HIV-1 벡터 입자들을 생산하였다. 음성 대조군으로, 단지 pCAH/SINd1/hlacZ(서열번호: 79) 전이 벡터 플라스미드만을 형질감염시켰다. 형질감염 후 48시간에, 감염을 위해 배양 상등액을 각 형질감염된 세포로부터 회수하여 8 μg/ml 폴리브렌 존재하에 신선한 293T 세포에 적용시켰다. 48시간 후, 유전체 DNA를 각 형질도입된 세포로부터 제조한 후, 제한 효소로 절단하여 서던 블롯 분석을 수행하였다. Dig-표지된 *lacZ* 프로브로 HIV-1-계 전이 벡터에 대한 3.15kb *Bam*H I-*Kpn* I 절편, 및 CAEV-계 전이 벡터 및 음성 대조군에 대한 1.35kb *Hind* III-*Ssp* I 절편을 검출하였다. 양성 대조군에 대해서는, pCAH/SINd1/hlacZ(서열번호: 79) 전이 벡터 플라스미드의 *Hind* III-*Ssp* I DNA 절편 0.3 ng 및 3 ng이 사용되었다. 도 12에 나타난 바와 같이, 본 발명의 CAEV-계 전이 벡터는 HIV-1-계 렌티바이러스 전이 벡터의 것과 비교할만한 수준으로 통합되었다.

<318> 실시예 7: 비분열 세포로의 유전자 전이

<319> 293T 세포를 DNA 합성 저해물질인 아피디콜린(aphidicolin)으로 처리하여 6-웰 배양 플레이트에 분주한 후,

lacZ 마커 유전자를 코딩하는 CAEV 벡터 입자로 형질도입시켰다. 대조군으로서, 세포들을 *lacZ* 발현 MuLV 레트로바이러스 벡터 및 HIV-1 렌티바이러스 벡터로 나란히 감염시켰다. 감염 후 48시간에, 형질도입 효율성을 시험하기 위해, 형질 도입된 *lacZ* 유전자의 발현을 X-gal 염색법으로 측정하였다. 도 14에 나타난 바와 같이, MuLV-유래 벡터는 DNA 합성 저해물질로 처리되지 않은 세포를 효율적으로 감염시켰다. 그러나, 세포가 DNA 합성 저해물질의 처리에 의해 세포주기에서 정체되었을 때, 형질도입 효율은 현저하게 떨어졌다. 반면, CAEV-계 벡터는 분열 세포 뿐 아니라 비분열 인간 세포를 HIV-1-계 벡터의 것과 비교할만한 수준으로 효율적으로 형질도입시킬 수 있었다.

- <320> 실시예 8: 근육 세포의 생체내 형질도입
- <321> 본 실시예에서, pCAH/SINd1/hlacZ(서열번호: 79) CAEV 벡터는 생체내에서 근육세포를 형질도입시키기 위해 사용된다. 마우스(Beige strain)의 뒷다리를 4 $\mu\text{g}/\text{ml}$ 의 폴리브렌 존재하에 상기 CAEV 벡터 100 μl 로 근육내 주사하였다. 이를 후 상기 마우스들을 희생시키고, 감염된 조직을 동결 절편법(frozen section) 및 β -갈락토시데이즈 분석을 위해 준비하였다. 예상된 결과대로 pCAH/SINd1lacZ(서열번호: 79) CAEV 벡터는 생체내에서 근육 세포로 효과적으로 형질도입된다.
- <322> 상기 특정 실시양태 및 실시예를 포함하는 상세한 설명은 본 발명을 설명하기 위한 것이지만 한정하기 위한 것은 아니다. 많은 다른 변형 및 수정이 본 발명의 사상 및 범위내에서 이루어질 수 있다. 본원에서 인용된 NCBI 데이터베이스에 기탁된 서열, 특허 및 특허출원은 인용에 의해 그 전체가 상세한 설명에 도입된다.
- <323> 인용문헌
- <324> Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**(3), 403-10.
- <325> Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**(17), 3389-402.
- <326> Anderson, M. P., Rich, D. P., Gregory, R. J., Smith, A. E., and Welsh, M. J. (1991). Generation of cAMP-activated chloride currents by expression of CFTR. *Science* **251**(4994), 679-82.
- <327> Bridges, S. H., and Sarver, N. (1995). Gene therapy and immune restoration for HIV disease. *Lancet* **345**(8947), 427-32.
- <328> Burns, J. C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J. K. (1993). Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc Natl Acad Sci U S A* **90**(17), 8033-7.
- <329> Carswell, S., and Alwine, J. C. (1989). Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. *Mol Cell Biol* **9**(10), 4248-58.
- <330> Cech, T. R. (1988). Ribozymes and their medical implications. *Jama* **260**(20), 3030-4.
- <331> Chang, M. W., Barr, E., Seltzer, J., Jiang, Y. Q., Nabel, G. J., Nabel, E. G., Parmacek, M. S., and Leiden, J. M. (1995). Cytostatic gene therapy for vascular proliferative disorders with a constitutively active form of the retinoblastoma gene product. *Science* **267**(5197), 518-22.
- <332> Curran, M. A., and Nolan, G. P. (2002). Nonprimate lentiviral vectors. *Curr Top Microbiol Immunol* **261**, 75-105.
- <333> Crawford, I., Maloney, P. C., Zeitlin, P. L., Guggino, W. B., Hyde, S. C., Turley, H., Gatter, K. C., Harris, A., and Higgins, C. F. (1991). Immunocytochemical localization of the cystic fibrosis gene product CFTR. *Proc Natl Acad Sci U S A* **88**(20), 9262-6.
- <334> DuBridge, R. B., Tang, P., Hsia, H. C., Leong, P. M., Miller, J. H., and Calos, M. P. (1987). Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol Cell Biol* **7**(1), 379-87.
- <335> Erlich, H. A. (1989). Polymerase chain reaction. *J Clin Immunol* **9**(6), 437-47.
- <336> Fagerhol, M. K., and Cox, D. W. (1981). The Pi polymorphism: genetic, biochemical, and clinical

aspects of human alpha 1-antitrypsin. *Adv Hum Genet* **11**, 1-62, 371-2.

- <337> Gilbert, J. R., and Wong-Staal, F. (2001). HIV-2 and SIV vector systems. *Somat Cell Mol Genet* **26**(1-6), 83-98.
- <338> Gillam, S., and Smith, M. (1979). Site-specific mutagenesis using synthetic oligodeoxyribonucleotide primers: I. Optimum conditions and minimum oligodeoxyribonucleotide length. *Gene* **8**(1), 81-97.
- <339> Hall, C. V., Jacob, P. E., Ringold, G. M., and Lee, F. (1983). Expression and regulation of Escherichia coli lacZ gene fusions in mammalian cells. *J Mol Appl Genet* **2**(1), 101-9.
- <340> Harmache, A., Bouyac, M., Audoly, G., Hieblot, C., Peveri, P., Vigne, R., and Suzan, M. (1995). The vif gene is essential for efficient replication of caprine arthritis encephalitis virus in goat synovial membrane cells and affects the late steps of the virus replication cycle. *J Virol* **69**(6), 3247-57.
- <341> Harmache, A., Russo, P., Guiguen, F., Vitu, C., Vignoni, M., Bouyac, M., Hieblot, C., Pepin, M., Vigne, R., and Suzan, M. (1996). Requirement of caprine arthritis encephalitis virus vif gene for in vivo replication. *Virology* **224**(1), 246-55.
- <342> Helene, C. (1991). The anti-gene strategy: control of gene expression by triplex-forming-oligonucleotides. *Anticancer Drug Des* **6**(6), 569-84.
- <343> Hoffman, E. P., Brown, R. H., Jr., and Kunkel, L. M. (1987). Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51**(6), 919-28.
- <344> Horwich, A. L., Fenton, W. A., Williams, K. R., Kalousek, F., Kraus, J. P., Doolittle, R. F., Konigsberg, W., and Rosenberg, L. E. (1984). Structure and expression of a complementary DNA for the nuclear coded precursor of human mitochondrial ornithine transcarbamylase. *Science* **224**(4653), 1068-74.
- <345> Karlin, S., and Altschul, S. F. (1990). Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc Natl Acad Sci U S A* **87**(6), 2264-8.
- <346> Karlin, S., and Altschul, S. F. (1993). Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc Natl Acad Sci U S A* **90**(12), 5873-7.
- <347> Kim, S. J., Sadelain, M., Choi, K. H., Kim, H. K., Lee, J. S., and Chung, H. Y. (1997). Tetracycline-mediated suppression of gene expression with a new dicistronic retroviral vector. *Mol Cells* **7**(4), 514-20.
- <348> Kim, S. Y., Lee, J. H., Shin, H. S., Kang, H. J., and Kim, Y. S. (2002). The human elongation factor 1 alpha (EF-1 alpha) first intron highly enhances expression of foreign genes from the murine cytomegalovirus promoter. *J Biotechnol* **93**(2), 183-7.
- <349> Kordower, J. H., Emborg, M. E., Bloch, J., Ma, S. Y., Chu, Y., Leventhal, L., McBride, J., Chen, E. Y., Palfi, S., Roitberg, B. Z., Brown, W. D., Holden, J. E., Pyzalski, R., Taylor, M. D., Carvey, P., Ling, Z., Trono, D., Hantraye, P., Deglon, N., and Aebischer, P. (2000). Neurodegeneration prevented by lentiviral vector delivery of GDNF in primate models of Parkinson's disease. *Science* **290**(5492), 767-73.
- <350> Landau, N. R., and Littman, D. R. (1992). Packaging system for rapid production of murine leukemia virus vectors with variable tropism. *J Virol* **66**(8), 5110-3.
- <351> Lee, C. C., Pearlman, J. A., Chamberlain, J. S., and Caskey, C. T. (1991). Expression of recombinant dystrophin and its localization to the cell membrane. *Nature* **349**(6307), 334-6.
- <352> Lemarchand, P., Jaffe, H. A., Danel, C., Cid, M. C., Kleinman, H. K., Stratford-Perricaudet, L. D., Perricaudet, M., Pavirani, A., Lecocq, J. P., and Crystal, R. G. (1992). Adenovirus-mediated transfer of a recombinant human alpha 1-antitrypsin cDNA to human endothelial cells. *Proc Natl Acad Sci U S A* **89**(14), 6482-6.

- <353> Maher, L. J., 3rd, Wold, B., and Dervan, P. B. (1991). Oligonucleotide-directed DNA triple-helix formation: an approach to artificial repressors. *Antisense Res Dev* **1**(3), 277-81.
- <354> Marcus-Sekura, C. J. (1988). Techniques for using antisense oligodeoxyribonucleotides to study gene expression. *Anal Biochem* **172**(2), 289-95.
- <355> Miller, A. D. (1992). Human gene therapy comes of age. *Nature* **357**(6378), 455-60.
- <356> Mitrophanous, K., Yoon, S., Rohll, J., Patil, D., Wilkes, F., Kim, V., Kingsman, S., Kingsman, A., and Mazarakis, N. (1999). Stable gene transfer to the nervous system using a non-primate lentiviral vector. *Gene Ther* **6**(11), 1808-18.
- <357> Mselli-Lakhali, L., Favier, C., Da Silva Teixeira, M. F., Chettab, K., Legras, C., Ronfort, C., Verdier, G., Mornex, J. F., and Chebloune, Y. (1998). Defective RNA packaging is responsible for low transduction efficiency of CAEV-based vectors. *Arch Virol* **143**(4), 681-95.
- <358> Mullen, C. A., Kilstrup, M., and Blaese, R. M. (1992). Transfer of the bacterial gene for cytosine deaminase to mammalian cells confers lethal sensitivity to 5-fluorocytosine: a negative selection system. *Proc Natl Acad Sci U S A* **89**(1), 33-7.
- <359> Mulligan, R. C. (1993). The basic science of gene therapy. *Science* **260**(5110), 926-32.
- <360> Mullis, K. B., and Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* **155**, 335-50.
- <361> Murakami, K., Amaya, Y., Takiguchi, M., Ebina, Y., and Mori, M. (1988). Reconstitution of mitochondrial protein transport with purified ornithine carbamoyltransferase precursor expressed in *Escherichia coli*. *J Biol Chem* **263**(34), 18437-42.
- <362> Naldini, L., Blomer, U., Galloway, P., Ory, D., Mulligan, R., Gage, F. H., Verma, I. M., and Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**(5259), 263-7.
- <363> Ory, D. S., Neugeboren, B. A., and Mulligan, R. C. (1996). A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. *Proc Natl Acad Sci U S A* **93**(21), 11400-6.
- <364> Persons, D. A., Mehaffey, M. G., Kaleko, M., Nienhuis, A. W., and Vanin, E. F. (1998). An improved method for generating retroviral producer clones for vectors lacking a selectable marker gene. *Blood Cells Mol Dis* **24**(2), 167-82.
- <365> Pfarr, D. S., Rieser, L. A., Woychik, R. P., Rottman, F. M., Rosenberg, M., and Reff, M. E. (1986). Differential effects of polyadenylation regions on gene expression in mammalian cells. *DNA* **5**(2), 115-22.
- <366> Rhodes, A., and James, W. (1991). Inhibition of heterologous strains of HIV by antisense RNA. *Aids* **5**(2), 145-51.
- <367> Riordan, J. R., Rommens, J. M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J. L., and et al. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **245**(4922), 1066-73.
- <368> Roberts, S., Cheetham, J. C., and Rees, A. R. (1987). Generation of an antibody with enhanced affinity and specificity for its antigen by protein engineering. *Nature* **328**(6132), 731-4.
- <369> Rommens, J. M., Iannuzzi, M. C., Kerem, B., Drumm, M. L., Melmer, G., Dean, M., Rozmahel, R., Cole, J. L., Kennedy, D., Hidaka, N., and et al. (1989). Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* **245**(4922), 1059-65.
- <370> Saltarelli, M., Querat, G., Konings, D. A., Vigne, R., and Clements, J. E. (1990). Nucleotide sequence

and transcriptional analysis of molecular clones of CAEV which generate infectious virus. *Virology* **179**(1), 347-64.

- <371> Saltarelli, M. J., Schoborg, R., Gdovin, S. L., and Clements, J. E. (1993). The CAEV tat gene trans-activates the viral LTR and is necessary for efficient viral replication. *Virology* **197**(1), 35-44.
- <372> Saltarelli, M. J., Schoborg, R., Pavlakis, G. N., and Clements, J. E. (1994). Identification of the caprine arthritis encephalitis virus Rev protein and its cis-acting Rev-responsive element. *Virology* **199**(1), 47-55.
- <373> Sauter, S. L., and Gasmi, M. (2001). FIV vector systems. *Somat Cell Mol Genet* **26**(1-6), 99-129.
- <374> Seroude, V., Audoly, G., Gluschankof, P., and Suzan, M. (2002). Viral and cellular specificities of caprine arthritis encephalitis virus Vif protein. *Virology* **292**(1), 156-61.
- <375> Smith, T. F., Waterman, M. S., and Fitch, W. M. (1981). Comparative biosequence metrics. *J Mol Evol* **18**(1), 38-46.
- <376> Soneoka, Y., Cannon, P. M., Ramsdale, E. E., Griffiths, J. C., Romano, G., Kingsman, S. M., and Kingsman, A. J. (1995). A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res* **23**(4), 628-33.
- <377> To, R. Y., Booth, S. C., and Neiman, P. E. (1986). Inhibition of retroviral replication by anti-sense RNA. *Mol Cell Biol* **6**(12), 4758-62.
- <378> Toole, J. J., Knopf, J. L., Wozney, J. M., Sultzman, L. A., Buecker, J. L., Pittman, D. D., Kaufman, R. J., Brown, E., Shoemaker, C., Orr, E. C., and et al. (1984). Molecular cloning of a cDNA encoding human antihemophilic factor. *Nature* **312**(5992), 342-7.
- <379> von Ruden, T., and Gilboa, E. (1989). Inhibition of human T-cell leukemia virus type I replication in primary human T cells that express antisense RNA. *J Virol* **63**(2), 677-82.
- <380> Wang, C. C. (1991). A novel suicide inhibitor strategy for antiparasitic drug development. *J Cell Biochem* **45**(1), 49-53.
- <381> Warne, P. H., Viciano, P. R., and Downward, J. (1993). Direct interaction of Ras and the amino-terminal region of Raf-1 in vitro. *Nature* **364**(6435), 352-5.
- <382> Weintraub, H. M. (1990). Antisense RNA and DNA. *Sci Am* **262**(1), 40-6.
- <383> Wills, K. N., Maneval, D. C., Menzel, P., Harris, M. P., Sutjipto, S., Vaillancourt, M. T., Huang, W. M., Johnson, D. E., Anderson, S. C., Wen, S. F., and et al. (1994). Development and characterization of recombinant adenoviruses encoding human p53 for gene therapy of cancer. *Hum Gene Ther* **5**(9), 1079-88.
- <384> Wion, K. L., Kelly, D., Summerfield, J. A., Tuddenham, E. G., and Lawn, R. M. (1985). Distribution of factor VIII mRNA and antigen in human liver and other tissues. *Nature* **317**(6039), 726-9.

표 1

```

PileUp
  MSF: 9300  Type: N  Check: 9398  ..
  Name: NC_001463 (SEQ ID NO: 1) Len: 9300  Check: 3957  Weight: 0
  Name: AF322109 (SEQ ID NO: 2) Len: 9300  Check: 5441  Weight: 0
  //
  1 50
  NC_001463 .....GAGTTC TAGG...AGA GTCCTCCTA GTCTCTCCTC
  AF322109 GTGAGTGCTC TGAGGAGCTC GAAGGAAAGA GTCC.TC..A GCCTCTCCTC
  51 100
  NC_001463 TCCGAGGAGG TACCGAGACC TCAAATAAA GGAGTGATTG CCTTACTGCC
  AF322109 TCCGAGGAGC TTCGG....C TCATAATAAA GGAGTGCTTG CTTCA..ACA
  101 150
  NC_001463 GAGTGGAGAG TGATTACTGA GCGGCCGGTG TATCGGGAGT CGTCCCTTAA
  AF322109 GAACTGAG... ..CTGG TCGTGGTTAT TATCGGG... .GACCGAAGT
  151 200
  NC_001463 TCTGTGCAAT ACCAGAGCGG CTCTCGCAGC TGGCGCCCAA CGTGGGGCCC
  AF322109 CCCGTGCAAC ACCGGGGCGG TTCTCGCAGC TGGCGCCCAA CGTGGGGCTC
  201 250
  NC_001463 GAGGAG....
  AF322109 GAGTAGCTTG AGAAGCTCGA CTGAGATCTG AATCCAAGAG CGACATCAGA
  251 300
  NC_001463 ....AAGAAA AGAAAGC... GGCCCTGAGA ACTCGGCTTC TG..AAAAAG
  AF322109 CAGCAAGAAA TGAGAGTAAT GAGACCGCGA GCTCTGCTGC TGTAAAAAAG
  301 350
  NC_001463 AGGAAGAGGA CAAGTTGCTA TAGCAACAAG AGAGAAGAAG TAGAGCAAAG
  AF322109 AGGAAGTAG. CGGGTTGCCG AGGCAACTGC TCAGAAGAAC CAGGGGAAAG
  351 400
  NC_001463 GTCCAGTGGC T.CGAAAAA GAGGAACTGA AACTTCGGGG ACGCCTGAAG
  AF322109 GGCTTCCAGC AACCTCAAAA GAGGAACCGA GACTTCGGGG ACGCCTGAA.
  401 450
  NC_001463 GAGTAAGGTA AGTGACTCTG CTGTACGCGG GCGGAGGCAG AGGTT.TCCT
  AF322109 ..GTAAGGTA AGTGACTCTG CTGTACGCGG GCGGAGGCAT AGGAGATCCT
  451 500
  NC_001463 TCTAAATT.G AAAGAGAAGT GTTGCTGCGA GAGGTCTTGG TGGTCGAGAA
  AF322109 TCTATTCTAG GAAGAGAAGC GCTGTTCTGG GAGGTCTTGG CGACCGAGAA
  501 550
  NC_001463 TCCTGTACAA AAAAAAGGAG GGATCTCGGT CAGGACCAGG ACCCCTGGGA
  AF322109 TCTTGT... AAATAAGCCA GGATCTCGAT CAGGACCAAG ACCCCTCAGG
  551 600
  NC_001463 GTAATACAAC AGCAACACCG TAAGAAAATC CGCCATGGTG AGTCTAGATA
  AF322109 AGAGGGTATA GACAGCGTGG TAAGAAA.TC CGCCGTGGTG AGTCTAGATA
  601 650
  NC_001463 GAGCATGGC GAGGCAAGTC TCCGGGGGA AAAGAGATTA TCCTGAGCTC
  
```

AF322109	GAGACATGGT	GAGGCAGGCC	TCCGGAAGGG	GAAAGGAGTA	CCCCGAGCTA
	651				700
NC_001463	GAAAAATGTA	TCAAGCATGC	ATGCAAGATA	AAAGTTCGAC	TCAGAGGGGA
AF322109	AAAGAATGTC	TGAAAAAGGC	ATGCAAAATA	AAAGTAAGGG	CTGGGGGGGA
	701				750
NC_001463	GCACTTGACA	GAAGGAAATT	GTTTATGGTG	CCTTAAAAACA	TTAGATTACA
AF322109	GCGCCTGACA	GAAGGAAATT	GTCTCTGGTG	TATAAAAAACA	CTAGAGTGTA
	751				800
NC_001463	TGTTTGAGGA	CCATAAAGAG	GAACCTTGGA	CAAAAGTAAA	ATTTAGGACA
AF322109	TGTATGAGGA	TTGTAGGGAG	GAACCTTGGA	CCCCAGAAAA	ATGTAACAA
	801				850
NC_001463	ATATGGCAGA	AGGTGAAGAA	TCTAACTCCT	GAGGAGAGTA	ACAAAAAAGA
AF322109	TTATGGAAAA	ACTTGAAGCA	GGTAGAGCCT	GAGGAGAGTA	GCAAAAGCAGA
	851				900
NC_001463	CTTTATGTCT	TTGCAGGCCA	CATTAGCGGG	TCTAATGTGT	TGCCAAATGG
AF322109	CTATAACTCG	TTAAAAGCAA	CCTTGGCGGG	GATAGTCTGT	GTGCAAATGG
	901				950
NC_001463	GGATGAGACC	TGAGACATTG	CAAGATGCAA	TGGCTACAGT	AATCATGAAA
AF322109	GAATGCAGCC	CGAGACACTG	CAGGATGCGA	TAGCAACCTT	AAACATGAGA
	951				1000
NC_001463	GATCGGTTAC	TGGAACAAGA	GGAAAAGAAG	GAAGACAAAA	GAGAAAAGGA
AF322109	GA.....	TGAAGTAAA	GGAAAGGAA	.AAGCCATCA	GAAGAAAAGA
	1001				1050
NC_001463	AGAGAGTGTC	TTCCCAATAG	TAGTGCAAGC	AGCAGGAGGG	AGRAGCTGGA
AF322109	AGGGAATATA	TCCC..ATAT	TAGTGCAAGC	AGGAGGAGGA	AGAGCATGGA
	1051				1100
NC_001463	AAGCAGTAGA	TTCTGTAATG	TTCCAGCAAC	TGCAAACAGT	AGCAATGCAG
AF322109	GAGCGGTAGA	GCCTGCTACC	TTTCAGCAGC	TCCAAACAGT	GGCAATGCAG
	1101				1150
NC_001463	CATGGCCTCG	TGTCTGAGGA	CTTGAAAGG	CAGTTGGCAT	ATTATGCTAC
AF322109	CATGGACTAG	TATCAGAAGA	ATTTGAAAGG	CAGCTAGCAT	ACTATGCCAC
	1151				1200
NC_001463	TACCTGGACA	AGTAAAGACA	TACTAGAAGT	ATTGGCCATG	ATGCCCTGGAA
AF322109	CACATGGACA	AGCAAGGATA	TCTTAGAAGT	ATTAGCCATG	ATGCCAGGAA
	1201				1250
NC_001463	ATAGAGCTCA	AAAGGAGTTA	ATTCAAGGGA	AATTAATGA	AGAAGCAGAA
AF322109	ATAGAGCCCA	AAAAGAACTA	ATACAAGGAA	AGTTAAATGA	GGAAGCAGAG
	1251				1300
NC_001463	AGGTGGAGAA	GGAATAATCC	ACCACCTCCA	GCAGGAGGAG	GATTAACAGT
AF322109	AGATGGAGAA	GGCAGAATCC	ACAACCT...	...GCGGGCG	GGTTAACCGT
	1301				1350
NC_001463	GGATCAAATT	ATGGGGGTAG	GACAAACAAA	TCAAGCAGCA	GCACAAGCTA
AF322109	GCATCAGATA	ATGGGGGTAG	GACAAACGAA	TCAGGCAGCG	GCACAGGCTA
	1351				1400
NC_001463	ACATGGATCA	GGCAAGGCAA	ATATGCCTGC	AATGGGTAAT	AAATGCATTA

AF322109	ATATGGATCA	AGCAAGACAA	ATATGCCTAC	AATGGGTTAT	AACAGCAATA
	1401				1450
NC_001463	AGAGCAGTAA	GACATATGGC	GCACAGGCCA	GGGAATCCAA	TGCTAGTAAA
AF322109	AGAGGAGTTA	GGCATATGGC	CCATAGACCA	GGAAATCCCA	TGCTGGTAAG
	1451				1500
NC_001463	GCAAAAACG	AATGAGCCAT	ATGAAGATT	TGCAGCAAGA	CTGCTAGAAG
AF322109	ACAAAAACCA	AATGAGAAT	ATGAAGATT	TGCCGCAAGG	TTGTTAGAAG
	1501				1550
NC_001463	CAATAGATGC	AGAGCCAGTT	ACACAGCCTA	TAAAAGATTA	TCTAAAGCTA
AF322109	CAGTGGATGC	AGAACCCTGT	ACCCAACCTA	TAAAAGAATA	TTTAAAGGTA
	1551				1600
NC_001463	ACACTATCTT	ATACAAATGC	ATCAGCAGAT	TGTCAGAAGC	AAATGGATAG
AF322109	ACTCTGTCTT	ACACAAATGC	AAATTCGGAA	TGTCAAAAAC	ATATGGACAG
	1601				1650
NC_001463	AACACTAGGA	CAAAGAGTAC	AACAAGCTAG	TGTAGAAGAA	AAAAATGCAAG
AF322109	AGTGTGGGG	CAAAGAGTAC	AGCAGGCCTC	AATAGAAGAA	AAAAATGCAAG
	1651				1700
NC_001463	CATGTAGAGA	TGTGGGATCA	GAAGGGTTCA	AAATGCAATT	GTTAGCACAA
AF322109	CATGCAGGGA	CATCGGGGGA	ACAGCATATC	AGATGCAGTT	GCTTGCACAA
	1701				1750
NC_001463	GCATTAAGGC	CAGGAAAAGG	AAAAGGGAAT	GGACAGCCAC	AAAGGTGTTA
AF322109	GCCCTCCGTG	GCGGAAAAGA	AGATGGGAAA	AAATCTGTAG	GGAAGTGTTA
	1751				1800
NC_001463	CAACTGTGGA	AAACCCGGAC	ATCAAGCAAG	GCAATGTAGA	CAAGGAATCA
AF322109	TAACTGTGGA	AGGCCCGGAC	ACAGAGCAAA	AGAATGCAGA	CAAGGCATTA
	1801				1850
NC_001463	TATGTCACAA	CTGTGGAAAG	AGAGGACATA	TGCAAAAAGA	ATGCAGAGGA
AF322109	TATGTCACAA	CTGTGGAAAA	AGAGGGCATA	TACAGAAAAA	CTGCA...A
	1851				1900
NC_001463	AAGAGAGACA	TAAGGGGAAA	ACAGCAGGGA	AACGGGAGGA	GGGGGATACG
AF322109	AC.AGAA...	AAGAAGAAA	GGAGCAGGGA	AACATGAGGA	GGGGGCTACG
	1901				1950
NC_001463	TGTGGTGCCG	TCCGCTCCTC	CTATGGAATA	ACTTCAGCAC	CACCTATGGT
AF322109	TGTGGTGCCG	TCCGCACCCC	CTATGGAGTA	ACGCAAGCAC	CACTAATAGT
	1951				2000
NC_001463	TCAGGTCCGC	ATAGGTTCCC	AGCAGAGGAA	CTTGTATT	GATACCGGGG
AF322109	TAGGTACAA	ATAGGGAATC	AGGAGAAAACA	ATTATTATT	GACACAGGGG
	2001				2050
NC_001463	CGGACCGAAC	TATAGTTAGA	TGGCATGAGG	GCTCGGGAAA	CCCAGCCGGA
AF322109	CAGATAAAC	GATAGTAAGA	ATGCATGATG	GAACAGGGAT	TCCAAACCGA
	2051				2100
NC_001463	AGGATAAAC	TGCAAGGAAT	AGGAGGAATA	GTAGAAGGAG	AAAAATGGAA
AF322109	AGAATAAAT	TACAAGGGAT	AGGAGGAATA	GTAGAAGGAG	AAAAATGGAA
	2101				2150
NC_001463	TAATGTAGAA	TTAGAATATA	AAGGAGAAC	AAGAAAGGGA	ACAATAGTAG

AF322109	TAAAGTACCC	ATGACATATA	AGGGACAAAC	ATCCTGCCCA	AGCTTGCCTG
	2151				2200
NC_001463	TGTTACCACA	AAGTCCAGTA	GAAGTATTAG	GACGAGATAA	CATGGCCCCG
AF322109	TGCTAAGAGA	TAGCCCAGTA	GAAGTATTGG	GAAGAGATAA	CATGGAAGCA
	2201				2250
NC_001463	TTTGGAAATA	AGATAATAAT	GGCAAATTTA	GAGGAAAAAA	GAATCCCAAT
AF322109	TTCCGCGTAA	CCCTAATAAT	GGCAAATTTA	GAAGATAAGA	AAATTCAC
	2251				2300
NC_001463	TACAAAAGTA	AAATTGAAAG	AGGGATGTAC	GGGTCCACAT	GTCCCAACAT
AF322109	AATACCAGTA	GAATTGAAAG	AAGGATGTAA	AGGGCCACAT	GTGCCCCAGT
	2301				2350
NC_001463	GGCCATTAAC	AGAAGAGAAA	TTAAAAGGTC	TAACAGAAAT	CATAGATAAA
AF322109	GGCCATTAAC	AGCAGAGAAA	TTACAAGGAC	TAACAGGAAT	AGTAGAAAAA
	2351				2400
NC_001463	TTAGTGGAAAG	AAGGAAAAC	AGGAAAGGCA	CCCCACATT	GGACATGTAA
AF322109	TTACTACAGG	AAGGAAAATT	GGCAGAGGCC	CCAGAGGGAT	GGACGTGGAA
	2401				2450
NC_001463	TACTCCAATC	TTTTGCATAA	AAAAGAAATC	AGGGAAGTGG	AGAATGTTAA
AF322109	CACGCCCATC	TTCTGCATAA	AAAAGAAGTC	AGGAAAATGG	AGAATGTTAA
	2451				2500
NC_001463	TAGATTTTAC	AGAATTGAAC	AAACAGACAG	AAGATTTAAC	AGAAGCCAG
AF322109	TAGATTTTAC	GGAATTTAAT	AAGCAACAG	CAGATTTAGC	AGAAGCCAG
	2501				2550
NC_001463	TTAGGACTCC	CGCATCCGGG	AGGACTACAA	AAGAAAAAAC	ATGTTACAAT
AF322109	CTAGGACTGC	CACACCCAGG	AGGTTGCAA	AGGAAAAAGA	ATGTAACAAT
	2551				2600
NC_001463	ATTGGACATA	GGAGATGCAT	ATTTTACTAT	ACCCCTATAT	GAACCATATC
AF322109	TCTGGACATA	GGAGATGCAT	ATTTTACAAT	TCCCTTATAC	GAGCCCTATC
	2601				2650
NC_001463	GAGAGTACAC	ATGTTTTACT	CTATTAAGTC	CTAATAATCT	AGGACCATGT
AF322109	AGAAATATAC	ATGCTTCACA	CTCCTAAGTC	CTAACAATTT	GGGACCATGT
	2651				2700
NC_001463	AAAAGATACT	ATTGGAAAGT	GCTGCCACAA	GGTTGGAAAT	TGAGTCCATC
AF322109	AAAAGGTATT	ATTGGAAAGT	ATTACCCAG	GGATGGAAAT	TGAGCCCAGC
	2701				2750
NC_001463	TGTATATCAA	TTTACTATGC	AGGAGATCTT	AGAGGATTGG	ATACAGCAGC
AF322109	TGTATATCAA	TTTACCATGC	AAAGTTTCTT	AAAAGGATGG	ATACAACAGC
	2751				2800
NC_001463	ATCCAGAAAT	TCAATTTGGC	ATATATATGG	ATGATATTTA	CATAGGAAGT
AF322109	ATAAAAACAT	ACAATTTGGA	ATATATATGG	ATGATATCTA	TATTGGAAGT
	2801				2850
NC_001463	GATTTAGAAA	TTAAAAAGCA	TAGAGAAATA	GTGAAAGATT	TAGCCAATTA
AF322109	GATCTAACGA	TAGCCCAACA	TAGGAAGATA	ATAGAAGAAT	TAGCCTCATT
	2851				2900
NC_001463	TATTGCCCAA	TATGGATTCA	CTCTGCCAGA	AGAGAAGAGA	CAAAAGGGAT

AF322109	TATAGAACAA	TTGGGTTTA	CATTACCAGA	AGATAAGAGA	CAAGAGGGCT	
		2901				2950
NC_001463	ATCCAGCAAA	ATGGCTAGGA	TTTGAACCTAC	ACCCGCAGAC	CTGGAAATTT	
AF322109	ATCCAGCAAA	ATGGCTAGGA	TTCGAGCTAC	ATCCAGAAAA	ATGGAAATAT	
		2951				3000
NC_001463	CAGAAGCATA	CATTACCTGA	ATTAACAAAG	GGAACAATAA	CATTAATAAA	
AF322109	CAAAAGCATA	AATTGCCGGA	ATTACAAGAG	GGGTAATAA	CCCTGAACAA	
		3001				3050
NC_001463	ATTACAGAAA	TTAGTAGGAG	AATTAGTATG	GAGACAATCC	ATAATTGGGA	
AF322109	ATTACAGAAG	ATAGTAGGGG	AATTAGTGTG	GAGACAATCC	TTGATAGGAA	
		3051				3100
NC_001463	AAAGCATTC	TAACATTCTG	AAATTAATGG	AAGGAGATAG	AGAATTACAA	
AF322109	AGAGCATCCC	CAATATCATA	AAATTAATGG	AAGGAGATCG	CGCATTACAA	
		3101				3150
NC_001463	AGTGAAAGAA	AAATGGAQA	AGTACATGTG	AAAGAATGGG	AAGCATGTAG	
AF322109	AGTGAAAGGA	AAATAGAAAG	AATACATGTA	CAAGAATGGG	AAGCATGTCA	
		3151				3200
NC_001463	GAAGAAATTA	GAAGAAATGG	AAGGAAATTA	TTATAATAAA	GACAAAGATG	
AF322109	AAAGAAATTA	GATGAAATGG	TAGGAAATTA	TTACAGAGAA	GAAGAAGATA	
		3201				3250
NC_001463	TCTATGGACA	ATTGGCTTGG	GGAGACAAAG	CTATAGAATA	TATAGTGTAT	
AF322109	TCTATGGACA	AATAACTTGG	GGGATAAAGG	CAATAAATA	CATAGTATTC	
		3251				3300
NC_001463	CAGGAGAAAG	GGAAACCATT	ATGGGTAAAT	GTGGTTCACA	ATATAAAGAA	
AF322109	CAAAGGAAAG	GGGAACCCTT	ATGGGTAAAT	GTAGTACATG	ACATAAAAAA	
		3301				3350
NC_001463	CCTAAGCATC	CCGCAACAGG	TTATTAAAGC	AGCGCAAAAA	TTAACCCTAAG	
AF322109	TTTGAGTCTC	CCACAGCAAG	TGATAAAAGC	AGCACAGAAA	TTAACCCTAAG	
		3351				3400
NC_001463	AAGTCATCAT	TAGGACAGGA	AAAATACCAT	GGATATTGTT	GCCAGGGAAA	
AF322109	AAGTAATCAT	AAGAACAGGA	AAAATCCCAT	GGCTGCTACT	ACCAGGAAGA	
		3401				3450
NC_001463	GAAGAAGATT	GGAGACTAGA	ATTGCAATTA	GGGAACATCA	CATGGATGCC	
AF322109	GAAGAAGACT	GGAGATTAGA	ACTGCAGGTA	GGGAACATCA	CGTGGATGCC	
		3451				3500
NC_001463	AAAATTTTGG	TCCTGTTATC	GAGGA.CATA	CAAGATGCAG	AAAAGAAAT	
AF322109	ATCATTTTGG	TCATGTTATC	GAGGAGCACC	CAAG.TGGAA	AAGAAGGAAC	
		3501				3550
NC_001463	ATAATAGAAG	AAGTAGTAGA	AGGGCCTACA	TATTATACAG	ATGGAGGAAA	
AF322109	ATAGTGGCAG	CAGTGGTAGA	TGGACCGACA	TATTATACAG	ATGGGGGAAA	
		3551				3600
NC_001463	AAAGAATAAA	GTAGGAAGTC	TAGGGTTCAT	AGTATCAACA	GGGAAAAAAT	
AF322109	GAAGAAACGCA	CAGGGAAGCT	TTGGCTTCAT	CTCCCAACA	GGAGAAAAGT	
		3601				3650
NC_001463	TTAGAAAGCA	TGAAGAGGGC	ACAAACCAGC	AACCTAGAATT	AAGAGCCATA	

AF322109	TCAGAAGGCA	TGAAGATGGA	ACTAATCAGG	TATTAGAATT	AAGGGCAATA
	3651				3700
NC_001463	GAGGAAGCTC	TAAAACAAGG	GCCTCAAACA	ATGAATTTAG	TAACAGATAG
AF322109	GAAGATCCAT	GTAACAAGG	ACCTGAAAGC	ATGAACATTG	TAACATGACAG
	3701				3750
NC_001463	TAGATATGCA	TTTGAATTTT	TATTAAGAAA	TTGGGATGAA	GAAGTAATAA
AF322109	CAGGTATGCT	TATGAATTC	TGCTCCGAAA	CTGGGATGAA	CAGGTCATAA
	3751				3800
NC_001463	AGAAATCCAAT	TCAAGCAAGA	ATTATGGAAA	TTGCCACAA	GAAAGATAGG
AF322109	GAAACCCCAT	TCAGGCAAGA	ATCATGGCAG	AAGTGCACAA	GAAAAAGCAG
	3801				3850
NC_001463	ATAGGAGTGC	ATTGGGTGCC	AGGACATAAA	GGGATCCCC	AAAATGAAGA
AF322109	GTAGGAATAC	ACTGGGTGCC	AGGCATAAA	GGAATACCTC	ACAATGAAGA
	3851				3900
NC_001463	AATAGACAAA	TATATTTCCG	AAATATTTCT	TGCAAAAGAA	GGAGAAGGAA
AF322109	GATAGACCAG	TACATATCAG	AAGTATTTCT	AGCACGAGAA	GGAACAGGGA
	3901				3950
NC_001463	TTCTCCAAA	AAGAGAAGAG	GATGCAGGGT	ATGATTTAAT	ATGCCAGAA
AF322109	TATGTGAAA	AAGGAAGGAA	GATGCTGGAT	ATGATTTATT	ATGCCGCAT
	3951				4000
NC_001463	GAGGTTACCA	TAGAGCCAG	ACAAGTAAA	TGCATCCCA	TAGAGCTAAG
AF322109	GAGGTAATAC	TTAAACCCA	AGAAGTAAA	CGGATCCCA	TAGACCTAAA
	4001				4050
NC_001463	ATTAATTTA	AAGAAATCAC	AATGGGCTAT	GATTGCTACA	AAAAGCAGCA
AF322109	ATTAATTTG	AAGAAAAGC	AATGGGCCAT	GATAAGTGGG	AAAAGTAGCG
	4051				4100
NC_001463	TGGCTGCCAA	AGGAGTGTTT	ACACAAGGAG	GAATCATAGA	CTCAGGATAT
AF322109	TTGCAGCAA	AGGAATATTT	GTACAAGGAG	GCATAATAGA	TTCAGGGTAT
	4101				4150
NC_001463	CAGGGACAAA	TACAGGTAAT	AATGTATAAT	AGCAATAAAA	TAGCAGTAGT
AF322109	CAGGGACAAG	TACAAGTCAT	CCTATATAAT	AGTAATAAGA	TAGAGGTCAA
	4151				4200
NC_001463	CATACCCCAA	GGGAGAAAAT	TTGCACAATT	AATATTAATG	GATAAAAAGC
AF322109	AATACCACAA	GGCAGGAAAT	TTGCCCAATT	AATATTAATG	AACTTACAAC
	4201				4250
NC_001463	ATGGAAAATT	GGAACCTTGG	GGGAAAGCA	GAAAAACAGA	AAGGGGAGAA
AF322109	ATGAAGAATT	AGAAGAATGG	GGAAAGGAAA	GAAAAACAGA	AAGAGGAACA
	4251				4300
NC_001463	AAAGGATTTG	GGTCTACAGG	AATGTATTGG	ATAGAAAATA	TTCTCTGGC
AF322109	AAAGGATTTG	GGTCTACAGG	AGCATTTTGG	ATAGAGAATA	TTCCCAAGC
	4301				4350
NC_001463	AGAGGAAGAC	CACACAAAAT	GGCATCAAGA	TGCCCGATCA	TTGCATCTAG
AF322109	AGAGGAAGAA	CATTACAAAT	GGCATCAAGA	TGCTAGATCT	CTGCAGCTAG
	4351				4400
NC_001463	AATTTGAAAT	TCCAAGACA	GCAGCAGAAG	ACATAGTAAA	TCAATGTGAA

AF322109	AATTCAAGAT	ACCTAGAGCA	GCAGCAGAAG	ACATTATACA	GCACTGTGAG	
	4401					4450
NC_001463	ATATGCAAAG	AAGCGAGGAC	ACCTGCAGTA	ATTAGAGGCG	GAACAAAAG	
AF322109	GTATGTCAAG	AAGGCAAACC	CGCAGCGATC	ACGAGAGGGG	GAATAAAAAG	
	4451					4500
NC_001463	GGGGGTAAAT	CATTGGCAAG	TGGATTATAC	CCATTATGAA	AATATCATA	
AF322109	AGGAATAGAT	CATTGGCAGG	TAGACTATAC	ACATTACAAA	GAACACATAA	
	4501					4550
NC_001463	TATTAGTATG	GGTAGAAA	AATTCAGGAC	TAATATATGC	AGAAAAAGTA	
AF322109	TATTAGTATG	GGTAGAGACT	AATTCAGGAT	TAATATTTGC	AGAGAAAGTA	
	4551					4600
NC_001463	AAAGGAGAAT	CAGGGCAAGA	ATTCAGAATA	AAAGTGATGC	ATTGGTATGC	
AF322109	AAAGGAGAAT	CAGGACAAGA	ATTTAGGATG	CAGACATTGA	AATGGTATGC	
	4601					4650
NC_001463	ATTATTTGGT	CCAGAGTCAT	TGCAGTCAGA	CAATGGACCT	GCATTTGCAG	
AF322109	TTTGTTCCT	CCAAAATCAG	TGCAATCAGA	TAATGGGACA	GCCTTCACAG	
	4651					4700
NC_001463	CAGAGCCAC	ACAGCTGTTA	ATGCAATACC	TAGGAGTAAA	ACACACAACA	
AF322109	CTGAGGCTAC	GCAGCATCTA	ATGAAGTATT	TAGGGATTCA	GCACACTACG	
	4701					4750
NC_001463	GGCATACTT	GGAATCCACA	GTCTCAGGCT	ATAGTAGAAA	GGGCACATCA	
AF322109	GGTATTCCGT	GGAACCCCA	GTCACAAAGT	TTAGTAGAAA	GAGCTCATCA	
	4751					4800
NC_001463	ACTATTGAAA	AGCACTTTAA	AGAAGTTCCA	GCCACAATTT	GTCGCTGTAG	
AF322109	AACATTAAAA	CACATGTTAG	AAAAATTAGA	ACCACAATTT	GTGGCCCTAC	
	4801					4850
NC_001463	AATCAGCCAT	AGCAGCAGCC	CTAGTCGCCA	TAAATATAAA	AAGAAAGGGT	
AF322109	AGTCTGCCAT	CGCAGCCACT	CTAGTTGCCG	TCAATATAAA	AAGAAAGGGT	
	4851					4900
NC_001463	GGGCTGGGGA	CAAGCCCTAT	GGATATTTT	ATATATAATA	AAGAACAGAA	
AF322109	GGACTAGGGG	CAAGCCCTAT	GGATATTTAC	ATATATAATA	AGGAGCAACA	
	4901					4950
NC_001463	AAGAATAAAT	AATAAATATA	ATAAAAATTC	TCAAAAAATT	CAATTCTGTT	
AF322109	AAGACAACAA	GATAATAGTA	ATAAATTAAT	TCAGAAAA..	.AATTTTGT	
	4951					5000
NC_001463	ATTACAGAAT	AAGGAAAAGA	GGACATC.AG	GAGAGTGAA	AGGACCAACC	
AF322109	ATTACAGGAT	CAGAAAAGA	GGCCATCCAG	GAGAGTGAA	CGGCCCAACT	
	5001					5050
NC_001463	CAGGTACTGT	GAAAGGGGA	AGGAGCCAAT	TGTGGTAAAG	GATATAGAAA	
AF322109	GAGGTACTGT	GGGAGGGGA	AGGAGCCA.T	AGTAGTTAAA	GACAAAGAAA	
	5051					5100
NC_001463	GTGAAAAGTA	TTTAGTAATA	CCTTACAAAG	ATGCAAAATT	CATCCCGCCA	
AF322109	GTGATAGATA	TCTAGTCATC	CCATATAAAG	ATGCAAAATT	TATFCCGCCA	
	5101					5150
NC_001463	CCAACAAAAG	AAAAGGAATA	AAAAACCTGG	ACCAGAATTA	CCCTTAGCAC	

AF322109	CCGTCGGAAC	AGAAGGGAT	GAGGAATAGG	TCCAGAAATTG	CCTTTATCTT
	5151				5200
NC_001463	TATGGATACA	TATAGCAGAA	AGCATTAAATG	GGGATAGCTC	ATGGTACATA
AF322109	TATGGACTTA	TACAGCATAAC	AGCATAAATA	AAGATCCCGC	ATGGTATACA
	5201				5250
NC_001463	ACAATGAGAC	TGCAACAGAT	GATGTGGGGA	AAAAGAGGAA	ATAAGTTACA
AF322109	ACCCTAAGAC	TGCAGCAAAT	GATGTGGCAT	AGGAGGGGAA	ATAAATTGAC
	5251				5300
NC_001463	ATATAAGAAT	GAAGACAGGG	AATATGAAAA	TTGGGAAATT	ACATCATGGG
AF322109	ATATGTCAGG	GAAAAATGCAC	AGTACGAGGA	GTGGGAAATG	ACCTCGTATG
	5301				5350
NC_001463	GATGGAAAAT	GCACCTAAGG	AGAGTGAAAC	AATGGATACA	AGACAACAGG
AF322109	AGTGGAGGAT	AAGAATGAGA	AGGGACAAAA	CAAAAAGTCA	TC.CAAGAGG
	5351				5400
NC_001463	AGAGGAAGC	CCATGGCAGT	ACAAAGTAGG	AGGAACATGG	AAAAGTATAG
AF322109	GCATACTTCG	CCATGGCAAT	ATCGGAGACA	GGATGGATGG	AAGGATGTGG
	5401				5450
NC_001463	GAGTGTGGTT	CCTGCAAGCA	GGAGATTACA	GAAAGGTAGA	CAGGCAC TTC
AF322109	GAACGTGGTT	CCTACAGCCA	GGGACTATA	GAAAGCCGGA	TCAGCAGTTC
	5451				5500
NC_001463	TGGTGGGCAT	GGAGGATACT	GATATGTTCC	TGCAGGAAAG	AAAAGTTTGA
AF322109	TGTTTCGCTT	GGAGAATAGT	GTCGTGTTCA	TGTAAAAAGG	AAGGATTTAA
	5501				5550
NC_001463	TATAAGAGAA	TTTATGAGAG	GAAGACATAG	ATGGGATTTG	TGCAAACTCT
AF322109	CATAAGAGAA	TTTATGCTAG	GTACCCATAG	ATGGGATTTG	TGTAAGTCGT
	5551				5600
NC_001463	GTGCTCAAGG	AGAAGTAGTA	AAGCATACTA	GAACAAAAG	TCTGGAAAGA
AF322109	FTTGCCAGGG	TGAAGTAGTA	AAGAGAACAC	AACCCATACAC	CTTGCAAAGG
	5601				5650
NC_001463	CTAGTACTGC	TACAGATGGT	AGAACAGCAT	GTGTTTCAAG	TATTGCCATT
AF322109	CTCACGTGGC	TTAAATTAAC	AGAAGACCAT	GTATTTCAAG	TAATGCCCTT
	5651				5700
NC_001463	GTGGAGAGCC	AGGAGAAGTA	GTACAACAGA	TTTCCCATGG	TGCAGGGACA
AF322109	GTGGAGAGCT	CGCAAAGGGA	TTACCATAGA	CTTTCCCTGG	TGCAGGGACA
	5701				5750
NC_001463	CAACGGGATA	CACGCATGCG	TGGTCTGTCC	AGGAGTGCTG	GTTGATGGAA
AF322109	CAAAAGGATT	CCTGGAGCCG	TGGACAACCG	AAGAGTCTTG	GCAAATAGAG
	5751				5800
NC_001463	TATCTCTTAG	AGGATGAGTG	AAGAACTGCC	TCAAAGAAGG	GAGACACATC
AF322109	TATCCCTTGG	AGGATGAGTG	AGGAAACCCC	AGCAGGAAGA	GAACCGACTG
	5801				5850
NC_001463	CAGAAGAACT	.TGTAAGGAA	CGTACGGGAA	AGAGAAAGGG	ATACATGGCA
AF322109	CAGAGGAAAT	ATTTGAGCAAGAA	GCAGAAAGT.TGGAA
	5851				5900
NC_001463	ATGGACAAGC	ATCAGAGTAC	CTGCGGAAAT	ACTGCAAAGA	TGGCTTGCTA

AF322109	GAGAACAAGC	GTGGAGTCC	CAAATGACAT	ATTACAAAGA	TGGCTAGCAA	
	5901					5950
NC_001463	TGCTTAGGTC	AGGCAGAAAT	AGAAAGAAAG	TGTATAGAGA	AATGCAAAAA	
AF322109	TGCTTAGGCA	AAGAGGAAAT	AGAAAGAAAG	TGCTTAGGGA	AATGCAAAAA	
	5951					6000
NC_001463	TGGATGTGGA	TACATCCCAA	GGCGCCTGTG	ATTAGGCCT	GTGGATGCAG	
AF322109	TGGCATGGA	GGAATCCAC	GGCGCGGTG	ATTGGCCGT	GTGGATGTCG	
	6001					6050
NC_001463	ACTATGTAAC	CCGGGTGGG	GAACATAATC	AAGGGAATAA	TAAATGCAA	
AF322109	GCTATGTAAC	CCCGGTGGG	GGAG.TAATT	AAT..CATAA	TAAA.GCAA	
	6051					6100
NC_001463	TAAATGTAAC	TAACAAGTAG	CAAAAGTGTC	TGTGTTAGAT	GGATGCTGGG	
AF322109	T...TGTAACATGCTGTG	
	6101					6150
NC_001463	GCCAGATACA	TGCGCTTAAC	TGGGAAGGAA	AACTGGGTG	AAGTAACCAT	
AF322109	TC.....A..GG.....	...TGTCTTG	CAGGAA...T	
	6151					6200
NC_001463	GGACGGAGAG	AAGGAAAGGA	AAAGAGAAGG	TTTCACTCGG	GGACAGCAAG	
AF322109	GG.CGGAGAT	AAGAAAAG..	AA.GCAAAGG	AGCCACT...	..AATCCAGG	
	6201					6250
NC_001463	GTAAGTATCA	ACCCAGGTA	AGTAAGCAAA	TAGGGAACAG	AAATACTAAC	
AF322109	GTAAGTATAA	AAAACAGGTA	AGTA.....	...G.....	AA....TAAC	
	6251					6300
NC_001463	CCATGCTTTG	CCTATAAAGG	GATATTCCTA	TGGAGGATAT	CACTAACAAAT	
AF322109TATAGT	TATATT....A...	..CTAACAGT	
	6301					6350
NC_001463	GTGATATTG	CTAGGGATAA	ATATGTGTGT	CAGTGCAGAG	GATTACATAA	
AF322109	AAGAGCAGCA	CTAGG.....	..A.....	...GCAGAA	...TACATAA	
	6351					6400
NC_001463	CACTAATATC	AGATCCCTAT	GGTTCTCAC	CCATAAAAAA	TGTGTCTGGG	
AF322109	CCATAATATC	AGACCCATAT	GGTTCTCTC	CCGTGAGAAA	TGTGTCTAGGA	
	6401					6450
NC_001463	GTACCAGTGA	CTTGTGTAAC	AAAAGAATTC	GCAAAATGGG	GATGTCAACC	
AF322109	GTACCTGTAA	CTTGTGTGAC	AAAAGAATTT	AGTAAGTGGG	GATGTCAAGC	
	6451					6500
NC_001463	ACTAGGAGCG	TACCTGATC	CAGAAATAGA	ATACAGAAAT	GTGAGTCAGG	
AF322109	AATAGGAGCC	TACCCAGACC	CAGACTTAGA	ATACAGAAAT	ATAAGTAAAG	
	6501					6550
NC_001463	AAGTAGTGAA	AGAAGTATAT	CAAGAGAATT	GGCCATGGAA	TACATATCAT	
AF322109	AAATATTAGA	GGAAGTATAT	CAACAAGACT	GGCCGTGGAA	TACTTATCAT	
	6551					6600
NC_001463	TGGCTCTCT	GGCAAATGGA	GAATGTTAGG	TACTGGTTAA	AAGAAAATAT	
AF322109	TGGCCATTAT	GGCAAATGGA	TAATGTAGTA	CAATGGGCAA	GGCAAATTT	
	6601					6650
NC_001463	GCAAGAAAT	CAACAGAGAA	AAAATAATAC	AAAAGAGGCT	ATAGAGGAAT	

AF322109	ACAGGATAAC CGCAAG.CAA AAAAG..... ..GGAC CTGGCAGACC
	6651 6700
NC_001463	TATTAGCAGG AACTATAAGG GGAAGATTCT GTGTACCATA CCCATTTGCC
AF322109	TATTAGCAGG AAAATAAGG GGAAGATTCT GTGTACCCTA CCCATTTGCG
	6701 6750
NC_001463	TTGTAAAAAT GCACAAAGTG GTGCTGGTAT ACAGCGGCCA TAAA..CAAC
AF322109	CTCCTGGAGT GCATGGAATG GTGCTGGTGG GTTAAGAACA CTAATGCAGG
	6751 6800
NC_001463	GAGTCA.GGA AAAGCAGGAA AAATAAAAAT AAATTGCACA GAAGCAAGAG
AF322109	GGGGTATGGA GAAGCAG..A .CATAAGAAT AAATTGCTCA AGGGCAAGAG
	6801 6850
NC_001463	CAGTCTCCTG TACAGAGGAC ATGCCATTAG CCTCAATACA AAGAGCATAT
AF322109	CAGTGAGCTG CACAAGTGAA ATGCCCTTAG CATCCCTACA GAGGGTATAT
	6851 6900
NC_001463	TGGGATGAGA AAGACAGAGA GAGCATGGCC TTTATGAATA TCAAAGCATG
AF322109	TGGGAAAAGG AGGAACGAAA AAACATGGAG AAAATGACCA TCAAACCTTG
	6901 6950
NC_001463	TGATAGCAAC CTAAGGTGTC AGAAAAGACC TGGAGGGTGT ATGGAAGGAT
AF322109	CAATAAAAAT TTGGAATGCA AGAACAGAA..G.GGGATGC GCAGAAGGGT
	6951 7000
NC_001463	ACCTATCCC AGTAGGACCA GAAATAATCC CTGAAAGTAT GAAATACCTA
AF322109	ATCCAGTACC TCCCAAGGCA GAGTTATTC CTCACGCGTT TCAGGATTTA
	7001 7050
NC_001463	AGGGGAGCAA AGAGTCAG.. TATGGGGGAA TAAAGATAA GAATGGAGAA
AF322109	CAGCCA..AA AGGTACGCA TATGGGCAC TTAGAG...G GAACAGCAAA
	7051 7100
NC_001463	TTAAAATTAC CATTAACTT AAGAGTGTGG GTAAAATTAG CAAATGTGTC
AF322109	TTTCCACAAA GAGTGTGCT AAGAACATGG GTGAAAATAG CTAACCTGAC
	7101 7150
NC_001463	AGAATGGGTA AATGGGACAC CCCCGGATTG GCAAGACAGA ATTAACGGAT
AF322109	AGGATGGGAA AAAGGAAAGC CAGCAGAATG GT.....GG AATACCAG..
	7151 7200
NC_001463	CCAAAGGAAT AAATGGGACG CTCTGGGGAG AGCTTAACAG TATGCATCAC
AF322109	CCAACAGGTT CATTGGTTTG ATACCACGCC ACAATATCAT TTAGGAT...
	7201 7250
NC_001463	CTAGGATTG CCCTTAGCCA GAACGGCAA TGGTGTAAC ACACCGGGGA
AF322109	.ATGTATTAT CCCGAGCGCC TGAGAACAGG AGTTGTAATT TCACAGGGGA
	7251 7300
NC_001463	AATAAAATTA GGGCAAGAAA CATTCCAATA TCATTACAAG CCAAACCTGGA
AF322109	AATACGAATA GGGCAACATC AGTTTGAGTA TAATTACACC CTGACAAAGA
	7301 7350
NC_001463	ACTGTACC.. .GGGAATTGG ACGCAATATC CGGTGTGGCA AGTGATTAGA
AF322109	ATTGCACAAA GGAGAAGTGG AAAGAGTACC CCATGTGGCA TGTCTGGAGG
	7351 7400
NC_001463	AACCTGGATA TGGTGAACA TATGACAGGA GAATGTGTGC AGAGACCACA

AF322109	CATTTAGATC	AAAATGACCA	CTTATCTAGC	ATATCTTTCA	AAAGACCGAG	
	7401					7450
NC_001463	AAGGCACAAT	ATAACAGTAG	GAAATGGAAC	CATAACAGGG	AATTGCAGTA	
AF322109	AAGAAATGCA	ACACAAATAG	GGAACAGTAC	ACTGCAAGGG	CAATGTAATA	
	7451					7500
NC_001463	CAACAAACTG	GGATGGATGT	AATTGCTCAC	GATCAGGAAA	CTACCTATAT	
AF322109	GAGTAATTTG	GACAGGATGC	CACTGCAATG	AGACAGGGAT	AAAC..AC..	
	7501					7550
NC_001463	AACAGCTCTG	AGGGAGGATT	GTTATTAATT	CTGTGCAGAC	AAAACAGCAC	
AF322109	AACA.....TGGAGAA	TAAATGGCAC	
	7551					7600
NC_001463	CCTAACAAAG	ATCCTGGGAA	CAAATACAAA	TGGACAACCT	ATGTGGGGAA	
AF322109AAAGGG	AGC.TT..AT	CTCTTA..AA	TAGCACTAATGGAAA	
	7601					7650
NC_001463	TATACAAAAA	TTGTTTCAGGA	TGCGAGAATG	CAACATTAGA	CAACACAGGA	
AF322109	CATCATGGTC	TTGTT....A	TGCTGGAACA	CAACAGTGG.CAGGG	
	7651					7700
NC_001463	GAAGGAACCT	TAGGAGGTGT	AGCTAATAAG	AACTGTAGTC	TGCCTCATAA	
AF322109	GTA.....	TATGAGAGTC	AGCTAA....	A.GTGGAAATG	AGAGTCTTAA	
	7701					7750
NC_001463	AAATGAGAGC	AACAAGTGGA	CTTGTGCCCC	AAGACAAAGA	GATGGAAAAA	
AF322109	AGACGGAGAC	TATGGGCTCT	GTTTTAATTC	AACAAACAGG	AATTGTACTA	
	7751					7800
NC_001463	CAGATTC.GC	TATACATAGC	AGGAGGAAAA	AAGTTTTGGA	CACGAATTAA	
AF322109	GAAATGGAGC	TCGGCACTAT	GTAACAAGA	GAGTGATAAA	AAACGAC.AC	
	7801					7850
NC_001463	GGCCCCAATTC	AGCTGTGAAA	GTAACATAGG	ACAATTAGAT	GGAAATGTTGC	
AF322109	AGCAGATCAT	AATTGTGATA	GCAGCATATC	AGCAATAGAT	GGAAATGGTAC	
	7851					7900
NC_001463	ATCAGCAAAT	ACTATTGCAA	AAATATCAAG	TAATTAAGGT	AAGAGCTTAT	
AF322109	ATCAACAAAT	ATTACTGCAA	AGGTATCAAG	TAATTAGAGT	AAGAGCTTAC	
	7901					7950
NC_001463	ACATATGGGG	TGATAGAAAT	GCCAGAAAAC	TATGCAAAAA	CAAGAATCAT	
AF322109	ACATACGGAG	TGATTGATAT	GCCAGACAAT	TATG.AGACC	CTACCAGGA.	
	7951					8000
NC_001463	AAACAGGAAA	AAAAGAGAAC	TCAGCCACAA	GAGGAAGAAG	AGAGGCGTTG	
AF322109AGGAGA	AGGAGAGATC	TCGCAAAGGC	CAGGAAAAAG	ACGGGCGTGG	
	8001					8050
NC_001463	GCTTGGTCAT	TATGCTAGTT	ATCATGGCAA	TAGTAGCTGC	CGCAGGGGCT	
AF322109	GCCTGGTCAT	CATGTTAGCT	ATCATGGCCA	TAGTGGCTGC	TGCAGGGAGCA	
	8051					8100
NC_001463	TCTCTGGGAG	TCGCAAACGC	GATTCAGCAG	TCTTACACTA	AGGCAGCTGT	
AF322109	TCTCTGGGAG	TCGCGAACGC	GATTCAGCAG	TCCTACACCA	GGGACGCTGT	
	8101					8150
NC_001463	CCAGACCCTT	GCTAATGCAA	CTGCTGCACA	GCAGGATGTG	TTAGAGGCAA	

AF322109	CCAGACTCTT	GCTAACCGA	CTGCTGTGCA	ACAGCAGGTG	TTAGAGGCGT	
	8151					8200
NC_001463	CCTATGCCAT	GGTACAGCAT	GTGGCTAAAG	GCGTACGAAT	CTTGGAAGCT	
AF322109	CCTATGCCAT	GATACAGCAT	GTGGCTAAGG	GAATACGCAT	CCTTGAAGCA	
	8201					8250
NC_001463	CGAGTGGCTC	GAGTGAAGC	TATCACAGAT	AGAATAATGC	TATACCAAGA	
AF322109	CGCGTGGCGA	GAATGGAAGT	TATGATGGAT	AGAATGATGT	TATATCAGGA	
	8251					8300
NC_001463	ATTGGATTGT	TGGCACTATC	ATCAATACTG	TATAACCTCT	ACAAAAACAG	
AF322109	AGTAGACTGC	TGGCATTATC	ACCAATATTG	TGTAACCTCT	ACAAGAGCAG	
	8301					8350
NC_001463	AAGTAGCAAA	ATATATCAAT	TGGACGAGGT	TTAAGGATAA	TTGCACATGG	
AF322109	ACATAGTGAA	TTACATTAAT	TGGACAAGGT	TTAAAGATAA	TTGCACATGG	
	8351					8400
NC_001463	CAGCAGTGGG	AGAGAGGATT	ACAGGGGTAT	GATACAAACT	TAACAATACT	
AF322109	CAAGAGTGGG	AAAGGGAGAT	AAGTGCGCAT	GAAGGAAACA	TCACTATATT	
	8401					8450
NC_001463	GTTRAAAGAA	TCAGCAGCAA	TGACACAACT	AGCAGAAGAG	CAAGCAAGGA	
AF322109	ACTCAAAGAA	TCAGCAAGGA	TAACACAATT	AGCACAACAA	AAGGTACAAA	
	8451					8500
NC_001463	GGATACCAGA	AGTATGGGAA	AGTTTAAAAG	ACGTCTTTGA	TTGGTCAGGA	
AF322109	GAATACCAGA	TGTGTGGACA	GCACTAAGGG	AGTCACTAGG	ATGGACACAA	
	8501					8550
NC_001463	TGGTTCTCAT	GGCTAAAGTA	TATTCCTATT	ATAGTAGTAG	GATTATTAGG	
AF322109	TGGCTGGCTT	GGATAAAATA	CCTTCCCATA	ATAGTAGTAG	GGATATTAGG	
	8551					8600
NC_001463	ATGCATTCTG	ATAAGAGCTG	TGATATGTGT	ATGTCAACCT	CTTGTGCAGA	
AF322109	ATGCATAATC	ATAAGAATAA	TGTGTGTGT	AGTACAACCA	GTTCTTCAGA	
	8601					8650
NC_001463	TATACAGAAC	TCTAAGTACC	CCGACATACC	AACGGGTCAC	AGTCATCATG	
AF322109	TTTACAGAAC	CTTGACTCAG	ACCAGGTATC	AACAAGTCAA	CTTGGTGATG	
	8651					8700
NC_001463	GAAACAAGAG	CAGACGTCGC	AGGAGAAAAT	CAGGATTTTG	GC...GATGG	
AF322109	GAGACCCGGG	TGCAACTAGA	AGAAGAAGAA	GAAGAAGACG	GAAGGGATGG	
	8701					8750
NC_001463	CTTAGAGGAA	TCAGACAA..	.CAGCGAAC	AAGCGAAAGA	GTGACAGTAC	
AF322109	TGCAGATGGC	TCAGAGAGAT	GCAGCGATCC	CGACAACAAA	GG...AATTA	
	8751					8800
NC_001463	AGAAAGCTTG	GAGCCGTGCC	TGGGAGCTTT	GCGAGAACTC	ACCTTGGGAG	
AF322109	TGAACGCCCTG	GAGGAGAGCT	TGGGTGACTT	GGAGAACTC	ACCTTGGCAG	
	8801					8850
NC_001463	GAGCCATGGA	AAAGGGGCCT	GCTGAGGCTG	CTCGTCCTTC	CGCTGACGAT	
AF322109	AACACATGGA	AGAATGTGGT	GGTGGCGCCG	TTGGTGATTC	CGCTGACAAAT	
	8851					8900
NC_001463	GGGAATCTGG	ATAAATGGAT	GGCTTGGAGA	ACACCACAAA	AATAAAAAAA	

AF322109	CAGAATTTGG	CTCCTTGGAG	AGAATGGAGA	GAACCCCTAA	AAGAAAATA
	8901				8950
NC_001463	GAAAGGGTG	ACTGTGAGAC	ATGGGCTAAA	GAGGACTAAT	AACAAGCTAG
AF322109	AAAAGGGTGG	ACTGTGAGGA	CTGTG.....	.AGGCCTAGG	AGCGAGATAG
	8951				9000
NC_001463	GCCAAATTCC	TGTAATCAC	TGGGGGGTT	ATAAGAAAAG	CAAGTTCACT
AF322109	...AAACTTA	TAGGCCTCTC	TTCCCGG...AAAG	CTAACTCACT
	9001				9050
NC_001463	ATGACAAAGC	AAAATGTAAA	GGCCAAATTC	CTGTAAATCA	CTTGGGGGGT
AF322109	GTG.....	.AGAGGAATA	G..CAAGTCA	CAGTGA..CA	CT.....GCT
	9051				9100
NC_001463	TATAAGAAAA	GCAAGTTCAC	TATGACAAAG	CAAAATGTAA	CCGCAAG...
AF322109	AATTGTACCC	GCAA...CCC	TGAGATCATG	CAAACCACAA	TCCTGAGATT
	9101				9150
NC_001463	.TGCTGACAG	ATGTAACAGC	TGACATATCA	GCTGATGCTT	GCTCATGCTG
AF322109	ATGCTGACAT	GTGTAACAGC	TGATGCCTCA	GCTGATGCTT	GCTCATGCTG
	9151				9200
NC_001463	ACACTGTAGC	TCTGAGCTGT	ATATAAGGAG	AAGCTTGCTG	CTTGC.ACTT
AF322109	ACAATGTAAC	TAGGAGCTCT	ATATAACAG	AGCCCTAGAG	CTTGCTACTT
	9201				9250
NC_001463	CAGAGTTCTA	GGAGAGTCCCTCCT.AGTC	TCTCCTCTCC
AF322109	CAGAGTGCTC	TGAGGAGCTC	GAAGGAAAGA	GTCCTCAGCC	TCTCCTCTCC
	9251				9300
NC_001463	GAGGAGGTAC	CGAGACCTCA	AAATAAAGGA	GTGATTGCCT	TACTGCCGA.
AF322109	GAGGAGCTTC	GG....CTCA	TAATAAAGGA	GTGCTTGCTT	CA..ACAGAA

<397>

표 2

PileUp	
MSF: 759 Type: N Check: 1376 ..	
Name: NC_001463 (gag720bp) (SEQ ID NO: 3) Len: 759	Check: 9060 Weight: 0
Name: AF322109 (gag720bp) (SEQ ID NO: 4) Len: 759	Check: 2316 Weight: 0
//	
NC_001463 (gag720bp)	1 50
AF322109 (gag720bp)ATGGTGAGG CAGGCCTCCG GAAGGGGAAA
NC_001463 (gag720bp)	51 100
AF322109 (gag720bp)	GGAGTACCCC GAGCTAAAAG AATGTCTGAA AAAGGCATGC AAAATAAAAG
NC_001463 (gag720bp)	101 150
AF322109 (gag720bp)	TAAGGGCTGG GGGGAGCGC CTGACAGAAG GAAATTGTCT CTGGTGTATA
NC_001463 (gag720bp)	151 200
AF322109 (gag720bp)	AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTGGACAAA
NC_001463 (gag720bp)	201 250
AF322109 (gag720bp)	AGAAAAATTT AGGACAATAT GGCAGAAGTT GAAGAATCTA ACCTCTGAGG
NC_001463 (gag720bp)	251 300
AF322109 (gag720bp)	AGAGTAACAA AAAAGACTTT ATGCTTTTGC AGGCCACATT AGCGGTCTA
NC_001463 (gag720bp)	301 350
AF322109 (gag720bp)	GTCTGTGTGC AAATGGGAAT GCAGCCCGAG AACTGCAGG ATGCCATAGC
NC_001463 (gag720bp)	351 400
AF322109 (gag720bp)	TACAGTAATC ATGAAAGATG GGTACTGGA ACAAGAGGAA AAGAAGGAA
NC_001463 (gag720bp)	401 450
AF322109 (gag720bp)	CCATCAGAAG AAAAGAAGG AATATAT..C CCATATTAGT GCAGGCAGGA
NC_001463 (gag720bp)	451 500
AF322109 (gag720bp)	GGAGGGAGAA GCTGGAAAGC AGTAGATTCT GTAATGTTCC AGCAACTGCA
NC_001463 (gag720bp)	501 550
AF322109 (gag720bp)	AACAGTAGCA ATGCAGCATG GCCTCGTGTG TGAGGACTTT GAAAGGCAGT
NC_001463 (gag720bp)	551 600
AF322109 (gag720bp)	TGGCATATTA TGCTACTACC TGGACAAGTA AAGACATACT AGAAGTATTG
NC_001463 (gag720bp)	601 650
AF322109 (gag720bp)	GCCATGATGC CTGGAAATAG AGCTCAAAA GAGTTAATTC AAGGGAAATT
NC_001463 (gag720bp)	651 700
AF322109 (gag720bp)	AAATGAAGAA GCAGAAAGGT GGAGAAGGAA TAATCCACCA CCTCCAGCAG
NC_001463 (gag720bp)	701 750
AF322109 (gag720bp)	GAGGAGGATT AACAGTGGAT GCGGGTTC AACCGTGGAT CAGATAATGG GGGTAGGACA AACGAATCAG
	751

```

NC_001463(gag720bp) .....
AF322109(gag720bp) GCAGCGGCA
.....

PileUp

MSF: 1347  Type: N  Check: 2008  ..

Name: NC_001463(gag) (SEQ ID NO: 5) Len: 1347  Check: 6959  Weight: 0
Name: AF322109(gag) (SEQ ID NO: 6) Len: 1347  Check: 5049  Weight: 0

//

      1                               50
NC_001463(gag)  ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGGAAAAG
AF322109(gag)  ..... .ATGGTGAGG CAGGCCTCCG GAAGGGGAAA

      51                               100
NC_001463(gag)  AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAAAG
AF322109(gag)  AAAACACTCC GAGCTAAAAG AATGTCTGAA AAAGGCATGC AAAATAAAAAG

      101                              150
NC_001463(gag)  TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCCTT
AF322109(gag)  TAAGGGCTGG GGGGGAGCGC CTGACAGAAG GAAATTGTCT CTGGTGTATA

      151                              200
NC_001463(gag)  AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTTGGACAAA
AF322109(gag)  AAAACACTAG AGTGTATGTA TGAGGATTGT AGGGAGGAAC CTTGGACCCC

      201                              250
NC_001463(gag)  AGTAAAATTT AGGACAATAT GGCAGAAGGT GAAGAATCTA ACTCCTGAGG
AF322109(gag)  AGAAAAATGT AAACAATTAT GGAAAAAGTT GAAGCAGGTA GAGCCTGAGG

      251                              300
NC_001463(gag)  AGAGTAACAA AAAAGACTTT ATGTCTTTGC AGGCCACATT AGCGGGTCTA
AF322109(gag)  AGAGTAGCAA AGCAGACTAT AACTCGTTAA AAGCAACCTT GCGGGGATA

      301                              350
NC_001463(gag)  ATGTGTTGCC AAATGGGGAT GAGACCTGAG ACATTGCAAG ATGCAATGGC
AF322109(gag)  GTCTGTGTGC AAATGGGAAT GCAGCCCGAG ACATCTCAGG ATGCGATAGC

      351                              400
NC_001463(gag)  TACAGTAATC ATGAAAGATG GGTACTGGA ACAAGAGGAA AAGAAGGAAG
AF322109(gag)  AACCTTAAAC ATGAGAGATG .....AA GTAAAAGGAA AGGAA..AAG

      401                              450
NC_001463(gag)  ACAAAAAGAGA AAAGGAAGAG AGTGTCTTCC CAATAGTAGT GCAAGCAGCA
AF322109(gag)  CCATCAGAAG AAAAGAAGGG AATATATCCC ..ATATTAGT GCAGGCAGGA

      451                              500
NC_001463(gag)  GGAGGGAGAA GCTGGAAAGC AGTAGATTCT GTAATGTTC AGCAACTGCA
AF322109(gag)  GGAGGAAGAG CATGGAGAGC GGTAGAGCCT GCTACCTTTC AGCAGCTCCA

      501                              550
NC_001463(gag)  AACAGTAGCA ATGCAGCATG GCCTCGTGTC TGAGGACTTT GAAAGGCAGT
AF322109(gag)  AACAGTGGCA ATGCAGCATG GACTAGTATC AGAAGAATT GAAAGGCAGC

      551                              600
NC_001463(gag)  TGGCATATTA TGCTACTACC TGGACAAGTA AAGACACTACT AGAAGTATTG
AF322109(gag)  TAGCATACTA TGCCACCACA TGGACAAGCA AGGATATCTT AGAAGTATTA

      601                              650
NC_001463(gag)  GCCATGATGC CTGGAATAG AGCTCAAAAG GAGTTAATTC AAGGGAAATT
AF322109(gag)  GCCATGATGC CAGGAAATAG AGCGCAAAA GAACATAAC AAGGAAAGTT
    
```

	651		700
NC_001463 (gag)	AAATGAAGAA GCAGAAAGGT GGAGAAGGAA TAATCCACCA CCTCCAGCAG		
AF322109 (gag)	AAATGAGGAA GCAGAGAGAT GGAGAAGGCA GAATCCACAA CCTGCGG...		
	701		750
NC_001463 (gag)	GAGGAGGATT AACAGTGGAT CAAATTATGG GGGTAGGACA AACAAATCAA		
AF322109 (gag)	...GCGGGTT AACCGTGGAT CAGATAATGG GGGTAGGACA AACGAATCAG		
	751		800
NC_001463 (gag)	GCAGCAGCAC AAGCTAACAT GGATCAGGCA AGGCAAATAT GCCTGCAATG		
AF322109 (gag)	GCAGCCGCAC AGGCTAATAT GGATCAAGCA AGACAAATAT GCCTACAATG		
	801		850
NC_001463 (gag)	GGTAATAAAT GCATTAAGAG CAGTAAGACA TATGGCGCAC AGGCCAGGGA		
AF322109 (gag)	GTTATAACA GCAATAAGAG GAGTTAGGCA TATGGCCCAT AGACCAGGAA		
	851		900
NC_001463 (gag)	ATCCAATGCT AGTAAAGCAA AAAACGAATG AGCCATATGA AGATTTTGCA		
AF322109 (gag)	ATCCCATGCT GGTAAAGCAA AAACCAATG AGAATATGA AGAGTTTGCC		
	901		950
NC_001463 (gag)	GCAAGACTGC TAGAAGCAAT AGATGCAGAG CCAGTTACAC AGCCTATAAA		
AF322109 (gag)	GCAAGTTTGT TAGAAGCAGT GGATGCAGAA CCCGTTACCC AACCTATAAA		
	951		1000
NC_001463 (gag)	AGATTATCTA AAGCTAACAC TATCTTATAC AAATGCATCA GCAGATTGTC		
AF322109 (gag)	AGAATATTTA AAGGTAACCT TGTCTTACAC AAATGCAAAT TCGGAATGTC		
	1001		1050
NC_001463 (gag)	AGAAGCAAAT GGATAGAACA CTAGGACAAA GAGTACAACA AGCTAGTGTA		
AF322109 (gag)	AAAAACATAT GGACAGAGTG TTGGGGCAAA GAGTACAGCA GGCCCAATA		
	1051		1100
NC_001463 (gag)	GAAGAAAAAA TGCAAGCATG TAGAGATGTG GGATCAGAAG GGTTCAAAAT		
AF322109 (gag)	GAAGAAAAAA TGCAGGCATG CAGGGACATC GGGGGAACAG CATATCAGAT		
	1101		1150
NC_001463 (gag)	GCAATTGTTA GCACAAGCAT TAAGGCCAGG AAAAGGAAAA GGGAAATGCAC		
AF322109 (gag)	GCAGTTGCTT GCACAAGCCC TCCGTGGCGG AAAAGAAGAT GGGAAAAAAT		
	1151		1200
NC_001463 (gag)	AGCCACAAAG GTGTTACAAC TGTGGAAAAC CGGGACATCA AGCAAGGCAA		
AF322109 (gag)	CTGTAGGAA GTGTTATAAC TGTGGAAGGC CCGGACACAG AGCAAAAGAA		
	1201		1250
NC_001463 (gag)	TGTAGACAAG GAATCATATG TCACAACCTGT GGAAAGAGAG GACATATGCA		
AF322109 (gag)	TGCAGACAAG GCATTATATG TCACAACCTGT GGAAAAAGAG GGCATATACA		
	1251		1300
NC_001463 (gag)	AAAAGAATGC AGAGGAAAGA GAGACATAAG GGGAAAACAG CAGGGAAACG		
AF322109 (gag)	GAAAACTGC A.....AACA GAAA....AG AAGAAAGGAG CAGGGAAACA		
	1301		1347
NC_001463 (gag)	GGAGGAGGGG GATACGTGTG GTGCCGTCCG CTCCTCCTAT GGAATAA		
AF322109 (gag)	TGAGGAGGGG GCTACGTGTG GTGCCGTCCG CACCCCTAT GGAATAA		

표 3

```

FileUp
MSF: 605   Type: N   Check: 9138   ..
Name: NC_001463(5') (SEQ ID NO: 7) Len: 605   Check: 5398   Weight: 0
Name: AF322109(5') (SEQ ID NO: 8) Len: 605   Check: 3740   Weight: 0
//
      1                               50
NC_001463(5') .....GAGTTC TAGG...AGA GTCCCTCCTA GTCTCTCCTC
AF322109(5') GTGAGTGCTC TGAGGAGCTC GAAGGAAAGA GTCC.TC..A GCCTCTCCTC

      51                               100
NC_001463(5') TCCGAGGAGG TACCGAGACC TCAAATAAA GGAGTGATTG CCTTACTGCC
AF322109(5') TCCGAGGAGC TTCGG....C TCATAATAAA GGAGTGCTTG CTTCA..ACA

      101                              150
NC_001463(5') GAGTGGAGAG TGATTACTGA GCGGCCGGTG TATCGGGAGT CGTCCCTTAA
AF322109(5') GAAGCTGAG.. .....CTGG TCGTGGTTAT TATCGGG... .GACCGAAGT

      151                              200
NC_001463(5') TCTGTGCAAT ACCAGAGCGG CTCTCGCAGC TGGCGCCCAA CGTGGGGCCC
AF322109(5') CCCGTGCAAC ACCGGGGCGG TTCTCGCAGC TGGCGCCCAA CGTGGGGCTC

      201                              250
NC_001463(5') GAGGAG.....
AF322109(5') GAGTAGCTTG AGAAGCTCGA CTGAGATCTG AATCCAAGAG CGACATCAGA

      251                              300
NC_001463(5') .....AAGAAA AGAAAGC... GGCCCTGAGA ACTCGGCTTC TG..AAAAAG
AF322109(5') CAGCAAGAAA TGAGAGTAAT GAGACCGCGA GCTCTGCTGC TGTAAAAAAG

      301                              -350
NC_001463(5') AGGAAGAGGA CAAGTTGCTA TAGCAACAAG AGAGAAGAAG TAGAGCAAAG
AF322109(5') AGGAAGTAG. CGGTTGCCG AGGCAACTGC TCAGAAGAAC CAGGGGAAAG

      351                              400
NC_001463(5') GTCCAGTGGC T.CGGAAAAA GAGGAAGTGA AACTTCGGGG ACGCCTGAAG
AF322109(5') GGCTTCCAGC AACCTCAAAA GAGGAACCGA GACTTCGGGG ACGCCTGAA.

      401                              450
NC_001463(5') GAGTAAGGTA AGTGAAGTCTG CTGTACGCGG GGCGAGGCAG AGGTT.TCCT
AF322109(5') ..GTAAGGTA AGTGAAGTCTG CTGTACGCGG GGCGAGGCAT AGGAGATCCT

      451                              500
NC_001463(5') TCTAAATT.G AAAGAGAAGT GTTGCTGCGA GAGGTCTTGG TGGTGAGAGAA
AF322109(5') TCTATTCTAG GAAGAGAAGC GCTGTTCTGG GAGGTCTTGG CGACCGAGAA

      501                              550
NC_001463(5') TCCTGTACAA AAAAAAGGAG GGATCTCGGT CAGGACCAGG ACCCCTGGGA
AF322109(5') TCTTGT... AAATAAGCCA GGATCTCGAT CAGGACCAAG ACCCCTCAGG

      551                              600
NC_001463(5') GTAATACAAC AGCAACACCG TAAGAAAATC CGCC.....
AF322109(5') AGAGGGTATA GACAGCGTGG TAAGAAA.TC CGCCGTGGTG AGTCTAGATA

      601
NC_001463(5') .....
AF322109(5') GAGAC
    
```

<401>

표 4

```

FileUp
MSF: 3338 Type: N Check: 5428 ..
Name: NC_001463(pol) (SEQ ID NO: 9) Len: 3338 Check: 8114 Weight: 0
Name: AF322109(pol) (SEQ ID NO: 10) Len: 3338 Check: 7314 Weight: 0
//
1 50
NC_001463(pol) .....AT GTCACAAC TGGAAGAGA GGACATATGC
AF322109(pol) ATGCAGACAA GGCATTATAT GTCACAAC TGGAAGAGA GGCATATAC
51 100
NC_001463(pol) AAAAGAATG CAGAGGAAAG AGAGACATAA GGGGAAACA GCAGGAAAC
AF322109(pol) AGAAAACTG CA....AAC AGAA...A GAAGAAAGGA GCAGGAAAC
101 150
NC_001463(pol) GGGAGGAGGG GGATACGTGT GGTCCCGTCC GCTCCTCCTA TGGATAACT
AF322109(pol) ATGAGGAGGG GGCTACGTGT GGTCCCGTCC GCACCCCTA TGGAGTAACG
151 200
NC_001463(pol) TCAGCACACC CTATGGTTCA GGTCCGCATA GGTTCACAGC AGAGGAACCT
AF322109(pol) CAAGCACACC TAATAGTTAG GGTACAAATA GGGAAATCAGG AGAAACAATT
201 250
NC_001463(pol) GTTATTTGAT ACCGGGGCGG ACCGAACATAT AGTTAGATGG CATGAGGGCT
AF322109(pol) ATTATTTGAC ACAGGGGCAG ATAAAACGAT AGTAAGAATG CATGATGGAA
251 300
NC_001463(pol) CGGGAAACCC AGCCGGAAGG ATAAAACATGC AAGGAATAGG AGGAATAGTA
AF322109(pol) CAGGGATTCC AAACGGAAGA ATAAAATTAC AAGGATAGG AGGAATAGTA
301 350
NC_001463(pol) GAAGGAGAAA AATGGAATA TGTAGAATTA GAATATAAAG GAGAAACAAG
AF322109(pol) GAAGGAGAAA AATGGAATA AGTACCCATG ACATATAAAG GAGAAACATC
351 400
NC_001463(pol) AAAGGGAACA ATAGTAGTGT TACCACAAAG TCCAGTAGAA GTATTAGGAC
AF322109(pol) CTGCCCAAGC TTGGTTGTGC TAAGAGATAG CCCAGTAGAA GTATTGGGAA
401 450
NC_001463(pol) GAGATAACAT GGCCCGATTT GGAATAAAGA TAATAATGGC AAATTTAGAG
AF322109(pol) GAGATAACAT GGAAGCATTG GCGTAACCC TAATAATGGC AAATTTAGAA
451 500
NC_001463(pol) GAAAAAGAAA TCCCAATTAC AAAAGTAAAA TTGAAAGAGG GATGTACGGG
AF322109(pol) GATAAGAAA TTCCACAAT ACCAGTAGAA TTGAAAGAGG GATGTAAGG
501 550
NC_001463(pol) TCCACATGTC CCACAATGGC CATTACAGA AGAGAAATTA AAAGGTCTAA
AF322109(pol) GCCACATGTC CCCAGTGGC CATTACAGC AGAGAAATTA CAAGACTAA
551 600
NC_001463(pol) CACAAATCAT AGATAAATTA GTGGAAGAAG GAAACTAGG AAAGGCACCC
AF322109(pol) CAGGAATAGT AGAAAAATTA CTACAGGAAG GAAATTTGGC AGAGGCCCCA
601 650
NC_001463(pol) CCACATTGGA CATGTAATAC TCCAATCTTT TGCATAAAAA AGAAATCAGG
AF322109(pol) GAGGGATGGA CGTGAACAC GCCCATCTTC TGCATAAAAA AGAAGTCAGG
    
```

<402>

	651		700
NC_001463 (pol)	GAAGTGGAGA ATGTTAATAG ATTCAGAGA ATTGAACAAA CAGACAGAAG		
AF322109 (pol)	AAAATGGAGA ATGTTAATAG ATTTAGGGA ATTAAATAAG CAAACAGCAG		
	701		750
NC_001463 (pol)	ATTTAACAGA AGCGCAGTTA GGACTCCCGC ATCCGGGAGG ACTACAAAAG		
AF322109 (pol)	ATTTAGCAGA AGCGCAGCTA GGACTGCCAC ACCCAGGAGG GTTGCAAAGG		
	751		800
NC_001463 (pol)	AAAAAACATG TTACAATATT GGACATAGGA GATGCATATT TTAATAACC		
AF322109 (pol)	AAAAAGAATG TAACAATTCT GGACATAGGA GATGCATATT TCACAATTCC		
	801		850
NC_001463 (pol)	CCTATATGAA CCATATCGAG AGTACACATG TTTACTCTA TTAAGTCCTA		
AF322109 (pol)	CTTATACGAG CCCTATCAGA AATATACATG CTTACACTC CTAAGTCCTA		
	851		900
NC_001463 (pol)	ATAATCTAGG ACCATGTAAA AGATACTATT GGAAAGTCTC GCCACAAGGT		
AF322109 (pol)	ACAATTTGGG ACCATGTAAA AGTATTATT GGAAAGTATT ACCCCAGGGA		
	901		950
NC_001463 (pol)	TGGAAATTGA GTCCATCTGT ATATCAATTT ACTATGCAGG AGATCTTAGA		
AF322109 (pol)	TGGAAATTGA GCCCAGCTGT ATATCAATTC ACCATGCAA GGTGTGTAAA		
	951		1000
NC_001463 (pol)	GGATTGGATA CAGCAGCATC CAGAAATTCA ATTTGGCATA TATATGGATG		
AF322109 (pol)	AGGATGGATA CAACAGCATA AAAACATACA ATTTGGAATA TATATGGATG		
	1001		1050
NC_001463 (pol)	ATATTTACAT AGGAAGTGAT TTAGAAATTA AAAAGCATAG AGAAATAGTG		
AF322109 (pol)	ATATCTATAT TGGAAAGTAT CTAACGATAG CCCAACATAG GAAGATAATA		
	1051		1100
NC_001463 (pol)	AAAGATTTAG CCAATTATAT TGCCCAATAT GGATTCACTC TGCCAGAAGA		
AF322109 (pol)	GAAGAATTAG CCTCATTAT AGAACAATTT GGGTTTACAT TACCAGAAGA		
	1101		1150
NC_001463 (pol)	GAAGAGACAA AAGGGATATC CAGCAAAATG GCTAGGATTT GAACTACACC		
AF322109 (pol)	TAAGAGACAA GAGGGCTATC CAGCAAAATG GCTAGGATTG GAGCTACATC		
	1151		1200
NC_001463 (pol)	CGCAGACCTG GAAATTTGAG AAGCATAAT TACCTGAATT AACAAAGGGA		
AF322109 (pol)	CAGAAAAATG GAAATATCAA AAGCATAAAT TGCCGGAATT ACAAGAGGGG		
	1201		1250
NC_001463 (pol)	ACAATAACAT TAAATAAATT ACAGAAATTA GTAGGAGAAT TAGTATGGAG		
AF322109 (pol)	GTAATAACCC TGAACAAATT ACAGAAGATA GTAGGGGAAT TAGTGTGGAG		
	1251		1300
NC_001463 (pol)	ACAATCCATA ATTGGGAAAA GCATTCTTAA CATTCTGAAA TTAATGGAAG		
AF322109 (pol)	ACAATCCTTG ATAGGAAAGA GCATCCCCAA TATCATAAAA TTAATGGAAG		
	1301		1350
NC_001463 (pol)	GAGATAGAGA ATTACAAAGT GAAAGAAAA TTGAAGAAGT ACATGTGAAA		
AF322109 (pol)	GAGATCGCCG ATTACAAAGT GAAAGAAAA TAGAAAGAAT ACATGTACAA		
	1351		1400
NC_001463 (pol)	GAATGGGAAG CATGTAGGAA AAAATTAGAA GAAATGGAAG GAAATTATTA		
AF322109 (pol)	GAATGGGAAG CATGTCAAAA GAAATTAGAT GAAATGGTAG GAAATTATTA		
	1401		1450

NC_001463 (pol)	TAATAACAC	AAAGATGTCT	ATGGACAATT	GGCTTGGGGA	GACAAAGCTA
AF322109 (pol)	CAGAGAAGAA	GAAGATATCT	ATGGACAAT	AACTTGGGG	GATAAGGCAA
	1451				1500
NC_001463 (pol)	TAGAAATAT	AGTGTATCAG	GAGAAAGGGA	AACCATTATG	GGTAAATGTG
AF322109 (pol)	TAAAATACAT	AGTATTCCAA	AGGAAAGGGG	AACCCCTATG	GGTAAATGTA
	1501				1550
NC_001463 (pol)	GTTACAATA	TAAAGAACCT	AAGCATCCCG	CAACAGGTTA	TTAAAGCAGC
AF322109 (pol)	GTACATGACA	TAAAAAATT	GAGTCTCCCA	CAGCAAGTGA	TAAAGCAGC
	1551				1600
NC_001463 (pol)	GCAAAAATTA	ACCCAAGAAG	TCATCATTAG	GACAGGAAAA	ATACCATGGA
AF322109 (pol)	ACAGAAAATTA	ACCCAGGAAG	TAATCATAAG	AACAGGAAAA	ATCCCATGGC
	1601				1650
NC_001463 (pol)	TATTGTTGCC	AGGGAAGAA	GAAGATTGGA	GACTAGAATT	GCAATTAGGG
AF322109 (pol)	TGCTACTACC	AGGAAGAGAA	GAAGACTGGA	GATTAGAACT	GCAGGTAGGG
	1651				1700
NC_001463 (pol)	AACATCACAT	GGATGCCAAA	ATTTTGGTCC	TGTTATCGAG	GA.CATACAA
AF322109 (pol)	AACATCACGT	GGATGCCATC	ATTTTGGTCA	TGTTATCGAG	GAGCACCCAA
	1701				1750
NC_001463 (pol)	GATGGAGAAA	AAGAAATATA	ATAGAAGAAG	TAGTAGAAGG	GCCTACATAT
AF322109 (pol)	G.TGGAAG	AAGAACATA	GTGGCAGCAG	TGGTAGATGG	ACCGACATAT
	1751				1800
NC_001463 (pol)	TATACAGATG	GAGGAAAAAA	GAATAAAGTA	GGAAGTCTAG	GGTTCATAGT
AF322109 (pol)	TATACAGATG	GGGGAAGAA	AAACGCACAG	GGAAGCTTGG	GCTTCATCTC
	1801				1850
NC_001463 (pol)	ATCAACAGGG	GAAAAATTTA	GAAAGCATGA	AGAGGGCACA	AACCAGCAAC
AF322109 (pol)	CCCAACAGGA	GAAAAGTTCA	GAAGGCATGA	AGATGGAACT	AATCAGGTAT
	1851				1900
NC_001463 (pol)	TAGAATTAAG	AGCCATAGAG	GAAGCTCTAA	AACAAGGGCC	TCAAACAATG
AF322109 (pol)	TAGAATTAAG	GGCAATAGAA	GATCCATGTA	AACAAGGACC	TGAAAGCATG
	1901				1950
NC_001463 (pol)	AATTTAGTAA	CAGATAGTAG	ATATGCATTT	GAATTTTAT	TAAGAAATG
AF322109 (pol)	AACATTTGTA	CTGACAGCAG	GTATGCTTAT	GAATTCATGC	TCCGAAACTG
	1951				2000
NC_001463 (pol)	GGATGAAGAA	GTAATAAAGA	ATCCAATTCA	AGCAAGAATT	ATGGAATG
AF322109 (pol)	GGATGAACAG	GTCATAAGAA	ACCCCATTC	GGCAAGAATC	ATGGCAGAAG
	2001				2050
NC_001463 (pol)	CCCACAAGAA	AGATAGGATA	GGAGTGCAIT	GGGTGCCAGG	ACATAAAGGG
AF322109 (pol)	TGCACAAGAA	AAAGCAGGTA	GGAATACACT	GGGTGCCAGG	GCATAAAGGA
	2051				2100
NC_001463 (pol)	ATTTCCCAAA	ATGAAGAAAT	AGACAAATAT	ATTTCCGAAA	TATTTCTTGC
AF322109 (pol)	ATACCTCAGA	ATGAAGAGAT	AGACCAGTAC	ATATCAGAAG	TATTTCTTAGC
	2101				2150
NC_001463 (pol)	AAAAGAAGGA	GAAGGAATTC	TCCCAAAAAG	AGAAGAGGAT	GCAGGTATG
AF322109 (pol)	ACGAGAAGGA	ACAGGGATAT	GTGAAAAAAG	GAAGGAAGAT	GCTGGATATG
	2151				2200
NC_001463 (pol)	ATTTAATATG	CCCAGAAGAG	GTTACCATAG	AGCCAGGACA	AGTGAATGC
AF322109 (pol)	ATTTATTATG	CCCGCATGAG	GTAATACTTA	AACCCCAAGA	AGTAAACCG
	2201				2250

NC_001463 (pol)	ATCCCATAG	AGCTAAGATT	AAATTTAAAG	AAATCACAAT	GGGCTATGAT
AF322109 (pol)	ATCCCAATAG	ACCTAAAATT	AAAATTGAAA	GAAAAGCAAT	GGGCCATGAT
	2251				2300
NC_001463 (pol)	TGCTACAAA	AGCAGCATGG	CTGCCAAAGG	AGTGTTCACA	CAAGGAGGAA
AF322109 (pol)	AAGTGGGAAA	AGTAGCGTTG	CAGCAAAAGG	AATATTTGTA	CAAGGAGGCA
	2301				2350
NC_001463 (pol)	TCATAGACTC	AGGATATCAG	GGACAAATAC	AGGTAATAAT	GTATAATAGC
AF322109 (pol)	TAATAGATTG	AGGGTATCAG	GGACAAGTAC	AAGTCATCCT	ATATAATAGT
	2351				2400
NC_001463 (pol)	AATAAAATAG	CAGTAGTCAT	ACCCCAAGGG	AGAAAATTTG	CACAATTAAT
AF322109 (pol)	AATAAGATAG	AGGTCAAAAT	ACCACAAGGC	AGGAAAATTTG	CCCAATTAAT
	2401				2450
NC_001463 (pol)	ATTAATGGAT	AAAAGCATG	GAAAATTGGA	ACCTGGGGG	GAAAGCAGAA
AF322109 (pol)	ATTAATGAAC	TTACAACATG	AAGAATTAGA	AGAATGGGGA	AAGGAAAGAA
	2451				2500
NC_001463 (pol)	AAACAGAAAG	GGGAGAAAA	GGATTGGGT	CTACAGGAAT	GTATTGGATA
AF322109 (pol)	AAACAGAAAG	AGGAACAAAA	GGATTGGGT	CTACAGGAGC	ATTTTGGATA
	2501				2550
NC_001463 (pol)	GAAAATATTC	CTCTGGCAGA	GGAAGACCAC	ACAAAATGGC	ATCAAGATGC
AF322109 (pol)	GAGAATATTC	CCCAAGCAGA	GGAAGAACAT	TACAAATGGC	ATCAAGATGC
	2551				2600
NC_001463 (pol)	CCGATCATTG	CATCTAGAAT	TTGAAATTC	AAGAACAGCA	GCAGAAGACA
AF322109 (pol)	TAGATCTCTG	CAGCTAGAAT	TCAAGATACC	TAGAGCAGCA	GCAGAAGACA
	2601				2650
NC_001463 (pol)	TAGTAAATCA	ATGTGAAATA	TGCAAAGAAG	CGAGGACACC	TGCAGTAATT
AF322109 (pol)	TTATACAGCA	CTGTGAGGTA	TGTCAAGAAG	GCAAACCCGC	AGCGATCAGG
	2651				2700
NC_001463 (pol)	AGAGGCGGAA	ACAAAAGGGG	GGTAAATCAT	TGGCAAGTGG	ATTATACCCA
AF322109 (pol)	AGAGGGGGAA	ATAAAAGAGG	AATAGATCAT	TGGCAGGTAG	ACTATACACA
	2701				2750
NC_001463 (pol)	TTATGAAAT	ATCATACTAT	TAGTATGGGT	AGAAACAAAT	TCAGGACTAA
AF322109 (pol)	TTACAAAGAA	CACATAATAT	TAGTATGGGT	AGAGACTAAT	TCAGGATTAA
	2751				2800
NC_001463 (pol)	TATATGCAGA	AAAAGTAAAA	GGAGAATCAG	GGCAAGAATT	CAGAATAAAA
AF322109 (pol)	TATTTGCAGA	GAAAGTAAAA	GGAGAATCAG	GACAAGAATT	TAGGATGCAG
	2801				2850
NC_001463 (pol)	GTGATGCATT	GGTATGCATT	ATTTGGTCCA	GAGTCATTGC	AGTCAGACAA
AF322109 (pol)	ACATTGAAAT	GGTATGCTTT	GTTTCAACCA	AAATCAGTGC	AATCAGATAA
	2851				2900
NC_001463 (pol)	TGGACCTGCA	TTTGCAGCAG	AGCCACACAC	GCTGTTAATG	CAATACCTAG
AF322109 (pol)	TGGGACAGCC	TTCACAGCTG	AGGCTACGCA	GCATCTAATG	AAGTATTTAG
	2901				2950
NC_001463 (pol)	GAGTAAAACA	CACAACAGGC	ATACCTTGGG	ATCCACAGTC	TCAGGCTATA
AF322109 (pol)	GGATTGAGCA	CACTACGGGT	ATTCGTGGA	ACCCCCAGTC	ACAAAGTTTA
	2951				3000
NC_001463 (pol)	GTAGAAAGGG	CACATCAACT	ATTGAAAAGC	ACTTTAAAGA	AGTTCAGGCC
AF322109 (pol)	GTAGAAAGAG	CTCATCAAAC	ATTAAAACAC	ATGTTAGAAA	AATTAGAACC
	3001				3050

```

NC_001463(pol) ACAATTTGTC GCTGTAGAAT CAGCCATAGC AGCAGCCCTA GTCGCCATAA
AF322109(pol) ACAATTTGTG GCCCTACAGT CTGCCATCGC AGCCACTCTA GTTGCGCTCA
...
3051 3100
NC_001463(pol) ATATAAAAAG AAAGGGTGGG CTGGGGACAA GCCCTATGGA TATTTTATA
AF322109(pol) ATATAAAAAG AAAGGGTGGG CTAGGGGCAA GCCCTATGGA TATTTACATA
...
3101 3150
NC_001463(pol) TATAATAAAG AACAGAAAAG AATAAATAAT AAATATAATA AAAATTCTCA
AF322109(pol) TATAATAAAG AGCAACAAAG ACAACAAGAT AATAGTAATA AATTAATTCA
...
3151 3200
NC_001463(pol) AAAAATTCAA TTCGTATT ACAGAATAAG GAAAAGAGGA CATC.AGGAG
AF322109(pol) GAAAA...AA TTTGTATT ACAGGATCAG AAAAAGAGGC CATCCAGGAG
...
3201 3250
NC_001463(pol) AGTGGAAAGG ACCAACCCAG GTACTGTGGA AAGGGGAAGG AGCCAATTGT
AF322109(pol) AGTGGAAAGG CCCAACTGAG GTACTGTGGG AAGGGGAAGG AGCCA.TAGT
...
3251 3300
NC_001463(pol) GGTAAAGGAT ATAGAAAAGT AAAAGTATT AGTAATACCT TACAAAGATG
AF322109(pol) AGTTAAAGAC AAAGAAAAGT ATAGATATCT AGTCATCCCA TATAAAGATG
...
3301 3338
NC_001463(pol) CAAAATTCAT CCCGCCACCA ACAAAGAAA AGGAATAA
AF322109(pol) CAAAATTTAT TCCGCCACCG TCGGAACAGA AGGGATAG

```

<406>

표 5

```

FileUp
MSF: 408 Type: N Check: 517 ..
Name: NC_001463(rev) (SEQ ID NO: 11) Len: 408 Check: 7287 Weight: 0
Name: AF322109(rev) (SEQ ID NO: 12) Len: 408 Check: 3230 Weight: 0
//
1 50
NC_001463(rev) ATGGATGCTG GGGCCAGATA CATGCGCTTA ACTGGGAAGG AAAACTGGGT
AF322109(rev) .....
51 100
NC_001463(rev) TGAAGTAACC ATGGACGGAG AGAAGGAAAAG GAAAAGAGAA GGTTCACTG
AF322109(rev) ..... ATGG.CGGAG ATAAGAAAAG ..A.AGCAAA GGAGCCACTA
101 150
NC_001463(rev) CGGGACAGCA AGATATACAG AACTCTAAGT ACCCCGACAT ACCAACGGGT
AF322109(rev) ATCCAGGACC AGGTATCAAC AAGTCAACTT GGTGATGGAG ACC..CGGTT
151 200
NC_001463(rev) CACAGTCATC ATGGAACAA GAGCAGACGT CGCAGGAGAA AATCAGGATT
AF322109(rev) ..... ..GCAACTAG AAGAAGAAGA AGAAGAAGAC GGAAGGGATG
201 250
NC_001463(rev) TTGGCGATGG CTTAGAGGAA TCAGACAACA GCGAAACAAG CGAAAGAGTG
AF322109(rev) GTGGAGATGG CTCAGAGAGA TG.....CA GCGATCCCGA CAACAAAGGA
251 300
NC_001463(rev) ACAGTACAGA AAGCTTGGAG CCGTGCCTGG GAGCTTTGGC AGAACTCACC
AF322109(rev) A...TTATGA ACGCCTGGAG GAGAGCTTGG GTGACTTGGG GAAACTCACC
301 350
NC_001463(rev) CTGGAAGGAG CCATGGAAAA GGGGCCTGCT GAGGCTGCTC GTCCTTCCGC
AF322109(rev) TTGGCAGAAC ACATGGAAGA ATGTGGTGGT GCGCCCGTTG GTGATTCCGC
351 400
NC_001463(rev) TGACGATGGG AATCTGGATA AATGGATGGC TTGGAGAACA CCACAAAAT
AF322109(rev) TGACAATCAG AATTTGGCTC CTTGGAGAGA ATGGAGAGAA CCCTAAAAG
401
NC_001463(rev) AA.....
AF322109(rev) AAAAAATA

```

<407>

표 6

```

FileUp
MSF: 691 Type: N Check: 6528 ..
Name: NC_001463(vif) (SEQ ID NO: 13) Len: 691 Check: 5882 Weight: 0
Name: AF322109(vif) (SEQ ID NO: 14) Len: 691 Check: 646 Weight: 0
//
1 50
NC_001463(vif) ATGCAAAATT CATCCCGCCA CCAACAAAAG AAAAGGAATA AAAAACCTGG
AF322109(vif) ATGCAAAATT TATTCCGCCA CCGTCGGAAC AGAAGGGATA GAAGAATAGG

51 100
NC_001463(vif) ACCAGAATTA CCCTTAGCAC TATGGATACA TATAGCAGAA AGCATTAAATG
AF322109(vif) TCCAGAATTG CCTTATCTT TATGGACTTA TACAGCATA CAGCATAAATA

101 150
NC_001463(vif) GGGATAGCTC ATGGTACATA ACAATGAGAC TGCAACAGAT GATGTGGGGA
AF322109(vif) AAGATCCCGC ATGGTATACA ACCCTAAGAC TGCAGCAAAT GATGTGGCAT

151 200
NC_001463(vif) AAAAGAGGAA ATAAGTTACA ATATAAGAAT GAAGACAGGG AATATGAAAA
AF322109(vif) AGGAGGGGAA ATAAATTGAC ATATGTCAGG GAAAATGCAC AGTACGAGGA

201 250
NC_001463(vif) TTGGGAAATT ACATCATGGG GATGGAAAAT GCACCTAAGG AGAGTGAAAC
AF322109(vif) GTGGGAAATG ACCTCGTATG AGTGGAGGAT AAGAATGAGA AGGGACAAAA

251 300
NC_001463(vif) AATGGATACA AGACAACAGG AGAGGAAGC CCATGGCAGT ACAAGTAGG
AF322109(vif) CAAAAGTCA TC.CAAGAGG GCATACTTCG CCATGGCAAT ATCGGAGACA

301 350
NC_001463(vif) AGGAACATGG AAAAGTATAG GAGTGTGGTT CCTGCAAGCA GGAGATTACA
AF322109(vif) GGATGGATGG AAGGATGTGG GAACGTGGTT CCTACAGCCA GGGGACTATA

351 400
NC_001463(vif) GAAAGGTAGA CAGGCACTTC TGGTGGGCAT GGAGGATACT GATATGTTCC
AF322109(vif) GAAAGGCGGA TCAGCAGTTC TGGTTCGCTT GGAGAATAGT GTCGTGTTCA

401 450
NC_001463(vif) TGCAGGAAAG AAAAGTTTGA TATAAGAGAA TTTATGAGAG GAAGACATAG
AF322109(vif) TGTA AAAAGG AAGGATTTAA CATAAGAGAA TTTATGCTAG GTACCCATAG

451 500
NC_001463(vif) ATGGGATTTG TGCAATCCT GTGCTCAAGG AGAAGTAGTA AAGCATACTA
AF322109(vif) ATGGGATTTG TGTAAGTCCT GTTGCCAGGG TGAAGTACTA AAGAGAACAC

501 550
NC_001463(vif) GAACAAAAG TCTGGAAAGA CTAGTACTGC TACAGATGGT AGAACAGCAT
AF322109(vif) AACCTACAC CTGCAAAGG CTCACGTGGC TTAATTTAAC AGAAGACCAT

551 600
NC_001463(vif) GTGTTTCAAG TATTGCCATT GTGGAGAGCC AGGAGAAGTA GTACAACAGA
AF322109(vif) GTATTTCAAG TAATGCCCTT GTGGAGAGCT CGCAAAGGGA TTACCATAGA

601 650
NC_001463(vif) TTTCCCATGG TGCAGGGACA CAACGGGATA CACGCATGCG TGGTCTGTCC
AF322109(vif) CTTTCCCTGG TGCAGGGACA CAAAAGGATT CCTGGAGCCG TGGACAACGC
    
```

```

651 691
NC_001463(vif) AGGAGTGCTG GTTGATGGAA TATCTCTTAG AGGATGAGTG A
AF322109(vif) AAGAGTGTG GCAAATAGAG TATCCCTTGG AGGATGACTG A
    
```

<408>

<409>

표 7

```

PileUp
MSF: 736  Type: N  Check: 513  ..
Name: NC_001463(gag720bp) (SEQ ID NO: 15) Len: 736 Check: 4701 Weight: 0
Name: >AF015181 (SEQ ID NO: 16) Len: 736 Check: 5812 Weight: 0

//

1 50
NC_001463(gag720bp) .ATGTTGAGT CTAGATAGAG ACATGGCGAG GCAAGTCTCC GGGGGGAAAA
>AF015181 GCTGTAGACT CTGTAATGTT CCAACAA.AT GCAA.....C AGTAGCAATG

51 100
NC_001463(gag720bp) GAGATTATCC TGAGCTCGAA AAATGTATCA AGCATGCATG CAAGATAAAA
>AF015181 CAGCATGGCC TCGTGTCCGA GGATTTTGAA AGACAGTTAG CAT.ATTATG

101 150
NC_001463(gag720bp) GTTCGACTCA GAGGGG..AG CACTTGACAG AAGGAAATG TTTATGGTGC
>AF015181 CTACTACCTG GACAAGTAAA GACATACTAG AAGTA..TTG GCCATGATGC

151 200
NC_001463(gag720bp) CTTAAAAACA...TTAGATT..ACATGTTT GAGGACCAT..AAAGAGGAA
>AF015181 CTGGGAATAG GGCTCAGAAA GAACCTATTC AAGGGAATT GAATGAAGAA

201 250
NC_001463(gag720bp) CCTTGGACAA AAGTAAATTT TAGGACAATA TGGCAGAAGG..TGAAGAATC
>AF015181 GCA..GACAG GTGGAGAAG..GAACAATC CACCAGGAGG ATTAACAGTG

251 300
NC_001463(gag720bp) TAACCTCTGA GGAGAGTAAC AAAAAAGACT TTATGTCTTT GCAGGCCACA
>AF015181 GATCAAATTA TGGGGTAGG ACAAACAAAT CAAGCA....GCAGCACAAG

301 350
NC_001463(gag720bp) TTAGCGGGTC TAATGTGTTG CCAAATGGGG ATGAGACCTG AGACATTGCA
>AF015181 CTAACATGGA TCAGGCAA.G ACAAATATGC CT...ACAAT.GGGTAATAAA

351 400
NC_001463(gag720bp) AGATGCAATG GCTACAGTAA TCATGAAAGA TGGTACTG GAACAAGAGG
>AF015181 CGCCTTAAGA GCAGTAAGGC ATATGGCTCA TAGGCCAGGG AATCCAATGC

401 450
NC_001463(gag720bp) AAAAGAAGGA AGACAAAAGA GAAAAGGAAG AGAGTGTCTT CCCAATAGTA
>AF015181 TAGTAAAGCA A...AAAACA AATGAGCCAT ATGAAGAATT TGCAGCAAGA

451 500
NC_001463(gag720bp) GTCCAAGCAG CA..GGAGGG AGAAGCTGGA AAGCAGTAGA TTCTGTAATG
>AF015181 CTGCTAGAAG CAATAGATGC AGAAGCGGTT ACACAGCCCA TAAAAGAGTA

501 550
NC_001463(gag720bp) T.TCCAGC.A ACTGCAAACA GTAGCAATGC AGCATGGCCT CGTGTCTGAG
>AF015181 TCTAAAGCTA ACATTATCCT ATACAAATGC AGC.....CT CA.....

551 600
NC_001463(gag720bp) GACTTTGAAA GGCAGTTGGC ATATTATGCT ACTACCTGGA CAAGTAAAGA
>AF015181 GATTGTCAAA AGCAAATGG. AGAGAGTGTCT AGGACAAAGA...GTACA.A

601 650
NC_001463(gag720bp) CATACTAGAA GTATTGGCCA TGATGCCTGG AAATAGAGCT CAAAAGGAGT
>AF015181 CAGGCTAGT. GTAGAAAAAA AAATGCAAGC ATGT.....

651 700
NC_001463(gag720bp) TAATTCAAGG GAAATTAAAT GAAGAAGCAG AAAGGTGGAG AACGAATAAT
    
```

<410>


```

>AF015181 .....
... 701 736
NC_001463(gag720bp) CCACCACCTC CAGCAGGAGG AGGATTAACA GTGGAT
>AF015181 .....

PileUp

MSF: 1347 Type: N Check: 939 ..

Name: NC_001463(gag) (SEQ ID NO: 17)Len: 1347 Check: 6959 Weight: 0
Name: >AF015181 (SEQ ID NO: 18) Len: 1347 Check: 3980 Weight: 0

//

1 50
NC_001463(gag) ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGGAAAAG
>AF015181 .....

51 100
NC_001463(gag) AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAG
>AF015181 .....

101 150
NC_001463(gag) TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCCT
>AF015181 .....

151 200
NC_001463(gag) AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTTGGACAAA
>AF015181 .....

201 250
NC_001463(gag) AGTAAATTT AGGACAATAT GGCAGAAGGT GAAGAATCTA ACTCCTGAGG
>AF015181 .....

251 300
NC_001463(gag) AGAGTAACAA AAAAGACTTT ATGTCTTTGC AGGCCACATT AGCGGGTCTA
>AF015181 .....

301 350
NC_001463(gag) ATGTGTTGCC AAATGGGGAT GAGACCTGAG ACATTGCAAG ATGCAATGGC
>AF015181 .....

351 400
NC_001463(gag) TACAGTAATC ATGAAAGATG GTTACTGGA ACAAGAGGAA AAGAAGGAAG
>AF015181 .....

401 450
NC_001463(gag) ACAAAGAGA AAAGGAAGAG AGTGTCTTCC CAATAGTAGT GCAAGCAGCA
>AF015181 .....

451 500
NC_001463(gag) GGAGGGAGAA GCTGGAAGC AGTAGATTCT GTAATGTTCC AGCAACTGCA
>AF015181 .....GC TGTAGACTCT GTAATGTTCC AACAAATGCA

501 550
NC_001463(gag) AACAGTAGCA ATGCAGCATG GCCTCGTGTG TGAGGACTTT GAAAGGCAGT
>AF015181 AACAGTAGCA ATGCAGCATG GCCTCGTGTG CGAGGATTTT GAAAGACAGT

551 600
NC_001463(gag) TGGCATATTA TGCTACTACC TGGACAAGTA AAGACATACT AGAAGTATTG
>AF015181 TAGCATATTA TGCTACTACC TGGACAAGTA AAGACATACT AGAAGTATTG

```

NC_001463 (gag)	661	650
>AF015181	GCCATGATGC CTGGAAATAG AGCTCAAAAG GAGTTAATTC AAGGGAAATT	GCCATGATGC CTGGAAATAG GGCTCAGAAA GAACTTATTC AAGGGAAATT
NC_001463 (gag)	651	700
>AF015181	AAATGAAGAA GCAGAAAGGT GGAGAAGGAA TAATCCACCA CCTCCAGCAG	GAATGAAGAA GCAGACAGGT GGAGAAGGAA CAATCCACCA
NC_001463 (gag)	701	750
>AF015181	GAGGAGGATT AACAGTGGAT CAAATTATGG GGGTAGGACA AACAAATCAA	..GGAGGATT AACAGTGGAT CAAATTATGG GGGTAGGACA AACAAATCAA
NC_001463 (gag)	751	800
>AF015181	GCAGCAGCAC AAGCTAACAT GGATCAGGCA AGGCAAATAT GCCTGCAATG	GCAGCAGCAC AAGCTAACAT GGATCAGGCA AGACAAATAT GCCTACAATG
NC_001463 (gag)	801	850
>AF015181	GGTAATAAAT GCATTAAGAG CAGTAAGACA TATGGCGCAC AGGCCAGGGA	GGTAATAAAT GCCTTAAGAG CAGTAAGGCA TATGGCTCAT AGGCCAGGGA
NC_001463 (gag)	851	900
>AF015181	ATCCAATGCT AGTAAAGCAA AAAACGAATG AGCCATATGA AGATTTTGCA	ATCCAATGCT AGTAAAGCAA AAAACAAATG AGCCATATGA AGAATTTGCA
NC_001463 (gag)	901	950
>AF015181	GCAAGACTGC TAGAAGCAAT AGATGCAGAG CCAGTTACAC AGCCTATAAA	GCAAGACTGC TAGAAGCAAT AGATGCAGAA GCGGTTACAC AGCCCATAAA
NC_001463 (gag)	951	1000
>AF015181	AGATTATCTA AAGCTAACAC TATCTTATAC AAATGCATCA GCAGATTGTC	AGAGTATCTA AAGCTAACAT TATCCTATAC AAATGCAGCC TCAGATTGTC
NC_001463 (gag)	1001	1050
>AF015181	AGAAGCAAAT GGATAGAACA CTAGGACAAA GAGTACAACA AGCTAGTGTA	AAAAGCAAAT GGAGAGAGTG CTAGGACAAA GAGTACAACA GGCTAGTGTA
NC_001463 (gag)	1051	1100
>AF015181	GAAGAAAAAA TGCAAGCATG TAGAGATGTG GGATCAGAAG GGTTCAAAAT	GAAAAAAA TGCAAGCATG T.....
NC_001463 (gag)	1101	1150
>AF015181	GCAATTGTTA GCACAAGCAT TAAGCCAGG AAAAGGAAAA GGAATGGAC
NC_001463 (gag)	1151	1200
>AF015181	AGCCACAAAG GTGTTACAAC TGTGGAAAAC CGGGACATCA AGCAAGGCAA
NC_001463 (gag)	1201	1250
>AF015181	TGTAGACAAG GAATCATATG TCACAACGTG GGAAAGAGAG GACATATGCA
NC_001463 (gag)	1251	1300
>AF015181	AAAAGAATGC AGAGGAAAGA GAGACATAAG GGGAAAACAG CAGGGAAACG
NC_001463 (gag)	1301	1347
>AF015181	GGAGGAGGGG GATACGTGTG GTGCCGTCCG CTCCTCCTAT GGAATAA

표 8

```

FileUp
MSF: 727  Type: N  Check: 1231  ..

Name: NC_001463 (gag720bp) (SEQ ID NO: 19) Len: 727 Check: 1714 Weight: 0
Name: >AF402664 (SEQ ID NO: 20) Len: 727 Check: 1659 Weight: 0
Name: >AF402665 (SEQ ID NO: 21) Len: 727 Check: 331 Weight: 0
Name: >AF402666 (SEQ ID NO: 22) Len: 727 Check: 7190 Weight: 0
Name: >AF402667 (SEQ ID NO: 23) Len: 727 Check: 9833 Weight: 0
Name: >AF402668 (SEQ ID NO: 24) Len: 727 Check: 504 Weight: 0

//

1 50
NC_001463 (gag720bp) ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGAAAAG
>AF402664 .....TC AAGCAGCAGG .AGGGAGAAG CTGAAAAGCA GTAGACTCAG
>AF402665 .....GC AAGCAGCAGG .AGGGAGAAG CTGAAAAGCA GTAGACTCAG
>AF402666 .....GC AAGCAGCAGG .AGGGAGAAG CTGAAAAGCA GTAGACTCAG
>AF402667 .....GC AAGCAGCAGG .AGGGAGAAG CTGAAAAGCA GTAGACTCAG
>AF402668 .....GC AAGCAGCAGG .AGGGAGAAG CTGAAAAGCA GTAGACTCAG

51 100
NC_001463 (gag720bp) AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAAG
>AF402664 TGATGTTCCA GCAACTGCAA AATGTAGCAA TGCAGCATGG CCTCGTGTCC
>AF402665 TGATGTTCCA GCAACTGCAA AATGTAGCAA TGCAGCATGG CCTCGTGTCC
>AF402666 TGATGTTCCA GCAACTGCAA AATGTAGCAA TGCAGCATGG CCTCGTGTCC
>AF402667 TGATGTTCCA GCAACTGCAA AATGTAGCAA TGCAGCATGG CCTCGTGTCC
>AF402668 TGATGTTCCA GCAACTGCAA AATGTAGCAA TGCAGCATGG CCTCGTGTCC

101 150
NC_001463 (gag720bp) TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCTT
>AF402664 GAGGATTTTG AAAGG..CAG TTAGTATATT ATGCTACTAC CTGGACAAGT
>AF402665 GAGGATTTTG AAAGG..CAG TTAGCATATT ATGCTACTAC CTGGACAAGT
>AF402666 GAGGATTTTG AAAGG..CAG TTGGCATATT ATGCTACTAC CTGGACAAGT
>AF402667 GAGGATTTTG AAAGG..CAG TTAGCATATT ATGCTACTAC CTGGACAAGT
>AF402668 GAGGATTTTG AAAGG..CAG TTAGCATATT ATGCTACTAC CTGGACAAGT

151 200
NC_001463 (gag720bp) AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTTGGACAAA
>AF402664 AAAGA..TAT ATTAGAAGTA TTGG..CCAT GATG.....C CTGGAAATAG
>AF402665 AAAGA..TAT ATTAGAAGTA TTGG..CCAT GATG.....C CTGGAAATAG
>AF402666 AAAGA..TAT ATTAGAAGTA TTGG..CCAT GATG.....C CTGGAAACAG
>AF402667 AAAGA..TAT ATTAGAAGTA TTGG..CCAT GATG.....C CTGGAAATAG
>AF402668 AAAGA..TAT ATTAGAAGTA TTGG..CCAT GATG.....C CTGGAAATAG

201 250
NC_001463 (gag720bp) AGTAAAAATTT AGGACAATAT GGCAGAAGGT GAAGAATCTA ACTCCTGAGG
>AF402664 AGCTCAAAAA GAGTTAATTC AAGGGAAATT GAATGAGGAA GCAGAAAAGGT
>AF402665 AGCTCAAAAA GAGTTAATTC AAGGGAAATT GAATGAGGAA GCAGAAAAGGT
>AF402666 AGCTCAAAAA GAGTTAATTC AAGGGAAATT GAATAAGGAA GCAGAAAAGGT
>AF402667 AGCTCAAAAA GAGTTAATTC AAGGGAAATT GAATGAGGAA GCAGAAAAGGT
>AF402668 AGCTCAAAAA GAGTTAATTC AAGGGAAATT GAATGAGGAA GCAGAAAAGGT

251 300
NC_001463 (gag720bp) AGAGTAACAA AAAAGACTTT ATGTCPTTGC AGGCCACATT AGCGGCTTA
>AF402664 GGAG..AAGAA ATAATCCACC ACCTCAA.GC AGGCG.....GAGGATTAA
>AF402665 GGAG..AAGGA ATAATCCACC ACCTCAA.GC AGGCG.....GAGGATTAA
>AF402666 GGAG..AAGGA ATAATCCACC ACCTCAA.GC ACAAG.....GAGGATTAA
>AF402667 GGAG..AAGGA ATAATCCACC ACCTCAA.GC AGGCG.....GAGGATTAA
>AF402668 GGAG..AAGGA ATAATCCACC ACCTCAA.GC AGGCG.....GAGGATTAA

301 350
NC_001463 (gag720bp) ATGTGTTGCC AA..ATGGGG ATGAGACCTG AGACATTGCA AGATGCAATG
    
```

<413>

>AF402664	CAGTGGATCA	AATTATGGGG	GTAGGACAAA	CAAATCAAGC	AGCGGCACAG
>AF402665	CAGTGGATCA	AATTATGGGG	GTAGGACAAA	CAAATCAAGC	AGCGGCACAG
>AF402666	CAGTGGATCA	AATTATGGGG	GTAGGACAAA	CAAATCAGGC	AGCGGCACAG
>AF402667	CAGTGGATCA	AATTATGGGG	GTAGGACAAA	CAAATCAAGC	AGCGGCACAG
>AF402668	CAGTGGATCA	AATTATGGGG	GTAGGACAAA	CAAATCAAGC	AGCGGCACAG
	351				400
NC_001463 (gag720bp)	GCTA.CAGTA	ATCATGAAAG	ATGGGTT..A	CTGGAACAAG	.AGGAAAAGA
>AF402664	GCTAACATGG	ATCAGGCAAG	ACAAATATGT	CTGCAATGGG	TAATAACAGC
>AF402665	GCTAACATGG	ATCAGGCAAG	ACAAATATGC	CTGCAATGGG	TAATAACAGC
>AF402666	GCTAACATGG	ATCAGGCAAG	ACAAATATGC	CTGCAATGGG	TAATAACAGC
>AF402667	GCTAACATGG	ATCAGGCAAG	ACAAATATGC	CTGCAATGGG	TAATAACAGC
>AF402668	GCTAACATGG	ATCAGGCAAG	ACAAATATGC	CTGCAATGGG	TAATAACAGC
	401				450
NC_001463 (gag720bp)	AGGAAGACAA	AAGAGAAAAG	GAAGAGAGTG	TCTTCCCAAT	AGTAGTGCAA
>AF402664	ACTAAGAGCA	GTGAGACATA	TGGCTCACAA	ACCAGGGAAT	CCAA.TGCTA
>AF402665	ACTAAGAGCA	GTGAGACATA	TGGCTCACAA	ACCAGGGAAT	CCAA.TGCTA
>AF402666	ACTAAGAGCA	GTGAGACATA	TGGCTCACAA	ACCAGGGAAT	CCAA.TGCTA
>AF402667	ACTAAGAGCA	GTGAGACATA	TGGCTCACAA	ACCAGGGAAT	CCAA.TGCTA
>AF402668	ACTAAGAGCA	GTGAGACATA	TGGCTCACAA	ACCAGGGAAT	CCAA.TGCTA
	451				500
NC_001463 (gag720bp)	GCAGCAGGAG	GGAGAAGCTG	GAAAGCAGTA	GATTCTGTAA	TGTTCACGCA
>AF402664	GTAAGCAAAA	AAACAAATGA	GTCCATATGAA	GATTTTGCCG	CAAGACTGCT
>AF402665	GTAAGCAAAA	AGACAAATGA	GTCCATATGAA	GATTTTGCCG	CAAGACTGCT
>AF402666	GTAAGCAAAA	AGACAAATGA	GTCCATATGAA	GATTTTGCCG	CAAGACTGCT
>AF402667	GTAAGCAAAA	AGACAAATGA	GTCCATATGAA	GATTTTGCCG	CAAGACTGCT
>AF402668	GTAAGCAAAA	AGACAAATGA	GTCCATATGAA	AAATTTTCAG	CAAGACTCCT
	501				550
NC_001463 (gag720bp)	ACTGCAAAACA	GTAGCA.ATG	CAGCATGGCC	TCGTGTCTGA	GGACTTTGAA
>AF402664	AGAAGCAATA	GATGCAGAAC	CAGTTACACA	GCAAATAAAA	GAATATTTAA
>AF402665	AGAAGCAATA	GATGCAGAAC	CAGTTACACA	GCAAATAAAA	GAATATTTAA
>AF402666	AGAAGCAATA	GATGCAGAAC	CAGTTACACA	GCAAATAAAA	GAATATTTAA
>AF402667	AGAAGCAATA	GATGCAGAAC	CAGTTACACA	GCAAATAAAA	GAATATTTAA
>AF402668	AGAAGCAATA	GATGCAGAAC	CAGTTACACA	GCCTATAAAA	GAATATTTAA
	551				600
NC_001463 (gag720bp)	AGGCAGTTGG	CATATTATGC	TACTACCTGG	ACAAGTAAAG	ACATACTAGA
>AF402664	AGTTA.....	.ACATTATCT	TAC.ACAAAT	GCATCCTCAG	ACTGTCAGAA
>AF402665	AGTTA.....	.ACATTATCT	TAC.ACAAAT	GCATCCTCAG	ACTGTCAGAA
>AF402666	AGTTA.....	.ACATTATCT	TAC.ACAAAT	GCATCCTCAG	ACTGTCAGAA
>AF402667
>AF402668	AGTTA.....	.ACATTATCT	TAC.ACAAAT	GCATCCTCAG	ACTGTCAGAA
	601				650
NC_001463 (gag720bp)	AGTATTGGCC	ATGATGCCTG	GAAATAGAGC	TCAAAAGGAG	TTAATTCAAG
>AF402664	ACAGATGGAT	AGAGTACTAG	GACAGAGAGT	GCAACAAGCT	AGTGTGGAAG
>AF402665	ACAAATGGAT	AGAATACTAG	GACAGAGAGT	GCAACAAGCT	AGTGTGGAAG
>AF402666	ACAAATGGAT	AGAGTACTAG	GACAGAGAGT	GCAACAAGCT	AGTGTGGAAG
>AF402667
>AF402668	ACAAATGGAT	AGAGTACTAG	GACAGAGAGT	GCAACAAGCT	AGTGTGGAAG
	651				700
NC_001463 (gag720bp)	GGAAATTAAG	TGAAGAAGCA	GAAAGGTGGA	GAAGGAATAA	TCCACCACCT
>AF402664	AAAAAATGCA	AGCAT..GCA	GAGATGTGGG	ATCAGAAGGA	TTCAGAATGC
>AF402665	AAAAAATGCA	AGCAT..GCA	GAGATGTGGG	ATCAGAAGGG	TTCAGAATGC
>AF402666	AAAAAATGCA	AGCAT..GCA	GAGATGTGGG	ATCAGAAGG.
>AF402667
>AF402668	AAAAAATGCA	AGCAT..GCA	GAGATGTGGG	ATCAGAAGGA	TTCAGAATGC
	701		727		
NC_001463 (gag720bp)	CCAGCAGGAG	GAGGATTAAC	AGTGGAT		
>AF402664		
>AF402665		
>AF402666		
>AF402667		

```

>AF402668 .....

FileUp
MSF: 1347 Type: N Check: 5320 ..

Name: NC_001463 (gag) (SEQ ID NO: 25) Len: 1347 Check: 6959 Weight: 0
Name: >AF402664 (SEQ ID NO: 26) Len: 1347 Check: 1590 Weight: 0
Name: >AF402665 (SEQ ID NO: 27) Len: 1347 Check: 9222 Weight: 0
Name: >AF402666 (SEQ ID NO: 28) Len: 1347 Check: 4950 Weight: 0
Name: >AF402667 (SEQ ID NO: 29) Len: 1347 Check: 3156 Weight: 0
Name: >AF402668 (SEQ ID NO: 30) Len: 1347 Check: 9443 Weight: 0

//

NC_001463 (gag) 1 50
>AF402664 ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGGAAAAG
>AF402665 .....
>AF402666 .....
>AF402667 .....
>AF402668 .....

NC_001463 (gag) 51 100
>AF402664 AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAAAG
>AF402665 .....
>AF402666 .....
>AF402667 .....
>AF402668 .....

NC_001463 (gag) 101 150
>AF402664 TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCTT
>AF402665 .....
>AF402666 .....
>AF402667 .....
>AF402668 .....

NC_001463 (gag) 151 200
>AF402664 AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTTGGACAAA
>AF402665 .....
>AF402666 .....
>AF402667 .....
>AF402668 .....

NC_001463 (gag) 201 250
>AF402664 AGTAAAATTT AGGACAATAT GGCAGAAGGT GAAGAATCTA ACTCCTGAGG
>AF402665 .....
>AF402666 .....
>AF402667 .....
>AF402668 .....

NC_001463 (gag) 251 300
>AF402664 AGAGTAACAA AAAAGACTTT ATGTCCTTGC AGGCCACATT AGCGGGTCTA
>AF402665 .....
>AF402666 .....

```

>AF402667
>AF402668
	301				350
NC_001463 (gag)	ATGTGTTGCC	AAATGGGGAT	GAGACCTGAG	ACATTGCAAG	ATGCAATGGC
>AF402664
>AF402665
>AF402666
>AF402667
>AF402668
	351				400
NC_001463 (gag)	TACAGTAATC	ATGAAAGATG	GGTTACTGGA	ACAAGAGGAA	AAGAAGGAAG
>AF402664
>AF402665
>AF402666
>AF402667
>AF402668
	401				450
NC_001463 (gag)	ACAAAAGAGA	AAAGGAAGAG	AGTGTCTTCC	CAATAGTAGT	GCAAGCAGCA
>AF402664	TCAAGCAGCA
>AF402665	GCAAGCAGCA
>AF402666	GCAAGCAGCA
>AF402667	GCAAGCAGCA
>AF402668	GCAAGCAGCA
	451				500
NC_001463 (gag)	GGAGGGAGAA	GCTGGAAAGC	AGTAGATTCT	GTAATGTTCC	AGCAACTGCA
>AF402664	GGAGGGAGAA	GCTGGAAAGC	AGTAGACTCA	GTGATGTTCC	AGCAACTGCA
>AF402665	GGAGGGAGAA	GCTGGAAAGC	AGTAGACTCA	GTGATGTTCC	AGCAACTGCA
>AF402666	GGAGGGAGAA	GCTGGAAAGC	AGTAGACTCA	GTGATGTTCC	AGCAACTGCA
>AF402667	GGAGGGAGAA	GCTGGAAAGC	AGTAGACTCA	GTGATGTTCC	AGCAACTGCA
>AF402668	GGAGGGAGAA	GCTGGAAAGC	AGTAGACTCA	GTGATGTTCC	AGCAACTGCA
	501				550
NC_001463 (gag)	AACAGTAGCA	ATGCAGCATG	GCCTCGTGTC	TGAGGACTTT	GAAAGGCAGT
>AF402664	AAATGTAGCA	ATGCAGCATG	GCCTCGTGTC	CGAGGATTTT	GAAAGGCAGT
>AF402665	AAATGTAGCA	ATGCAGCATG	GCCTCGTGTC	CGAGGATTTT	GAAAGGCAGT
>AF402666	AAATGTAGCA	ATGCAGCATG	GCCTCGTGTC	CGAGGATTTT	GAAAGGCAGT
>AF402667	AAATGTAGCA	ATGCAGCATG	GCCTCGTGTC	CGAGGATTTT	GAAAGGCAGT
>AF402668	AAATGTAGCA	ATGCAGCATG	GCCTCGTGTC	CGAGGATTTT	GAAAGGCAGT
	551				600
NC_001463 (gag)	TGGCATATTA	TGCTACTACC	TGGACAAGTA	AAGACATACT	AGAAGTATTG
>AF402664	TAGTATATTA	TGCTACTACC	TGGACAAGTA	AAGATATATT	AGAAGTATTG
>AF402665	TAGCATATTA	TGCTACTACC	TGGACAAGTA	AAGATATATT	AGAAGTATTG
>AF402666	TGGCATATTA	TGCTACTACC	TGGACAAGTA	AAGATATATT	AGAAGTATTG
>AF402667	TAGCATATTA	TGCTACTACC	TGGACAAGTA	AAGATATATT	AGAAGTATTG
>AF402668	TAGCATATTA	TGCTACTACC	TGGACAAGTA	AAGATATATT	AGAAGTATTG
	601				650
NC_001463 (gag)	GCCATGATGC	CTGGAAATAG	AGCTCAAAA	GAGTTAATTC	AAGGGAAATT
>AF402664	GCCATGATGC	CTGGAAATAG	AGCTCAAAA	GAGTTAATTC	AAGGGAAATT
>AF402665	GCCATGATGC	CTGGAAATAG	AGCTCAAAA	GAGTTAATTC	AAGGGAAATT
>AF402666	GCCATGATGC	CTGGAAACAG	AGCTCAAAA	GAGTTAATTC	AAGGGAAATT
>AF402667	GCCATGATGC	CTGGAAATAG	AGCTCAAAA	GAGTTAATTC	AAGGGAAATT
>AF402668	GCCATGATGC	CTGGAAATAG	AGCTCAAAA	GAGTTAATTC	AAGGGAAATT
	651				700

NC_001463 (gag)	AAATGAAGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCCAGCAG
>AF402664	GAATGAGGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCAAGCAG
>AF402665	GAATGAGGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCAAGCAG
>AF402666	GAATAAGGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCAAGCAC
>AF402667	GAATGAGGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCAAGCAG
>AF402668	GAATGAGGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCAAGCAG
	701				750
NC_001463 (gag)	GAGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAA
>AF402664	GCGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAA
>AF402665	GCGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAA
>AF402666	AAGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAG
>AF402667	GCGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAA
>AF402668	GCGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAA
	751				800
NC_001463 (gag)	GCAGCAGCAC	AAGCTAACAT	GGATCAGGCA	AGGCAAATAT	GCCTGCAATG
>AF402664	GCAGCGGCAC	AGGCTAACAT	GGATCAGGCA	AGACAAATAT	GTCTGCAATG
>AF402665	GCAGCGGCAC	AGGCTAACAT	GGATCAGGCA	AGACAAATAT	GCCTGCAATG
>AF402666	GCAGCGGCAC	AGGCTAACAT	GGATCAGGCA	AGACAAATAT	GCCTGCAATG
>AF402667	GCAGCGGCAC	AGGCTAACAT	GGATCAGGCA	AGACAAATAT	GCCTGCAATG
>AF402668	GCAGCGGCAC	AGGCTAACAT	GGATCAGGCA	AGACAAATAT	GCCTGCAATG
	801				850
NC_001463 (gag)	GGTAATAAAT	GCATTAAGAG	CAGTAAGACA	TATGGCGCAC	AGGCCAGGGA
>AF402664	GGTAATAACA	GCACTAAGAG	CAGTGAGACA	TATGGCTCAC	AAACCAGGGA
>AF402665	GGTAATAACA	GCACTAAGAG	CAGTGAGACA	TATGGCTCAC	AAACCAGGGA
>AF402666	GGTAATAACA	GCACTAAGAG	CAGTGAGACA	TATGGCTCAC	AAACCAGGGA
>AF402667	GGTAATAACA	GCACTAAGAG	CAGTGAGACA	TATGGCTCAC	AAACCAGGGA
>AF402668	GGTAATAACA	GCACTAAGAG	CAGTGAGACA	TATGGCTCAC	AAACCAGGGA
	851				900
NC_001463 (gag)	ATCCCAATGCT	AGTAAAGCAA	AAAACGAATG	AGCCATATGA	AGATTTTGCA
>AF402664	ATCCCAATGCT	AGTAAAGCAA	AAAACAAATG	AGTCATATGA	AGATTTTGCC
>AF402665	ATCCCAATGCT	AGTAAAGCAA	AAGACAAATG	AGTCATATGA	AGATTTTGCC
>AF402666	ATCCCAATGCT	AGTAAAGCAA	AAGACAAATG	AGTCATATGA	AGATTTTGCC
>AF402667	ATCCCAATGCT	AGTAAAGCAA	AAGACAAATG	AGTCATATGA	AGATTTTGCC
>AF402668	ATCCCAATGCT	AGTAAAGCAA	AAGACAAATG	AGTCATATGA	AAAAITTTCA
	901				950
NC_001463 (gag)	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAG	CCAGTTACAC	AGCCTATAAA
>AF402664	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAA	CCAGTTACAC	AGCAAATAAA
>AF402665	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAA	CCAGTTACAC	AGCAAATAAA
>AF402666	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAA	CCAGTTACAC	AGCAAATAAA
>AF402667	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAA	CCAGTTACAC	AGCAAATAAA
>AF402668	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAA	CCAGTTACAC	AGCCTATAAA
	951				1000
NC_001463 (gag)	AGAATTATCTA	AAGCTAACAC	TATCTTATAC	AAATGCATCA	GCAGATTGTC
>AF402664	AGAATATTTA	AAGTTAACAT	TATCTTACAC	AAATGCATCC	TCAGACTGTC
>AF402665	AGAATATTTA	AAGTTAACAT	TATCTTACAC	AAATGCATCC	TCAGACTGTC
>AF402666	AGAATATTTA	AAGTTAACAT	TATCTTACAC	AAATGCATCC	TCAGACTGTC
>AF402667	.GAATATTTA	A.....
>AF402668	AGAATATTTA	AAGTTAACAT	TATCTTACAC	AAATGCATCC	TCAGACTGTC
	1001				1050
NC_001463 (gag)	AGAAGCAAAT	GGATAGAACA	CTAGGACAAA	GAGTACAACA	AGCTAGTGTA
>AF402664	AGAAACAGAT	GGATAGAGTA	CTAGGACAGA	GAGTGCAACA	AGCTAGTGTG
>AF402665	AAAAACAAAT	GGATAGAATA	CTAGGACAGA	GAGTGCAACA	AGCTAGTGTG
>AF402666	AGAAACAAAT	GGATAGAGTA	CTAGGACAGA	GAGTGCAACA	AGCTAGTGTG

>AF402667
>AF402668	AAAAACAAAT	GGATAGAGTA	CTAGGACAGA	GAGTGCAACA	AGCTAGTGTG
	1051				1100
NC_001463 (gag)	GAAGAAAAAA	TGCAAGCATG	TAGAGATGTG	GGATCAGAAG	GGTTCAAAT
>AF402664	GAAGAAAAAA	TGCAAGCATG	CAGAGATGTG	GGATCAGAAG	GATTCAGAAT
>AF402665	GAAGAAAAAA	TGCAAGCATG	CAGAGATGTG	GGATCAGAAG	GGTTCAGAAT
>AF402666	GAAGAAAAAA	TGCAAGCATG	CAGAGATGTG	GGATCAGAAG	G.....
>AF402667
>AF402668	GAAGAAAAAA	TGCAAGCATG	CAGAGATGTG	GGATCAGAAG	GATTCAGAAT
	1101				1150
NC_001463 (gag)	GCAATTGTTA	GCACAAGCAT	TAAGGCCAGG	AAAAGGAAAA	GGGAATGGAC
>AF402664	GC.....
>AF402665	GC.....
>AF402666
>AF402667
>AF402668	GC.....
	1151				1200
NC_001463 (gag)	AGCCACAAAG	GTGTTACAAC	TGTGGAAAAC	CGGGACATCA	AGCAAGGCAA
>AF402664
>AF402665
>AF402666
>AF402667
>AF402668
	1201				1250
NC_001463 (gag)	TGTAGACAAG	GAATCATATG	TCACAACGTG	GGAAAGAGAG	GACATATGCA
>AF402664
>AF402665
>AF402666
>AF402667
>AF402668
	1251				1300
NC_001463 (gag)	AAAAGAATGC	AGAGGAAAGA	GAGACATAAG	GGGAAAACAG	CAGGGAAACG
>AF402664
>AF402665
>AF402666
>AF402667
>AF402668
	1301				1347
NC_001463 (gag)	GGAGGAGGGG	GATACGTGTG	GTGCCGTCCG	CTCCTCCTAT	GGAATAA
>AF402664
>AF402665
>AF402666
>AF402667
>AF402668

표 9

```

FileUp
MSF: 742 Type: N Check: 6523 ..
Name: NC_001463 (gag720bp) (SEQ ID NO: 31) Len: 742 Check: 3818 Weight: 0
Name: >AJ305040 (SEQ ID NO: 32) Len: 742 Check: 1263 Weight: 0
Name: >AJ305041 (SEQ ID NO: 33) Len: 742 Check: 9126 Weight: 0
Name: >AJ305042 (SEQ ID NO: 34) Len: 742 Check: 2316 Weight: 0

//

NC_001463 (gag720bp) 1 50
>AJ305040 .....GCAG TCGATGCTGT AATGTTCCAG CAAATGCAAA CAGTAGCCAT
>AJ305041 .....GCAG TAGACTCAGT AATGTTCCAG CAACTGCAAA CAGTAGCAAT
>AJ305042 .....GCAG TCGATGCTGT AATGTTCCAG CAAATGCAAA CAGTAGCCAT

NC_001463 (gag720bp) 51 100
>AJ305040 AGAGATTATC CTGAGCTCGA AAAATGTATC AAGCATGCAT GCAAGATAAA
>AJ305041 GCAGCATGGT CTTGTGTCTG AGGACTTTGA AAGGCAGTTA GCAT.ATTGT
>AJ305042 GCAGCATGGC CTCGTGTCCG AGGATTTTGA AAGGCAGTTG GCAT.ATTAT
>AJ305042 GCAGCATGGT CTTGTGTCTG AGGACTTTGA AAGGCAGTTA GCAT.ATTAT

NC_001463 (gag720bp) 101 150
>AJ305040 AGTTCGACTC AGAGGGG..A GCACITGACA GAAGGAAATT GTTTATGGTG
>AJ305041 GCTACTACCT GGACAAGTAA AGATATATTA GAAGTA..TT GGCCATGATG
>AJ305042 GCTACTACCT GGACGAGTAA AGACATACTA GAAGTA..TT GGCCATGATG
>AJ305042 GCTACTACCT GGACAAGTAA AGATATATTA GAAGTA..TT GGCCATGATG

NC_001463 (gag720bp) 151 200
>AJ305040 CCTTAAAACA TTAGATTACA TGTTTGAGGA CCATAAAGAG GAACCTTGGA
>AJ305041 CCTGGAAATA G.AGCTCAAA AA...GAGTT AATTCAAG.G AAAATTAATC
>AJ305042 CCTGGAAACA G.AGCTCAAA AG...GAGTT AATTCAAG.G GAAATTAATC
>AJ305042 CCTGGAAATA G.AGCTCAAA AA...GAGTT AATTCAAG.G AAAATTAATC

NC_001463 (gag720bp) 201 250
>AJ305040 CAAAAGTAAA ATTTAGGACA ATATGGCAGA AGGTGAAGAA TCTAACTCCT
>AJ305041 GAGGAAGCAG AA..AGGTGG AGAAGGAATA A..TCCACCC CCTCCACAGC
>AJ305041 GAGGAGGCAG AA..AGGTGG AGAAGACATA A..TCCACCC CCTCCGCGG
>AJ305042 GAGGAAGCAG AA..AGGTGG AGAAGGAATA A..TCCACCC CCTCCACAGC

NC_001463 (gag720bp) 251 300
>AJ305040 GAGGAG.AGT AACAAAAAAG .ACTTTATGT CTTTGCAAGC CACATTAGCG
>AJ305041 GAGGGGGATT AACAGTGGAT CAAATTATGG GGAT..AGGA CAAACAAATC
>AJ305041 GAGGAGGATT AACAGTGGAT CAAATTATGG GGGT..AGGA CAAACAAATC
>AJ305042 GAGGGGGATT AACAGTGGAT CAAATTATGG GGAT..AGGA CAAACAAATC

NC_001463 (gag720bp) 301 350
>AJ305040 GGTCTAATGT GTTGCCAAAT GGGATGAGA CCTGAGACAT .....TGCAA
>AJ305041 AAGCAGCAGC ACAAGCTAAC ATGGATCAGG CAAGACACAT ATGCCTGCAA
>AJ305041 AAGCAGCAGC ACAAGCTAAC ATGGATCAGG CAAGACAAAT ATGCCTGCAA
>AJ305042 AAGCAGCAGC ACAAGCTAAC ATGGATCAGG CAAGACACAT ATGCCTGCAA

NC_001463 (gag720bp) 351 400
>AJ305040 GATGCAATGG CTACAGTAAT ..CA.TGAAA GATGGGTTAC TGGAACAAGA
>AJ305041 TGGGTAATAA CAGCATTAAAG AGCAGTAAGA CATATGGCTC ACAGACCAGG
>AJ305041 TGGGTAATAA CAGCATTAAAG AGCAGTGAAG TATATGACTC ACAACCCAGG
>AJ305042 TGGGTAATAA CAGCATTAAAG AGCAGTAAGA CATATGGCTC ACAGACCAGG

NC_001463 (gag720bp) 401 450
>AJ305040 GGA...AAG A.AGGAAGAC AAAAGAGAAA AGGAAGAGAG T..GTCTTCC
>AJ305041 GAATCCAATG CTCGTAAAAC AAAAAACAAA TGAGCCATAT GAAGAGTTTG
>AJ305041 GAATCCAATG CTAGTAAAAC AAAAAACAAA TGAAGCATAT GAAGAGTTTA
    
```

```

>AJ305042 GAATCCAATG CTCGTAAAAC AAAAAACAAA TGAGCCATAT GAAGAGTTTG
NC_001463 (gag720bp) 451 500
>AJ305040 CAATAGTAGT GCAAGCAGCA GGAG..GGAG AAGCTGGAAA GCAGTAGATT
>AJ305041 CAGCAAAACT ATTAGAAGCA ATAGATGCAG AACCAGTAAC ACAGCCCATATA
>AJ305042 CAGCGAGACT GCTAGAAGCA ATAGATGCAG AGCCAGTAAC ACAGCCCATATA
NC_001463 (gag720bp) 501 550
>AJ305040 CTGTAATGTT CCAGCAACTG CAAACAGTAG CAATGCAGCA TGGCCTCGTG
>AJ305041 AAAGACTAT..CTAAAGTT..AACATTAT CT.TATACAA ATGCCTC...
>AJ305042 AAAGACTAT..CTAAAGTT..AACATTAT CT.TATACAA ATGCCTC...
NC_001463 (gag720bp) 551 600
>AJ305040 TCTGAGGACT TTGAAAGGCA GTTGGCATAT TATGCTACTA CCTGGACAAG
>AJ305041 .CTCAG.ACT GTCAAAAGCA AATGG.ATAG AGTGCTGGGA CAAAG...AG
>AJ305042 .CTCAG.ACT GTCAAAAGCA AATGG.ATAG AGTACTAGGA CAAAG...AG
NC_001463 (gag720bp) 601 650
>AJ305040 TAAAGACATA CTAGAAGTAT TGGCCATGAT GCCTGGAAAT AGAGCTCAA
>AJ305041 TGCA.ACAAG CTAGT.GTAG ACGAGAAAAT GCAA.....
>AJ305042 TGCA.ACAAG CTAGT.GTAG ACGAGAAAGAT GCAA.....
NC_001463 (gag720bp) 651 700
>AJ305040 AGGAGTTAAT TCAAGGGAAA TTAATGAAG AAGCAGAAAG GTGGAGAAGG
>AJ305041 .....
>AJ305042 .....
NC_001463 (gag720bp) 701 742
>AJ305040 AATAATCCAC CACCTCCAGC AGGAGGAGGA TTAACAGTGG AT
>AJ305041 .....
>AJ305042 .....

-PileUp
MSF: 1347 Type: N Check: 9510 ..

Name: NC_001463(gag) (SEQ ID NO: 35) Len: 1347 Check: 6959 Weight: 0
Name: >AJ305040 (SEQ ID NO: 36) Len: 1347 Check: 1930 Weight: 0
Name: >AJ305041 (SEQ ID NO: 37) Len: 1347 Check: 7682 Weight: 0
Name: >AJ305042 (SEQ ID NO: 38) Len: 1347 Check: 2939 Weight: 0

//

NC_001463 (gag) 1 50
>AJ305040 ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGGAAAAG
>AJ305041 .....
>AJ305042 .....

NC_001463 (gag) 51 100
>AJ305040 AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAAAG
>AJ305041 .....
>AJ305042 .....

NC_001463 (gag) 101 150
>AJ305040 TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCCTT
>AJ305041 .....

```

>AJ305041
>AJ305042
	151				200
NC_001463 (gag)	AAAACATTAG	ATTACATGTT	TGAGGACCAT	AAAGAGGAAC	CTTGGACAAA
>AJ305040
>AJ305041
>AJ305042
	201				250
NC_001463 (gag)	ACTAAAATTT	AGGACAATAT	GGCAGAAGGT	GAAGAATCTA	ACTCCTGAGG
>AJ305040
>AJ305041
>AJ305042
	251				300
NC_001463 (gag)	AGAGTAACAA	AAAAGACTTT	ATGTCCTTGC	AGGCCACATT	AGCGGGTCTA
>AJ305040
>AJ305041
>AJ305042
	301				350
NC_001463 (gag)	ATGTGTTGCC	AAATGGGGAT	GAGACCTGAG	ACATTGCAAG	ATGCAATGGC
>AJ305040
>AJ305041
>AJ305042
	351				400
NC_001463 (gag)	TACAGTAATC	ATGAAAGATG	GGTTACTGGA	ACAAGAGGAA	AAGAAGGAAG
>AJ305040
>AJ305041
>AJ305042
	401				450
NC_001463 (gag)	ACAAAAGAGA	AAAGGAAGAG	AGTGTCTTCC	CAATAGTAGT	GCAAGCAGCA
>AJ305040
>AJ305041
>AJ305042
	451				500
NC_001463 (gag)	GGAGGGAGAA	GCTGGAAAGC	AGTAGATTCT	GTAATGTTCC	AGCAACTGCA
>AJ305040GC	AGTCGATGCT	GTAATGTTCC	AGCAAAATGCA	
>AJ305041GC	AGTAGACTCA	GTAATGTTCC	AGCAACTGCA	
>AJ305042GC	AGTCGATGCT	GTAATGTTCC	AGCAAAATGCA	
	501				550
NC_001463 (gag)	AACAGTAGCA	ATGCAGCATG	GCCTCGTGTC	TGAGGACTTT	GAAAGGCAGT
>AJ305040	AACAGTAGCC	ATGCAGCATG	GTCTTGTGTC	TGAGGACTTT	GAAAGGCAGT
>AJ305041	AACAGTAGCA	ATGCAGCATG	GCCTCGTGTC	CGAGGATTTT	GAAAGGCAGT
>AJ305042	AACAGTAGCC	ATGCAGCATG	GTCTTGTGTC	TGAGGACTTT	GAAAGGCAGT
	551				600
NC_001463 (gag)	TGGCATATTA	TGCTACTACC	TGGACAAGTA	AAGACATACT	AGAAGTATTG
>AJ305040	TAGCATATTG	TGCTACTACC	TGGACAAGTA	AAGATATATT	AGAAGTATTG
>AJ305041	TGGCATATTA	TGCTACTACC	TGGACGAGTA	AAGACATACT	AGAAGTATTG
>AJ305042	TAGCATATTA	TGCTACTACC	TGGACAAGTA	AAGATATATT	AGAAGTATTG
	601				650
NC_001463 (gag)	GCCATGATGC	CTGGAAATAG	AGCTCAAAAG	GAGTTAATTC	AAGGGAAATT
>AJ305040	GCCATGATGC	CTGGAAATAG	AGCTCAAAA	GAGTTAATTC	AAGGAAAATT

>AJ305041	GCCATGATGC	CTGGAAACAG	AGCTCAAAAG	GAGTTAATTC	AAGGGAAATT
>AJ305042	GCCATGATGC	CTGGAAATAG	AGCTCAAAAA	GAGTTAATTC	AAGGAAATT
NC_001463 (gag)	651				700
>AJ305040	AAATGAAGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCCAGCAG
>AJ305041	AAACGAGGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCG	CCTCCACAAG
>AJ305042	AAATGAAGAG	GCAGAAAGGT	GGAGAAGACA	TAATCCACCC	CCTCCGGCGG
NC_001463 (gag)	701				750
>AJ305040	AAGTGAAGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCG	CCTCCACAAG
>AJ305041	AAATGAAGAG	GCAGAAAGGT	GGAGAAGACA	TAATCCACCC	CCTCCGGCGG
>AJ305042	AAATGAGGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCG	CCTCCACAGG
NC_001463 (gag)	751				800
>AJ305040	GAGGAGGATT	AACAGTGGAT	CAAATTATGG	GGTAGGACA	AACAAATCAA
>AJ305041	GAGGGGGATT	AACAGTGGAT	CAAATTATGG	GGTAGGACA	AACAAATCAA
>AJ305042	GAGGAGGATT	AACAGTGGAT	CAAATTATGG	GGTAGGACA	AACAAATCAA
NC_001463 (gag)	801				850
>AJ305040	GCAGCAGCAC	AAGCTAACAT	GGATCAGGCA	AGGCAAATAT	GCCTGCAATG
>AJ305041	GCAGCAGCAC	AAGCTAACAT	GGATCAGGCA	AGACACATAT	GCCTGCAATG
>AJ305042	GCAGCAGCAC	AAGCTAACAT	GGATCAGGCA	AGACACATAT	GCCTGCAATG
NC_001463 (gag)	851				900
>AJ305040	GGAATAAAT	GCATTAAGAG	CAGTAAGACA	TATGGGCGAC	AGGCCAGGGA
>AJ305041	GGAATAACA	GCATTAAGAG	CAGTAAGACA	TATGGCTCAC	AGACCAGGGA
>AJ305042	GGAATAACA	GCATTAAGAG	CAGTAAGACA	TATGGCTCAC	AGACCAGGGA
NC_001463 (gag)	901				950
>AJ305040	ATCCAATGCT	AGTAAAGCAA	AAAACGAATG	AGCCATATGA	AGATTTTGCA
>AJ305041	ATCCAATGCT	CGTAAACAA	AAAACAAATG	AGCCATATGA	AGAGTTTGCA
>AJ305042	ATCCAATGCT	AGTAAACAA	AAAACAAATG	AAGCATATGA	AGAGTTTACA
NC_001463 (gag)	951				1000
>AJ305040	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAG	CCAGTTACAC	AGCCTATAAA
>AJ305041	GCAAAACTAT	TAGAAGCAAT	AGATGCAGAA	CCAGTAAACAC	AGCCCATAAA
>AJ305042	GCGAGACTGC	TAGAAGCAAT	AGATGCAGAG	CCAGTAAACAC	AGCCCACAAA
NC_001463 (gag)	1001				1050
>AJ305040	GCAAAACTAT	TAGAAGCAAT	AGATGCAGAA	CCAGTAAACAC	AGCCTATAAA
>AJ305041	GCGAGACTGC	TAGAAGCAAT	AGATGCAGAG	CCAGTAAACAC	AGCCCACAAA
>AJ305042	GCAAAACTAT	TAGAAGCAAT	AGATGCAGAA	CCAGTAAACAC	AGCCTATAAA
NC_001463 (gag)	1051				1100
>AJ305040	AGATTATCTA	AAGCTAACAC	TATCTTATAC	AAATGCATCA	GCAGATTGTC
>AJ305041	AGACTATCTA	AAGTTAACAT	TATCTTATAC	AAATGCGTCC	TCAGACTGTC
>AJ305042	AGAATATCTA	AAACTAACAT	TATCTTATAC	AAATGCATCC	TCAGACTGTC
NC_001463 (gag)	1101				1150
>AJ305040	GAAGAAAAA	TGCAAGCATG	TAGAGATGTG	GGATCAGAAG	GGTTCAAAAT
>AJ305041	GACGAGAAAA	TGCAA.....
>AJ305042	GAAGAAAAA	TGCAA.....
NC_001463 (gag)	1151				1200
>AJ305040	GCAATTGTTA	GCACAAGCAT	TAAGGCCAGG	AAAAGGAAAA	GGGAATGGAC
>AJ305041
>AJ305042

>AJ305041
>AJ305042
	1151				1200
NC_001463 (gag)	AGCCACAAAG	GTGTTACAAC	TGTGGAAAAC	CGGGACATCA	AGCAAGGCAA
>AJ305040
>AJ305041
>AJ305042
	1201				1250
NC_001463 (gag)	TGTAGACAAG	GAATCATATG	TCACAACGTG	GGAAAGAGAG	GACATATGCA
>AJ305040
>AJ305041
>AJ305042
	1251				1300
NC_001463 (gag)	AAAAGAATGC	AGAGCAAAGA	GAGACATAAG	GGGAAAACAG	CAGGGAAAACG
>AJ305040
>AJ305041
>AJ305042
	1301				1347
NC_001463 (gag)	GGAGGAGGGG	GATACGTGTG	GTGCCGTCCG	CTCCTCCTAT	GGAATAA
>AJ305040
>AJ305041
>AJ305042

<423>

표 10

```

FileUp
MSF: 728  Type: N  Check: 9403  ..
Name: NC_001463 (gag720bp) (SEQ ID NO: 39) Len: 728 Check: 5765 Weight: 0
Name: >AY047362 (SEQ ID NO: 40) Len: 728 Check: 3638 Weight: 0

//

1 50
NC_001463 (gag720bp) ATGGTGAGTC TAGATAGAGA CATGGCCAGG CAAGTCTCCG GGGGAAAAG
>AY047362 .....

51 100
NC_001463 (gag720bp) AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAAG
>AY047362 .....TAAA GATATATTAG AA.GTATTGG CCATG.ATGC CTGGAAATAG

101 150
NC_001463 (gag720bp) TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCTT
>AY047362 AGC...TCAA AAAGAGTTAA TTCA...AGG GAAATTGAAT GAAGAAGCAG

151 200
NC_001463 (gag720bp) AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTGGACAAA
>AY047362 AAAGGTGGAG AAGGAATAAT CCACCACCTC AAGCAGG..C GGAGGATTAA

201 250
NC_001463 (gag720bp) AGTAAATTT AGGACAATAT GCCAGAAGGT GAAG..AATC TAACTCCTGA
>AY047362 C..AG..TGG ATCAAATTAT GGGGGTAGGA CAAACAAATC AAGCAGCGGC

251 300
NC_001463 (gag720bp) GGAGAGTAAC AAAAAAGACT TTATGCTTTT GCAGGCCACA TTAGCGGGTC
>AY047362 ACAGGCTAAC ATGGATCAG..... GCAAGACAAA TATGC..CTG

301 350
NC_001463 (gag720bp) TAATGTGTTG CCAAATGGGG ATGAGACCTG AGACATTGCA AGATGCAATG
>AY047362 CAATGGGTAA TAACAGCACT AAGAGCAGTG AGACAT....A....TG

351 400
NC_001463 (gag720bp) GCTACAGTAA TCATGAAAGA TGGGTTACTG GAACAAGAGG AAAAGAAGGA
>AY047362 GCT.CACAAA CCAGGGA..A TCCGATGCT.....AGT.. AAAGCAA..A

401 450
NC_001463 (gag720bp) AGACAAAAGA GAAA.AGGAA GAGAGTGTCT TCCCAATAGT AGTGCAAGCA
>AY047362 AAACAAATGA GTCATATGAA GATTTTGCCG...CAAGACT GCTAGAAGCA

451 500
NC_001463 (gag720bp) GCAGGAGGGA GAAGCTGGAA AGCAGTAGAT TCTGTAATGT TCCAGCAACT
>AY047362 ATAG.ATGCA GAACCAAGTTA CAAAGCAAAT AAAAGAATAT TT....AAA

501 550
NC_001463 (gag720bp) GCAAACAGTA GCAATGCAGC ATGGCCTCGT GTCTGAGGAC TTTGAAAGGC
>AY047362 GTTAACATTA TCT.TACACA AATGCATC...CTCAG.AC TGTAAGAAAC

551 600
NC_001463 (gag720bp) AGTTGGCATA TTATGCTACT ACCTGGACAA GTA.AAGACA TACTAGAAGT
>AY047362 AGATGG.ATA GAGTACTAGG ACAGAGAGTG CAACAGCTA GTGTGGAAGA

601 650
NC_001463 (gag720bp) ATTG..GCCA TGATGCCTGG AAATAGAGCT CAAAAGGAGT TA..ATTCAA
>AY047362 AAAAATGCAA GCATGCAGAG ATGT.GGGAT CAGAAGGATT CAGAATGC..

651 700
NC_001463 (gag720bp) GGGAAATTAA ATGAAGAAGC AGAAAGGTGG AGAAGGAATA ATCCACCACC
    
```

```

>AY047362 .....
              701                      728
NC_001463 (gag720bp) TCCAGCAGGA GGAGGATTAA CAGTGGAT
>AY047362 .....

PileUp

MSF: 1347  Type: N  Check: 3238  ..

Name: NC_001463 (gag) (SEQ ID NO: 41) Len: 1347 Check: 6959 Weight: 0
Name: >AY047362 (SEQ ID NO: 42) Len: 1347 Check: 6279 Weight: 0

//

      1                      50
NC_001463 (gag) ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGAAAAG
>AY047362 .....

      51                      100
NC_001463 (gag) AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAG
>AY047362 .....

     101                      150
NC_001463 (gag) TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATGTGTT ATGGTGCCTT
>AY047362 .....

     151                      200
NC_001463 (gag) AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTGGACAAA
>AY047362 .....

     201                      250
NC_001463 (gag) AGTAAAATTT AGGACAATAT GGCAGAAGGT GAAGAATCTA ACTCCTGAGG
>AY047362 .....

     251                      300
NC_001463 (gag) AGAGTAACAA AAAAGACTTT ATGTCTTTGC AGGCCACATT AGCGGGTCTA
>AY047362 .....

     301                      350
NC_001463 (gag) ATGTGTTGCC AAATGGGGAT GAGACCTGAG ACATTGCAAG ATGCAATGGC
>AY047362 .....

     351                      400
NC_001463 (gag) TACAGTAATC ATGAAAGATG GGTACTGGA ACAAGAGGAA AAGAAGGAAG
>AY047362 .....

     401                      450
NC_001463 (gag) ACAAAGAGA AAAGGAAGAG AGTGTCTTCC CAATAGTAGT GCAAGCAGCA
>AY047362 .....

     451                      500
NC_001463 (gag) GGAGGGAGAA GCTGGAAAGC AGTAGATTCT GTAATGTTC AGCAACTGCA
>AY047362 .....

     501                      550
NC_001463 (gag) AACAGTAGCA ATGCAGCATG GCCTCGTGTC TGAGGACTTT GAAAGGCAGT
>AY047362 .....

     551                      600
NC_001463 (gag) TGGCATATTA TGCTACTACC TGGACAAGTA AAGACATACT AGAAGTATTG
>AY047362 .....TA AAGATATATT AGAAGTATTG

```

NC_001463 (gag) >AY047362	601 GCCATGATGC CTGGAATAG AGCTCAAAAG GAGTTAATTC AAGGGAAT GCCATGATGC CTGGAATAG AGCTCAAAA GAGTTAATTC AAGGGAAT	650
NC_001463 (gag) >AY047362	651 AAATGAAGAA GCAGAAAGGT GGAGAAGGAA TAATCCACCA CCTCCAGCAG GAATGAAGAA GCAGAAAGGT GGAGAAGGAA TAATCCACCA CCTCAAGCAG	700
NC_001463 (gag) >AY047362	701 GAGGAGGATT AACAGTGGAT CAAATTATGG GGGTAGGACA AACAAATCAA GCGGAGGATT AACAGTGGAT CAAATTATGG GGGTAGGACA AACAAATCAA	750
NC_001463 (gag) >AY047362	751 GCAGCAGCAC AAGCTAACAT GGATCAGGCA AGGCAAATAT GCCTGCAATG GCAGCGGCAC AGGCTAACAT GGATCAGGCA AGACAAATAT GCCTGCAATG	800
NC_001463 (gag) >AY047362	801 GGTAATAAAT GCATTAAGAG CAGTAAGACA TATGGCGCAC AGGCCAGGGA GGTAATAACA GCACTAAGAG CAGTGAGACA TATGGCTCAC AAACCAGGGA	850
NC_001463 (gag) >AY047362	851 ATCCAATGCT AGTAAAGCAA AAAACGAATG AGCCATATGA AGATTTTGCA ATCCGATGCT AGTAAAGCAA AAAACAAATG AGTCATATGA AGATTTTGCC	900
NC_001463 (gag) >AY047362	901 GCAAGACTGC TAGAAGCAAT AGATGCAGAG CCAGTTACAC AGCCTATAAA GCAAGACTGC TAGAAGCAAT AGATGCAGAA CCAGTTACAA AGCAAATAAA	950
NC_001463 (gag) >AY047362	951 AGATTATCTA AAGCTAACAC TATCTTATAC AAATGCATCA GCAGATTGTC AGAATATTTA AAGTTAACAT TATCTTACAC AAATGCATCC TCAGACTGTA	1000
NC_001463 (gag) >AY047362	1001 AGAAGCAAAT GGATAGAACA CTAGGACAAA GAGTACAACA AGCTAGTGTA AGAAACAGAT GGATAGAGTA CTAGGACAGA GAGTGCAACA AGCTAGTGTG	1050
NC_001463 (gag) >AY047362	1051 GAAGAAAAAA TGCAAGCATG TAGAGATGTG GGATCAGAAG GGTTCAAAAT GAAGAAAAAA TGCAAGCATG CAGAGATGTG GGATCAGAAG GATTCAGAAT	1100
NC_001463 (gag) >AY047362	1101 GCAATTGTTA GCACAAGCAT TAAGGCCAGG AAAAGGAAAA GGGAAATGGAC GC.....	1150
NC_001463 (gag) >AY047362	1151 AGCCACAAAG GTGTTACAAC TGTGGAAAAC CGGGACATCA AGCAAGGCAA	1200
NC_001463 (gag) >AY047362	1201 TGTAGACAAG GAATCATATG TCACAACGTG GGAAAGAGAG GACATATGCA	1250
NC_001463 (gag) >AY047362	1251 AAAAGAATGC AGAGGAAAGA GAGACATAAG GGGAAAACAG CAGGGAAACG	1300
NC_001463 (gag) >AY047362	1301 GAGGAGGGG GATACGTGTG GTGCCGTCCG CTCCTCCTAT GGAATAA	1347

표 11

```

PileUp
MSF: 733  Type: N  Check: 5855  ..
Name: NC_001463(gag720bp) (SEQ ID NO: 43) Len: 733 Check: 9482 Weight: 0
Name: >AY081139 (SEQ ID NO: 44) Len: 733 Check: 6373 Weight: 0

//
1 50
NC_001463(gag720bp) ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGAAAAG
>AY081139 .....TGCC GTAGACTCTG

51 100
NC_001463(gag720bp) AGATTATCCT G.AGCTCGAA AAATGTATCA AGCATGCATG CAAGATAAAA
>AY081139 TGATGTTCCTA CCAGCT.GCA TACAGTAGCA ATGCCGCATG GCCTCGTGTC

101 150
NC_001463(gag720bp) GTTCGACTCA GAGGGGAGCA CTTGACAGAA GGAAATTGTT TATGGTGCCT
>AY081139 TGAGGACTTT GAAAGG..CA GTTGGCATAT TATGCTACTA CCTGGACAAG

151 200
NC_001463(gag720bp) TAAACATTA GATTACATGT .TTGAGGACC ATAAAGAGGA ACCTGGACA
>AY081139 TAAAGA..TA TACTGGAAGT ATTGGCCATG ATGCCTGGGA ATAGAGCTCA

201 250
NC_001463(gag720bp) AAAGTAA..A ATTTAGGACA ATATGGCAGA AGGTGAAGAA TCTAACTCCT
>AY081139 AAAAGAATTA ATTCAAGGAA AATTAATGA AGAAGCAGAA .....

251 300
NC_001463(gag720bp) GAGGAGAGTA ACAAAAAAGA CTTTATGTCT TTGCAGGCCA CATTAGCGGG
>AY081139 .AGGTGGAGA AGGAATAATC CACCACCTCA A.GCAGCCG. ....GAGGA

301 350
NC_001463(gag720bp) TCTAATGTGT TGCCAA..AT GGGGATGAGA CCTGAGACAT TGCAAGATGC
>AY081139 TTAACAGTGG ATCAAATTAT GGGGGTAGGA CAAACAAATC AAGCAGCTGC

351 400
NC_001463(gag720bp) AATGGCTA.C AGTAATCATG AAAGATGGGT TACTGGAACA AGAGGAAAAG
>AY081139 ACAAGCTAAC ATGGATCAGG CAAGACAAAT A..TGCCTGC AATGGGTAAT

401 450
NC_001463(gag720bp) AAGGAAGACA AAAGAGAAA GGAAGAGAGT GTCTTCCCAA TAGTAGTCA
>AY081139 ATC..AGCCT TAAGAGCAGT GAGACATA.T GTCT..CATA AACCAGGG.A

451 500
NC_001463(gag720bp) AGCAGCAGGA GGGACAAGCT GGAAGAGCAGT AGATTCTGTA ATGTTCCAGC
>AY081139 ATCCGCTGCT AGTA.AAGCA AAAAAACAAAT GAGTCATATG AAGATTTTGC

501 550
NC_001463(gag720bp) AACTGCAAAC ..AGTAGCAA TGCAGCATGG CCTCGTGTCT GAGGACTTTC
>AY081139 AGCTAGACTG CTAGAAGCAA TAGATCCAGC CCCAGTAGCA CATC.CTATA

551 600
NC_001463(gag720bp) AAAGGCAGTT GGCATATTAT GCTAC...T ACCTGGACAA GTAAGACAT
>AY081139 AAAGATTATT TAAAGTTAAC ACTATCTTAT ACGAATGCAT CATCAGATTG

601 650
NC_001463(gag720bp) ACTAGAAGTA TTGGCCATGA TGCCTGGAAA TAGAGCTCAA AAGGAGTTAA
>AY081139 TCAAAAGCAA ATGGGTAGAA TGCTAGGATC GAGAGTCCAT CA..AGCCAG

651 700
NC_001463(gag720bp) TTCAAGGCAA ATTAATGAA GAAGCAGAAA GGTGGAGAAG GAATAATCCA
    
```

<427>

```

>AY081139 TGTGGGCCAA AAAA.....
              701                               733
NC_001463 (gag720bp) CCACCTCCAG CAGGAGGAGG ATTAACAGTG GAT
>AY081139 .....

PileUp

MSF: 1347  Type: N  Check: 2072  ..

Name: NC_001463 (gag) (SEQ ID NO: 45) Len: 1347 Check: 6959 Weight: 0
Name: >AY081139 (SEQ ID NO: 46) Len: 1347 Check: 5113 Weight: 0

//

1 50
NC_001463 (gag) ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGGAAAAG
>AY081139 .....

51 100
NC_001463 (gag) AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAAG
>AY081139 .....

101 150
NC_001463 (gag) TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCTT
>AY081139 .....

151 200
NC_001463 (gag) AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTGGACAAA
>AY081139 .....

201 250
NC_001463 (gag) AGTAAAATTT AGGACAATAT GGCAGAAGGT GAAGAATCTA ACTCCTGAGG
>AY081139 .....

251 300
NC_001463 (gag) AGAGTAACAA AAAAGACTTT ATGTCTTTGC AGGCCACATT AGCGGGTCTA
>AY081139 .....

301 350
NC_001463 (gag) ATGTGTTGCC AAATGGGGAT GAGACCTGAG ACATTGCAAG ATGCAATGGC
>AY081139 .....

351 400
NC_001463 (gag) TACAGTAATC ATGAAAGATG GGTACTGGA ACAAGAGGAA AAGAAGGAAG
>AY081139 .....

401 450
NC_001463 (gag) ACAAAGAGA AAAGGAAGAG AGTGTCTTCC CAATAGTAGT GCAAGCAGCA
>AY081139 .....

451 500
NC_001463 (gag) GGAGGGAGAA GCTGGAAAGC AGTAGATTCT GTAATGTTCC AGCAACTGCA
>AY081139 .....TGC CGTAGACTCT GTGATGTTCC ACCAGCTGCA

501 550
NC_001463 (gag) AACAGTAGCA ATGCAGCATG GCCTCGTGTC TGAGGACTTT GAAAGGCAGT
>AY081139 TACAGTAGCA ATGCCGCATG GCCTCGTGTC TGAGGACTTT GAAAGGCAGT

551 600
NC_001463 (gag) TGGCATATTA TGCTACTACC TGGACAAGTA AAGACATACT AGAAGTATTG
>AY081139 TGGCATATTA TGCTACTACC TGGACAAGTA AAGATATACT GGAAGTATTG

601 650

```

<428>

NC_001463 (gag) >AY081139	GCCATGATGC CTGGAAATAG AGCTCAAAAG GAGTTAATTC AAGGGAAATT GCCATGATGC CTGGAAATAG AGCTCAAAA GAATTAATTC AAGGAAAATT		
		651	700
NC_001463 (gag) >AY081139	AAATGAAGAA GCAGAAAGGT GGAGAAGGAA TAATCCACCA CCTCCAGCAG AAATGAAGAA GCAGAAAGGT GGAGAAGGAA TAATCCACCA CCTCAAGCAG		
		701	750
NC_001463 (gag) >AY081139	GAGGAGGATT AACAGTGGAT CAAATTATGG GGGTAGGACA AACAAATCAA GCGGAGGATT AACAGTGGAT CAAATTATGG GGGTAGGACA AACAAATCAA		
		751	800
NC_001463 (gag) >AY081139	GCAGCAGCAC AAGCTAACAT GGATCAGGCA AGGCAAATAT GCCTGCAATG GCAGCTGCAC AAGCTAACAT GGATCAGGCA AGACAAATAT GCCTGCAATG		
		801	850
NC_001463 (gag) >AY081139	GGTAATAAAT GCATTAAGAG CAGTAAGACA TATGGCCGAC AGGCCAGGGA GGTAATATCA GCCTTAAGAG CAGTGAGACA TATGTCTCAT AAACCAGGGA		
		851	900
NC_001463 (gag) >AY081139	ATCCAATGCT AGTAAAGCAA AAAACGAATG AGCCATATGA AGATTTTGCA ATCCGCTGCT AGTAAAGCAA AAAACAAATG AGTCATATGA AGATTTTGCA		
		901	950
NC_001463 (gag) >AY081139	GCAAGACTGC TAGAAGCAAT AGATGCAGAG CCAGTTACAC AGCCTATAAA GCTAGACTGC TAGAAGCAAT AGATCCAGCC CCAGTAGCAC ATCCTATAAA		
		951	1000
NC_001463 (gag) >AY081139	AGATTATCTA AAGCTAACAC TATCTTATAC AAATGCATCA GCAGATTGTC AGATTATTTA AAGTTAACAC TATCTTATAC GAATGCATCA TCAGATTGTC		
		1001	1050
NC_001463 (gag) >AY081139	AGAAGCAAAT GGATAGAACA CTAGGACAAA GAGTACAACA AGCTAGTGTA AAAAGCAAAT GGGTAGAATG CTAGGATCGA GAGTCCATCA AGCCAGTGTG		
		1051	1100
NC_001463 (gag) >AY081139	GAAGAAAAA TGCAAGCATG TAGAGATGTG GGATCAGAAG GGTTCAAAAT GGCAAAAAA		
		1101	1150
NC_001463 (gag) >AY081139	GCAATTGTTA GCACAAGCAT TAAGGCCAGG AAAAGGAAAA GGAATGGAC		
		1151	1200
NC_001463 (gag) >AY081139	AGCCACAAAG GTGTTACAAC TGTGGAAAAC CGGGACATCA AGCAAGGCAA		
		1201	1250
NC_001463 (gag) >AY081139	TGTAGACAAG GAATCATATG TCACAACCTGT GGAAAGAGAG GACATATGCA		
		1251	1300
NC_001463 (gag) >AY081139	AAAAGAATGC AGAGGAAAGA GAGACATAAG GGGAAAACAG CAGGGAAACG		
		1301	1347
NC_001463 (gag) >AY081139	GGAGGAGGGG GATACGTGTG GTGCCGTCGG CTCCTCCTAT GGAATAA		

표 12

```

PileUp
MSF: 731  Type: N  Check: 9546  ..
Name: NC_001463(gag720bp) (SEQ ID NO: 47) Len: 731 Check: 7595 Weight: 0
Name: >AY101347 (SEQ ID NO: 48)      Len: 731      Check: 7962      Weight: 0
Name: >AY101348 (SEQ ID NO: 49)      Len: 731      Check: 3989      Weight: 0

//

NC_001463(gag720bp)      1                               50
>AY101347                .....
>AY101348                .....

NC_001463(gag720bp)      51                               100
>AY101347                .....AGCAG TAGATTCTGT AATGTTCCAG
>AY101348                .....AGCCG TAGATTCTGT AATGTTCCAG

NC_001463(gag720bp)     101                              150
>AY101347                TTCGACT.CA GAGGGGAGCA CTTGACAGAA GGAAATTGTT TATGGTGCCT
>AY101348                ..CAACTGCA AACAGTAGCA AT..GCAGCA TGGACTCCTG TATGAAGACT

NC_001463(gag720bp)     151                              200
>AY101347                TAAACATTA GATTACATGT TTGAGGACCA TAAAGAGGAA CCTTGGACAA
>AY101348                TTGAAAGGCT GTCGGCATAT TATGCTACTA CCTGGAC.AA GTAAAGATAT

NC_001463(gag720bp)     201                              250
>AY101347                AAGTAAAT  TAGGACAATA TGGCAGAAGG TGAAGAATCT AACTCCTGAG
>AY101348                CTTAGAAGTA CTGGCCATGA TGCCTG..G. ....GAATAG AGCTCAAAAA

NC_001463(gag720bp)     251                              300
>AY101347                GAGAGTAA.. CAAAAAGAC TTTATGTCTT TGCAGGCCAC ATTAGCGGGT
>AY101348                GA.ATTAAT  CAAGGAAAT  TAAATGAAGA AGCAGAAAGG TGGAGAAGGA

NC_001463(gag720bp)     301                              350
>AY101347                CTAATGTGTT GCCAAATGGG GATGAGACCT GAGACATTGC AAGATGCAAT
>AY101348                ATAATCCACC ACCTCAAGCA GGCC.GAGGA TTAACAGTGG ATCAAATTAT

NC_001463(gag720bp)     351                              400
>AY101347                GGCTACAGTA ATCATGAAAG ATGG.GTTAC TGAACAAGA GGAAAAGAAG
>AY101348                GGGGGTAGGA CAAACAAATC AAGCAGCTGC ACAAGCTAAC ATGGATCAGG

NC_001463(gag720bp)     401                              450
>AY101347                GAAGACAAA  GAGAAAAGGA AGAG.AGTGT CTTCCC.AAT AGTAGTGCAA
>AY101348                CAAGACAAAT ATGCTGCAA TGGGTAAATAT CAGCCTTAAG AGCAGTGAGA

NC_001463(gag720bp)     451                              500
>AY101347                GCAGCAGGAG GGAGAAGCTG GAAAGCAGTA GATTCTGTAA TGTTCAGCA
>AY101348                .CATATGTCT CATAAACAG GGAATCCGCT GCTAGTA.AA GCAAAAAACA

NC_001463(gag720bp)     501                              550
>AY101347                ACTG...CAA ACAGTAGCAA TGCAGCATGG CCTCGTGTCT GAGGACTTTG
>AY101347                AATGAGTCAT ATGAAGATT  TGCAGCAAGA CTGCTAGAAG CAATAGATGC
    
```

<430>


```

>AY101348 AATGAGCCAT ATGAAGAATT TGCAGCAAGG CTGCTGGAAAT CAATAGRTGC
551 600
NC_001463 (gag720bp) AAAGGCAGTT GG.CATATTA TGCTACTACC TGGACAAGTA AAGAC.ATAC
>AY101347 AGAGCCAGTA GCACATCCTA TAAAAGAATA CTTA.AAGTT AACACTATCT
>AY101348 CGAGCCAGTT AATCAGCCCA TAAAAGAATA TCTA.AAACT AACGTTGTCT

601 650
NC_001463 (gag720bp) TAGAAGTATT GGCCATGATG CCTGGAAATA GAGCTCAAAA GGAGTTAATT
>AY101347 TATACGAATG CATCATCA.G ATTGTCAAAA G....CAAAT GGATAGAATG
>AY101348 TATACGAATG CATCCTCA.G ATTGTCAGAA G....CAAAT GGATAGAACA

651 700
NC_001463 (gag720bp) CAAGGGAAT TAAATGAAGA AGCAGAAAAG TGGAGAAGGA ATAATCCACC
>AY101347 CTGG...AAT CAAGAGTACA ACAAGCTAG. TGTAGAACAA AAAA.....
>AY101348 CTAG...GAC AAAGAGTCAA ACAAGCTAG. TGTAGAACAA AAAA.....

701 731
NC_001463 (gag720bp) ACCTCCAGCA GGAGGAGGAT TAACAGTGGGA T
>AY101347 .....
>AY101348 .....

FileUp
MSF: 1347 Type: N Check: 2815 ..

Name: NC_001463 (gag) (SEQ ID NO: 50) Len: 1347 Check: 6959 Weight: 0
Name: >AY101347 (SEQ ID NO: 51) Len: 1347 Check: 969 Weight: 0
Name: >AY101348 (SEQ ID NO: 52) Len: 1347 Check: 4887 Weight: 0

//

1 50
NC_001463 (gag) ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGAAAAG
>AY101347 .....
>AY101348 .....

51 100
NC_001463 (gag) AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAAAG
>AY101347 .....
>AY101348 .....

101 150
NC_001463 (gag) TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCTT
>AY101347 .....
>AY101348 .....

151 200
NC_001463 (gag) AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTTGGACAAA
>AY101347 .....
>AY101348 .....

201 250
NC_001463 (gag) AGTAAAATTT AGGACAATAT GGCAGAAGGT GAAGAATCTA ACTCCTGAGG
>AY101347 .....
>AY101348 .....

251 300
NC_001463 (gag) AGAGTAACAA AAAAGACTTT ATGTCTTTGC AGGCCACATT AGCGGGTCTA
>AY101347 .....
>AY101348 .....

301 350
NC_001463 (gag) ATGTGTTGCC AAATGGGGAT GAGACCTGAG ACATTGCAAG ATGCAATGGC

```

<431>

>AY101347
>AY101348
	351				400
NC_001463 (gag)	TACAGTAATC	ATGAAAGATG	GGTTACTGGA	ACAAGAGGAA	AAGAAGGAAG
>AY101347
>AY101348
	401				450
NC_001463 (gag)	ACAAAAGAGA	AAAGGAAGAG	AGTGTCTTCC	CAATAGTAGT	GCAAGCAGCA
>AY101347
>AY101348
	451				500
NC_001463 (gag)	GGAGGGAGAA	GCTGGAAAGC	AGTAGATTCT	GTAATGTTCC	AGCAACTGCA
>AY101347AGC	AGTAGATTCT	GTAATGTTCC	AGCAACTGCA
>AY101348AGC	CGTAGATTCT	GTAATGTTCC	AGCAGCTGCA
	501				550
NC_001463 (gag)	AACAGTAGCA	ATGCAGCATG	GCCTCGTGTC	TGAGGACTTT	GAAAGGCAGT
>AY101347	AACAGTAGCA	ATGCAGCATG	GACTCGTGTA	TGAAGACTTT	GAAAGGCTGT
>AY101348	AACAGTAGCA	ATGCAGCATG	GCCTCGTGTC	AGAGGACTTT	GAAAGGCTTC
	551				600
NC_001463 (gag)	TGGCATATTA	TGCTACTACC	TGGACAAGTA	AAGACATACT	AGAAGTATTG
>AY101347	CGGCATATTA	TGCTACTACC	TGGACAAGTA	AAGATATACT	GGAGTATTG
>AY101348	CAGCATATCA	TGCTACTACC	TGGGCAAGTA	AAGATATCTT	AGAAGTACTG
	601				650
NC_001463 (gag)	GCCATGATGC	CTGGAAATAG	AGCTCAAAAG	GAGTTAATTC	AAGGGAAATT
>AY101347	GCCATGATGC	CTGGGAATAG	AGCTCAAAAA	GAATTAATTC	AAGGAAAATT
>AY101348	GCCATGATGC	CTGGAAATAG	AGCTCAAAAA	GAGTTAATTC	AAGGGAAATT
	651				700
NC_001463 (gag)	AAATGAAGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCCAGCAG
>AY101347	AAATGAAGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCAAGCAG
>AY101348	AAATGAAGAA	GCAGAGAGGT	GGAGAAGGAA	TAATCCACCA	CCTCCAGCAG
	701				750
NC_001463 (gag)	GAGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAA
>AY101347	GCGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAA
>AY101348	GAGGAGGGTT	AACAGTGGAT	CAAATTATGG	GAGTAGGACA	AACAAATCAG
	751				800
NC_001463 (gag)	GCAGCAGCAC	AAGCTAACAT	GGATCAGGCA	AGGCAAATAT	GCCTGCAATG
>AY101347	GCAGCTGCAC	AAGCTAACAT	GGATCAGGCA	AGACAAATAT	GCCTGCAATG
>AY101348	GCAGCGGCAC	AAGCAAACAT	GGATCAGGCA	AGACAAATAT	GCCTACAATG
	801				850
NC_001463 (gag)	GGTAATAAAT	GCATTAAGAG	CAGTAAGACA	TATGGCGCAC	AGGCCAGGGA
>AY101347	GGTAATATCA	GCCTTAAGAG	CAGTGAGACA	TATGTCTCAT	AAACCAGGGA
>AY101348	GGTGATATCA	GCACTAAGAG	CAGTAAGGCA	TATGGCTCAC	AAGCCAGGGA
	851				900
NC_001463 (gag)	ATCCAATGCT	AGTAAAGCAA	AAAACGAATG	AGCCATATGA	AGATTTTGCA
>AY101347	ATCCGCTGCT	AGTAAAGCAA	AAAACAAATG	AGTCATATGA	AGATTTTGCA
>AY101348	ATCCAATGTT	AGTAAAGCAA	AAAGCAAATG	AGCCATATGA	AGAATTTGCA
	901				950
NC_001463 (gag)	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAG	CCAGTTACAC	AGCCTATAAA
>AY101347	GCAAGACTGC	TAGAAGCAAT	AGATGCAGAG	CCAGTAGCAC	ATCCTATAAA
>AY101348	GCAAGGCTGC	TGGAAGCAAT	AGATGCCGAG	CCAGTTAATC	AGCCCATAAA
	951				1000

NC_001463 (gag)	AGATTATCTA	AAGCTAACAC	TATCTTATAC	AAATGCATCA	TCAGATTGTC
>AY101347	AGAATACTTA	AAGTTAACAC	TATCTTATAC	GAATGCATCA	TCAGATTGTC
>AY101348	AGAATATCTA	AAACTAACGT	TGTCTTATAC	GAATGCATCC	TCAGATTGTC
	1001				1050
NC_001463 (gag)	AGAAGCAAAT	GGATAGAACA	CTAGGACAAA	GAGTACAACA	AGCTAGTGTA
>AY101347	AAAAGCAAAT	GGATAGAATG	CTGGAATCAA	GAGTACAACA	AGCTAGTGTA
>AY101348	AGAAGCAAAT	GGATAGAACA	CTAGGACAAA	GAGTCAAACA	AGCTAGTGTA
	1051				1100
NC_001463 (gag)	GAAGAAAAAA	TGCAAGCATG	TAGAGATGTG	GGATCAGAAG	GGTTCAAAAT
>AY101347	GAACAAAAAA
>AY101348	GAACAAAAAA
	1101				1150
NC_001463 (gag)	GCAATTGTGA	GCACAAGCAT	TAAGGCCAGG	AAAAGGAAAA	GGGAATGGAC
>AY101347
>AY101348
	1151				1200
NC_001463 (gag)	AGCCACAAAG	GTGTTACAAC	TGTGGAAAAC	CGGGACATCA	AGCAAGGCAA
>AY101347
>AY101348
	1201				1250
NC_001463 (gag)	TGTAGACAAG	GAATCATATG	TCACAACGTG	GGAAAGAGAG	GACATATGCA
>AY101347
>AY101348
	1251				1300
NC_001463 (gag)	AAAAGAATGC	AGAGGAAAGA	GAGACATAAG	GGGAAAACAG	CAGGGAAACG
>AY101347
>AY101348
	1301				1347
NC_001463 (gag)	GGAGGAGGGG	GATACGTGTG	GTGCCGTCCG	CTCCTCCTAT	GGAATAA
>AY101347
>AY101348

표 13

```

PileUp
MSF: 720  Type: N  Check: 3690  ..

Name: NC_001463 (gag720bp) (SEQ ID NO: 53) Len: 720 Check: 5792 Weight: 0
Name: >L78446 (SEQ ID NO: 54) Len: 720 Check: 272 Weight: 0
Name: >L78447 (SEQ ID NO: 55) Len: 720 Check: 1999 Weight: 0
Name: >L78450 (SEQ ID NO: 56) Len: 720 Check: 9633 Weight: 0
Name: >L78451 (SEQ ID NO: 57) Len: 720 Check: 5177 Weight: 0
Name: >L78453 (SEQ ID NO: 58) Len: 720 Check: 817 Weight: 0

//

NC_001463 (gag720bp) 1 50
ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGAAAAG
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

NC_001463 (gag720bp) 51 100
AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAG
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

NC_001463 (gag720bp) 101 150
TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCTT
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

NC_001463 (gag720bp) 151 200
AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTTGGACAAA
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

NC_001463 (gag720bp) 201 250
AGTAAAATTT AGGACAATAT GGCAGAAGGT GAACAATCTA ACTCCTGAGG
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

NC_001463 (gag720bp) 251 300
AGAGTAACAA AAAAGACTTT ATGCTTTGTC AGGCCACATT AGCGGGTCTA
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

NC_001463 (gag720bp) 301 350
ATGTGTTGCC AAATGGGGAT GAGACCTGAG ACATTGCAAG ATGCAATGCC
  
```

>L78446
>L78447
>L78450
>L78451
>L78453
	351				400
NC_001463 (gag720bp)	TACAGTAATC	ATGAAAGATG	GTTACTGGA	ACAAGAGGAA	AAGAAGGAAG
>L78446
>L78447
>L78450
>L78451
>L78453
	401				450
NC_001463 (gag720bp)	ACAAAAGAGA	AAAGGAAGAG	AGTGTCTTCC	CAATAGTAGT	GCAAGCAGCA
>L78446
>L78447
>L78450
>L78451
>L78453
	451				500
NC_001463 (gag720bp)	GGAGGGAGAA	GCTGGAAAGC	AGTAGATTCT	GTAATGTTCC	AGCAACTGCA
>L78446
>L78447
>L78450
>L78451
>L78453
	501				550
NC_001463 (gag720bp)	AACAGTAGCA	ATGCAGCATG	GCCTCGTGTC	TGAGGACTTT	GAAAGGCAGT
>L78446CAGCATG	GCCTCGTGTC	CGAGGACTTT	GAAAGGCAGT
>L78447CAGCATG	GAATAGTATC	AGAAGACTTT	GAGAGGCAAC
>L78450CAACATG	GGATAGTATC	AGAGGAATTT	GAGAGACAAA
>L78451CAGCATG	GACTAGTATC	AGAAGAATTT	GAAAGGCAGC
>L78453CAGCATG	GACTTGTGTC	CGAAGATTTT	GAGAGGCAAT
	551				600
NC_001463 (gag720bp)	TGGCATATTA	TGCTACTACC	TGGACAAGTA	AAGACATACT	AGAAGTATTG
>L78446
>L78447
>L78450
>L78451
>L78453
	601				650
NC_001463 (gag720bp)	GCCATGATGC	CTGGAAATAG	AGCTCAAAAAG	GAGTTAATTC	AAGGGAAATT
>L78446
>L78447
>L78450
>L78451
>L78453
	651				700
NC_001463 (gag720bp)	AAATGAAGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCCAGCAG
>L78446
>L78447
>L78450
>L78451
>L78453
	701	720			
NC_001463 (gag720bp)	GAGGAGGATT	AACAGTGGAT			
>L78446			
>L78447			
>L78450			
>L78451			

```

>L78453 .....

FileUp
  MSF: 1347  Type: N  Check: 6947  ..

Name: NC_001463(gag) (SEQ ID NO: 59) Len: 1347 Check: 6959 Weight: 0
Name: >L78446 (SEQ ID NO: 60) Len: 1347 Check: 6690 Weight: 0
Name: >L78447 (SEQ ID NO: 61) Len: 1347 Check: 8417 Weight: 0
Name: >L78450 (SEQ ID NO: 62) Len: 1347 Check: 6051 Weight: 0
Name: >L78451 (SEQ ID NO: 63) Len: 1347 Check: 1595 Weight: 0
Name: >L78453 (SEQ ID NO: 64) Len: 1347 Check: 7235 Weight: 0

//

      1                               50
NC_001463(gag) ATGGTGAGTC TAGATAGAGA CATGGCGAGG CAAGTCTCCG GGGGAAAAG
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

      51                               100
NC_001463(gag) AGATTATCCT GAGCTCGAAA AATGTATCAA GCATGCATGC AAGATAAAG
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

     101                               150
NC_001463(gag) TTCGACTCAG AGGGGAGCAC TTGACAGAAG GAAATTGTTT ATGGTGCCTT
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

     151                               200
NC_001463(gag) AAAACATTAG ATTACATGTT TGAGGACCAT AAAGAGGAAC CTTGGACAAA
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

     201                               250
NC_001463(gag) AGTAAAATTT AGGACRATAI GGCAGAAGGT GAAGAATCTA ACTCCTGAGG
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

     251                               300
NC_001463(gag) AGAGTAAACAA AAAAGACTTT ATGTCTTTGC AGGCCACATT AGCGGGTCTA
>L78446 .....
>L78447 .....
>L78450 .....
>L78451 .....
>L78453 .....

     301                               350
NC_001463(gag) ATGTGTTGCC AAATGGGGAT GAGACCTGAG ACATTGCAAG ATGCAATGCC
>L78446 .....
>L78447 .....

```

>L78450
>L78451
>L78453
	351			400	
NC_001463 (gag)	TACAGTAATC	ATGAAAGATG	GTTTACTGGA	ACAAGAGGAA	AAGAAGGAAG
>L78446
>L78447
>L78450
>L78451
>L78453
	401			450	
NC_001463 (gag)	ACAAAAGAGA	AAAGGAAGAG	AGTGTCTTCC	CAATAGTAGT	GCAAGCAGCA
>L78446
>L78447
>L78450
>L78451
>L78453
	451			500	
NC_001463 (gag)	GGAGGGAGAA	GCTGGAAAGC	AGTAGATTCT	GTAATGTTCC	AGCAACTGCA
>L78446
>L78447
>L78450
>L78451
>L78453
	501			550	
NC_001463 (gag)	AACAGTAGCA	ATGCAGCATG	GCCTCGTGTC	TGAGGACTTT	GAAGGGCAGT
>L78446
>L78447
>L78450
>L78451
>L78453
	551			600	
NC_001463 (gag)	TGTCATATTA	TGCTACTACC	TGGACAAGTA	AAGACATACT	AGAAGTATTG
>L78446
>L78447
>L78450
>L78451
>L78453
	601			650	
NC_001463 (gag)	GCCATGATGC	CTGGAAATAG	AGCTCAAAG	GAGTTAATTC	AAGGGAAATT
>L78446
>L78447
>L78450
>L78451
>L78453
	651			700	
NC_001463 (gag)	AAATGAAGAA	GCAGAAAGGT	GGAGAAGGAA	TAATCCACCA	CCTCCAGCAG
>L78446
>L78447
>L78450
>L78451
>L78453
	701			750	
NC_001463 (gag)	GAGGAGGATT	AACAGTGGAT	CAAATTATGG	GGGTAGGACA	AACAAATCAA
>L78446
>L78447
>L78450
>L78451
>L78453

	751		800
NC_001463 (gag)	GCAGCAGCAC AAGCTAACAT GGATCAGGCA AGGCCAAATAT GCCTGCAATG		
>L78446		
>L78447		
>L78450		
>L78451		
>L78453		
	801		850
NC_001463 (gag)	GGTAATAAAT GCATTAAGAG CAGTAAGACA TATGGCGCAC AGGCCAGGGA		
>L78446		
>L78447		
>L78450		
>L78451		
>L78453		
	851		900
NC_001463 (gag)	ATCCAATGCT AGTAAAGCAA AAAACGAATG AGCCATATGA AGATTTTGCA		
>L78446		
>L78447		
>L78450		
>L78451		
>L78453		
	901		950
NC_001463 (gag)	GCAAGACTGC TAGAAGCAAT AGATGCAGAG CCAGTTACAC AGCCTATAAA		
>L78446		
>L78447		
>L78450		
>L78451		
>L78453		
	951		1000
NC_001463 (gag)	AGATTATCTA AAGCTAACAC TATCTTATAC AAATGCATCA GCAGATTGTC		
>L78446		
>L78447		
>L78450		
>L78451		
>L78453		
	1001		1050
NC_001463 (gag)	AGAAGCAAAT GGATAGAACA CTAGGACAAA GAGTACAACA AGCTAGTGTA		
>L78446		
>L78447		
>L78450		
>L78451		
>L78453		
	1051		1100
NC_001463 (gag)	GAAGAAAAA TGCACGATG TAGAGATGTG GGATCAGAAG GGTTCAAAAAT		
>L78446		
>L78447		
>L78450		
>L78451		
>L78453		
	1101		1150
NC_001463 (gag)	GCAATTGTTA GCACAAGCAT TAAGCCAGG AAAAGGAAAA GGGAATGGAC		
>L78446		
>L78447		
>L78450		
>L78451		
>L78453		
	1151		1200

NC_001463(gag)	AGCCACAAAG	GTGTTACAAC	TGTGGAAAPC	CGGGACBPA	AGCAAGGCAA
>L78446
>L78447
>L78450
>L78451
>L78453
	1201				1250
NC_001463(gag)	TGTAGACAAG	GAATCATATG	TCACAACGTG	GGAAGAGAG	GACATATGCA
>L78446
>L78447
>L78450
>L78451
>L78453
	1251				1300
NC_001463(gag)	AAAAGAATGC	AGAGGAAAGA	GAGACATAAG	GGGAAAACAG	CAGGAAAACG
>L78446
>L78447
>L78450
>L78451
>L78453
	1301				1347
NC_001463(gag)	GGAGGAGGGG	GATACGTGTG	GTGCCGTCCG	CTCCTCCTAT	GGAATAA
>L78446
>L78447
>L78450
>L78451
>L78453

<439>

표 14

NC_001463(gag720bp) 대 AF015181		
양성신호: 41.0% 상동성: 41.0%		
	NC_001463(gag720bp)	>AF015181
NC_001463(gag720bp)	100	41
>AF015181		100

<440>

NC_001463(gag) 대 AF015181		
양성신호: 40.6% 상동성 : 40.6%		
	NC_001463(gag)	>AF015181
NC_001463(gag)	100	41
>AF015181		100

<441>

NC_001463(gag 720bp) 대 AF402664~8						
양성신호: 91.1% 상동성: 32.2%						
	NC_001463(gag720bp)	>AF402664	>AF402665	>AF402666	>AF402667	>AF402668
NC_001463(gag720bp)	100	43	44	43	33	43
>AF402664		100	99	96	80	98
>AF402665			100	97	80	99
>AF402666				100	81	96
>AF402667					100	80
>AF402668						100

<442>

NC_001463(gag) 대 AF402664~8
 양성신호: 49.1% 상동성: 35.0%

	NC_001463 (gag)	>AF402664	>AF402665	>AF402666	>AF402667	>AF402668
NC_001463(gag)	100	45	46	44	36	45
>AF402664		100	99	98	89	99
>AF402665			100	98	89	99
>AF402666				100	90	98
>AF402667					100	89
>AF402668						100

<443>

NC_001463(gag720bp) 대 AJ305040~2
 양성신호: 80.5% 상동성: 38.1%

	NC_001463 (gag720bp)	>AJ305040	>AJ305041	>AJ305042
NC_001463(gag720bp)	100	39	42	39
>AJ305040		100	93	99
>AJ305041			100	93
>AJ305042				100

<444>

NC_001463(gag) 대 AJ305040~2
 양성신호: 44.3% 상동성: 38.8%

	NC_001463 (gag)	>AJ305040	>AJ305041	>AJ305042
NC_001463(gag)	100	41	41	40
>AJ305040		100	96	100
>AJ305041			100	96
>AJ305042				100

<445>

NC_001463(gag720bp) 대 AY047362
 양성신호: 40.2% 상동성: 40.2%

	NC_001463 (gag720bp)	>AY047362
NC_001463(gag720bp)	100	40
>AY047362		100

<446>

NC_001463(gag) 대 AY047362
 양성신호: 35.7% 상동성: 35.7%

	NC_001463 (gag)	>AY047362
NC_001463(gag)	100	36
>AY047362		100

<447>

NC_001463(gag720bp) 대 AY081139
 양성신호: 40.0% 상동성: 40.0%

	NC_001463 (gag720bp)	>AY081139
NC_001463(gag720bp)	100	40
>AY081139		100

<448>

NC_001463(gag) 대 AY081139 양성신호: 39.8% 상동성: 39.8%		
	NC_001463 (gag)	>AY081139
NC_001463(gag)	100	40
>AY081139		100

<449>

NC_001463(gag720bp) 대 AY101347~8 양성신호: 78.1% 상동성: 35.0%			
	NC_001463 (gag720bp)	>AY101347	>AY101348
NC_001463(gag720bp)	100	40	36
>AY101347		100	91
>AY101348			100

<450>

NC_001463(gag) 대 AY101347~8 양성신호: 43.9 % 상동성: 37.9%			
	NC_001463 (gag)	>AY101347	>AY101348
NC_001463(gag)	100	41	40
>AY101347		100	95
>AY101348			100

<451>

NC_001463(gag720bp) 대 L78446, 7, 50, 51, 53 양성신호: 17.6 % 상동성: 11.9%						
	NC_001463 (gag720bp)	>L78446	>L78447	>L78450	>L78451	>L78453
NC_001463(gag720bp)	100	17	14	14	15	15
>L78446		100	96	96	97	97
>L78447			100	98	97	96
>L78450				100	96	96
>L78451					100	97
>L78453						100

<452>

NC_001463(gag) 대 L78446, 47, 50, 51, 53 양성신호: 9.4% 상동성: 6.4%						
	NC_001463 (gag)	>L78446	>L78447	>L78450	>L78451	>L78453
NC_001463(gag)	100	9	7	7	8	8
>L78446		100	98	98	98	98
>L78447			100	99	98	98
>L78450				100	98	98
>L78451					100	98
>L78453						100

<453>

표 15

NC_001463(전장 유전체) 대 AF322109(전장 유전체) 양성신호: 68.2% 상동성: 68.2%		
	NC_001463	AF322109
NC_001463	100	68
AF322109		100

<454>

NC_001463(gag) 대 AF322109(gag)
양성신호: 73.1% 상동성: 73.1%

	NC_001463 (gag)	AF322109 (gag)
NC_001463(gag)	100	73
AF322109(gag)		100

<455>

NC_001463(5'LTR 영역) 대 AF322109(5'LTR 영역)
양성신호: 59.8% 상동성: 59.8%

	NC_001463 (5')	AF322109 (5')
NC_001463(5')	100	60
AF322109(5')		100

<456>

NC_001463(pol) 대 AF322109(pol)
양성신호: 74.9% 상동성: 74.9%

	NC_001463 (pol)	AF322109 (pol)
NC_001463(pol)	100	75
AF322109(pol)		100

<457>

NC_001463(rev) 대 AF322109(rev)
양성신호: 48.3% 상동성: 48.3%

	NC_001463 (rev)	AF322109 (rev)
NC_001463(rev)	100	48
AF322109(rev)		100

<458>

NC_001463(vif) 대 AF322109(vif)
양성신호: 66.0% 상동성: 66.0%

	NC_001463 (vif)	AF322109 (vif)
NC_001463(vif)	100	66
AF322109(vif)		100

<459>

도면의 간단한 설명

<94>

도 1은 CAEV 프로바이러스 유전체 구조에 대한 개략도이다.

<95>

도 2a는 플라스미드 pMGP/RRE(서열번호: 77)의 개략도이다. pMGP/RRE(서열번호: 77)는 CAEV *gag-pol* 코딩 영역(bp 709-5,243)의 상류에 위치하는 MCMV MIEP 영역(bp 1-660), RRE 영역(bp 5,426-5,627 또는 bp 5,368-5,669), 및 소 성장호르몬(BGH) 폴리아데닐화 신호(bp 5,751-5,984)를 포함하는 9,446 bp의 플라스미드이다. 상기 벡터는 또한 네오마이신 내성 유전자 코딩 영역(bp 8,151-7,155), SV40 복제 기점(origin of replicaiton)(bp 8,509-8,152), Col E1 복제 기점(bp 6,115-6,698), 및 암피실린 내성 유전자 영역(bp 9,362-8,528)을 포함한다.

<96>

도 2b는 플라스미드 pMGP/REV/RRE의 개략도이다. pMGP/REV/RRE는 CAEV *gag-pol* 코딩 영역(bp 726-5,258)의 상류에 위치하는 MCMV MIEP 영역(bp 1-660에 위치) 및 CAEV의 주 스플라이싱 제공 위치(bp 688-704), *rev* 코딩 영역의 제1 엑손(exon)(bp 5,383-5,494), RRE 영역(bp 5,540-5,841), *rev* 코딩 영역의 제2 엑손(bp 5,888-6,177), 및 소 성장호르몬(BGH) 폴리아데닐화 신호(bp 6,229-6,462)를 포함하는 9,924 bp의 플라스미드이다. 상기 벡터는 또한 네오마이신 내성 유전자 코딩 영역(bp 7,633-8,629), SV40 복제 기점(bp 8,987-8,630), Col E1 복제 기점(bp 6,593-7,176), 및 암피실린 내성 유전자 영역(bp 9,840-9,006)을 포함한다.

<97>

도 3a는 플라스미드 pCAH/SINd(서열번호: 73)의 개략도이다. pCAH/SINd(서열번호: 73)는 HCMV MIEP(bp 1-588), CAEV 5'LTR에서의 R-U5 서열 영역(bp 611-772), RRE 영역(bp 796-1,154), 및 U3-삭제된 CAEV 3'LTR 영역(bp 1,275-1,458)을 포함하는 3,566 bp의 플라스미드이다. 상기 벡터는 또한 Col E1 복제 기점(bp 1,863-

2,466), 및 카나마이신 내성 유전자 코딩 영역(bp 2,698-3,510)을 포함한다.

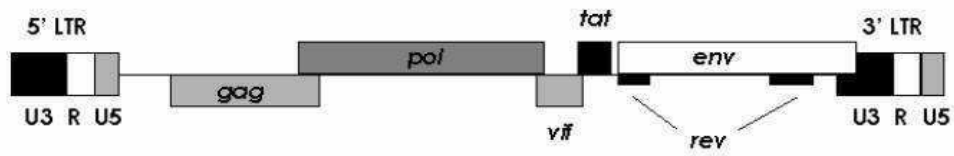
- <98> 도 3b는 플라스미드 pCAH/SINd0(서열번호: 67)의 개략도이다. pCAH/SINd0(서열번호: 67)은 HCMV MIEP(bp 1-588), CAEV 5'LTR에서의 R-U5 서열 영역(bp 611-772), 프라이머 결합 위치(PBS)를 포함하는 잔여 비번역 서열(bp 773-789), RRE 영역(bp 1,141-1,499), 및 U3-삭제된 CAEV 3'LTR 영역(bp 1,620-1,803)을 포함하는 3,911 bp의 플라스미드이다. 상기 벡터는 또한 Col E1 복제 기점(bp 2,208-2,791) 및 카나마이신 내성 유전자 코딩 영역(bp 3,043-3,855)을 포함한다.
- <99> 도 3c는 플라스미드 pCAH/SINd1(서열번호: 68)의 개략도이다. pCAH/SINd1(서열번호: 68)은 HCMV MIEP(bp 1-588) 프로모터; CAEV 5'LTR에서의 R-U5 서열 영역(bp 610-772); PBS 위치를 포함하는 잔여 비번역 영역(bp 773-789); 시작 ATG 코돈(bp1121-1123) 및 상기 시작 ATG 코돈의 하류에 위치하는 ATG 코돈(bp1142-1144)에서 ATG로부터 TAG로의 점변이(point mutation)를 갖는, *gag* 유전자의 327 bp의 절편(bp 1,121-1,448); RRE 영역(bp 1,468-1,826); 및 U3-삭제된 CAEV 3'LTR 영역(bp 1,947-2,130)을 포함하는 4,238 bp의 플라스미드이다. 상기 벡터는 또한 Col E1 복제 기점(bp 2,535-3,118) 및 카나마이신 내성 유전자 영역(bp 3,370-4,182)을 포함한다.
- <100> 도 3d는 플라스미드 pCAH/SINd2(서열 번호: 69)의 개략도이다. 플라스미드 pCAH/SINd2(서열 번호: 69)는 HCMV MIEP(bp 1-588); CAEV 5'LTR에서의 R-U5 서열 영역(bp 610-772); PBS 위치를 포함하는 잔여 비번역 영역(bp 773-789); 시작 ATG 코돈(bp1121-1123) 및 상기 시작 ATG 코돈의 하류에 위치하는 ATG 코돈(bp1142-1144)에서 점변이를 갖는, *gag* 유전자의 612 bp의 절편(bp 1,121-1,733); RRE 영역(bp 1,753-2,111); 및 U3-삭제된 CAEV 3'LTR 영역(bp 2,232-2,415)을 포함하는 4,523 bp 플라스미드이다. 상기 벡터는 또한 Col E1 복제 기점(bp 2,820-3,403) 및 카나마이신 내성 유전자 코딩 영역(bp 3,655-4,467)을 포함한다.
- <101> 도 3e는 플라스미드 pCAH/SINd3(서열 번호: 70)의 개략도이다. pCAH/SINd3(서열 번호: 70)는 HCMV MIEP(bp 1-588); CAEV 5'LTR에서의 R-U5 서열 영역(bp 610-772); PBS 위치를 포함하는 잔여 비번역 영역(bp 773-789); 시작 ATG 코돈(bp 1121-1123) 및 상기 시작 ATG 코돈의 하류에 위치하는 ATG 코돈(bp 1142-1144)에서 점변이를 갖는, *gag* 유전자의 908 bp의 절편(bp1,121-2,029); RRE 영역(bp 2,049-2,407); 및 U3-삭제된 CAEV 3'LTR 영역(bp 2,549-2,711)을 포함하는 4,819 bp의 플라스미드이다. 상기 벡터는 또한 Col E1 복제 기점(bp 3,116-3,699) 및 카나마이신 내성 유전자 코딩 영역(bp 3,951-4,763)을 포함한다.
- <102> 도 3f는 플라스미드 pCAH/SINd4(서열 번호: 71)의 개략도이다. pCAH/SINd4(서열 번호: 71)는 HCMV MIEP(bp 1-588); CAEV 5'LTR에서의 R-U5 서열 영역(bp 610-772); PBS 위치를 포함하는 잔여 비번역 영역(bp 773-1,120); 시작 ATG 코돈(bp 1121-1123) 및 상기 시작 ATG 코돈의 하류에 위치하는 ATG 코돈(bp 1142-1144)에서 점변이를 갖는, *gag* 유전자의 1198 bp의 절편(bp 1,121-2,319); RRE 영역(bp 2,342-2,700); 및 U3-삭제된 CAEV 3'LTR 영역(bp 2,842-3,004)을 포함하는 5,112 bp의 플라스미드이다. 상기 벡터는 또한 Col E1 복제 기점(bp 3,409-3,992) 및 카나마이신 내성 유전자 코딩 영역(bp 4,244-5,056)을 포함한다.
- <103> 도 3g는 플라스미드 pCAH/SINd1/hlacZ(서열 번호: 79)의 개략도이다. pCAH/SINd1/hlacZ(서열 번호: 79)는 lacZ 리포터 유전자(reporter 유전자)를 발현하는 pCAH/SINd1(서열 번호: 68)로부터 유래된 8,127 bp의 플라스미드이다. 상기 벡터는 두 개의 HCMV MIEP 프로모터 영역(각각 bp 1-588 및 bp 1,866-2,460에 위치); CAEV 5'LTR에서의 R-U5 서열 영역(bp 610-772); PBS 위치를 포함하는 잔여 비번역 영역(bp 773-789); 시작 ATG 코돈(bp 1121-1123) 및 상기 시작 ATG 코돈의 하류에 위치하는 ATG 코돈(bp 1142-1144)에서 점변이를 갖는, *gag* 유전자의 325 bp의 절편(bp 1,121-1,446); *lacZ* 유전자 코딩 서열(bp 2,541-5,711); 및 U3-삭제된 CAEV 3'LTR 영역(bp 5,782-6,019)을 포함한다. 상기 벡터는 또한 Col E1 복제 기점(bp 6,424-7,007), 및 카나마이신 내성 유전자 코딩 영역(bp 7,259-8,071)을 포함한다.
- <104> 도 3h는 플라스미드 pCAH/SINd60/hlacZ(서열 번호: 78)의 개략도이다. 플라스미드 pCAH/SINd60/hlacZ(서열 번호: 78)는 두 개의 프로모터 영역 HCMV MIEP(각각 bp 1-588 및 bp 1,595-2,189에 위치); CAEV 5'LTR에서의 R-U5 서열 영역(bp 610-772); PBS 위치를 포함하는 잔여 비번역 영역(bp 773-789 bp); 시작 ATG 코돈(bp 1121-1123) 및 상기 시작 ATG 코돈의 하류에 위치하는 ATG 코돈(bp 1142-1144)에서 점변이를 갖는, *gag* 유전자의 60 bp의 절편(bp 1,121-1,181); RRE 영역(bp 1,195-1,565); *lacZ* 유전자 코딩 서열(bp 2,270-5,440); 및 U3-삭제된 CAEV 3'LTR 영역(bp 5,511-5,748)을 포함하는 7,856 bp이다. 상기 벡터는 또한 Col E1 복제 기점(bp 6,153-6,736), 및 카나마이신 내성 유전자 코딩 영역(bp 6,988-7,800)을 포함한다.
- <105> 도 4는 플라스미드 pHYK/vif(서열 번호: 76)의 개략도이다. pHYK/vif(서열 번호: 76)는 HCMV MIEP(bp 1-596),

vif 유전자 코딩 영역(bp 691-1,380), BGH 폴리아데닐화 신호(bp 1,467-1,695), Col E1 복제 기점(bp 1,826-2,409), 네오마이신 내성 유전자 코딩 영역(bp 3,862-2,866), 및 암피실린 내성 유전자 코딩 영역(bp 5,270-4,239)을 포함하는 5,729 bp의 플라스미드이다.

- <106> 도 5는 플라스미드 pHYK/rev(서열 번호: 75)의 개략도이다. pHYK/rev(서열 번호: 75)는 HCMV MIEP(bp 1-596), *rev* 유전자 코딩 영역(bp 672-1,073), BGH 폴리아데닐화 신호(bp 1,157-1,385), Col E1 복제 기점(bp 1,516-2,099), 네오마이신 내성 유전자 코딩 영역(bp 3,552-2,556), 및 암피실린 내성 유전자 코딩 영역(bp 4,960-3,929)을 포함하는 5,419 bp의 플라스미드이다.
- <107> 도 6a는 플라스미드 pHGVSV-G(서열 번호: 74)의 개략도이다. pHGVSV-G(서열 번호: 74)는 HCMV MIEP(bp 1-596), β -글로빈 인트론 영역(bp 714-1,599), VSV-G 코딩 영역(bp 1,632-3,312), BGH 폴리아데닐화 신호(bp 3,361-3,589), Col E1 복제 기점(bp 3,720-4,303), 네오마이신 내성 유전자 코딩 영역(bp 5,756-4,760), 암피실린 내성 유전자 코딩 영역(bp 7,164-6,133), 및 F1 복제 기점(bp 7,165-7,621)을 포함하는 7,623 bp의 플라스미드이다.
- <108> 도 6b는 플라스미드 pMYKEF1/env(서열 번호: 72)의 개략도이다. pMYKEF1/env(서열 번호: 72)는 MCMV MIEP (bp 1-665), 인간 EF1- α 인트론 영역(bp 668-1,618), GaLV *env* 코딩 영역(bp 1,699-3701), BGH 폴리아데닐화 신호 (bp 3,885-4,118), Col E1 복제 기점(bp 4,349-4,832), 네오마이신 내성 유전자 코딩 영역(bp 6,290-5,284), 및 암피실린 내성 유전자 코딩 영역(bp 7,496-6,666)을 포함하는 7,579 bp의 플라스미드이다.
- <109> 도 7은 인간 293T 표적 세포로 형질감염된 유전자 전이 벡터로부터 전사된, 전이 벡터 RNA의 상대량을 나타내는 사진이다.
- <110> 도 8은 CAEV (A) 및 MuLV (B) 벡터에 의해 유전자 전이시킨 인간 293T 표적 세포를 나타내는 두 개의 사진이다.
- <111> 도 9는 형질감염된 293T 세포에서 발현되거나(라인 1, 2 및 3) 293T 패키징 세포에서 캡시드화되고 방출된(라인 4, 5 및 6), 전이 벡터 RNA의 상대량을 나타내는 사진이다.
- <112> 도 10은 인간 293T 패키징 세포에서 캡시드화되고 방출된 전이벡터 RNA의 상대량을 나타내는 사진이다.
- <113> 도 11은 VSV-G 또는 GaLV 외피 단백질에 의해 의사형화된(pseudotyped) 렌티바이러스 벡터의 감염 및 역전사 후, 통합된 레트로바이러스 cDNA의 상대량을 나타내는 사진이다.
- <114> 도 12는 감염된 숙주 세포 염색체(chromosome)내로 통합된 바이러스 벡터 cDNA의 상대량을 나타내는 사진이다.
- <115> 도 13은 (A) 대조군 세포 및 (B) G1-정체된 세포에 대한 FACS 분석 결과를 나타내는 두 개의 그래프이다.
- <116> 도 14는 (A) 형질도입된 세포의 수 및 (B) 분열 및 비분열 세포에 대한 HIV-1-, CAEV-, 및 MuLV-유래 바이러스 벡터의 상대적 형질도입 효율성을 나타내는 두 개의 그래프이다.

도면

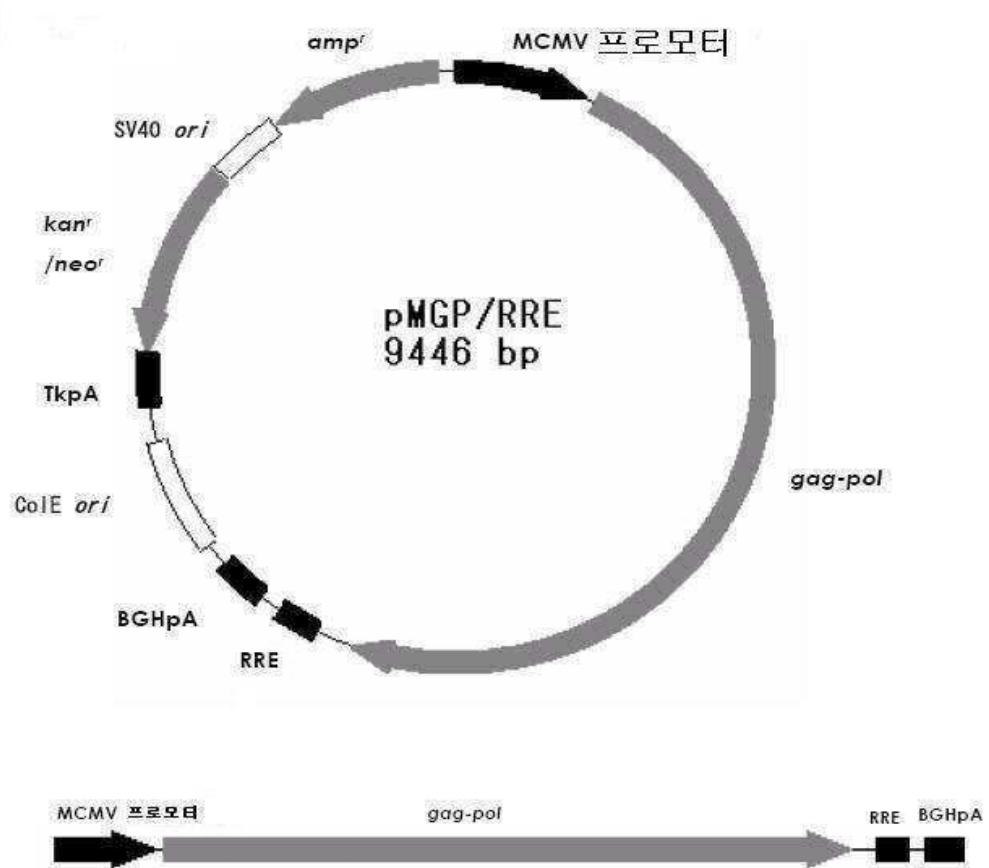
도면1



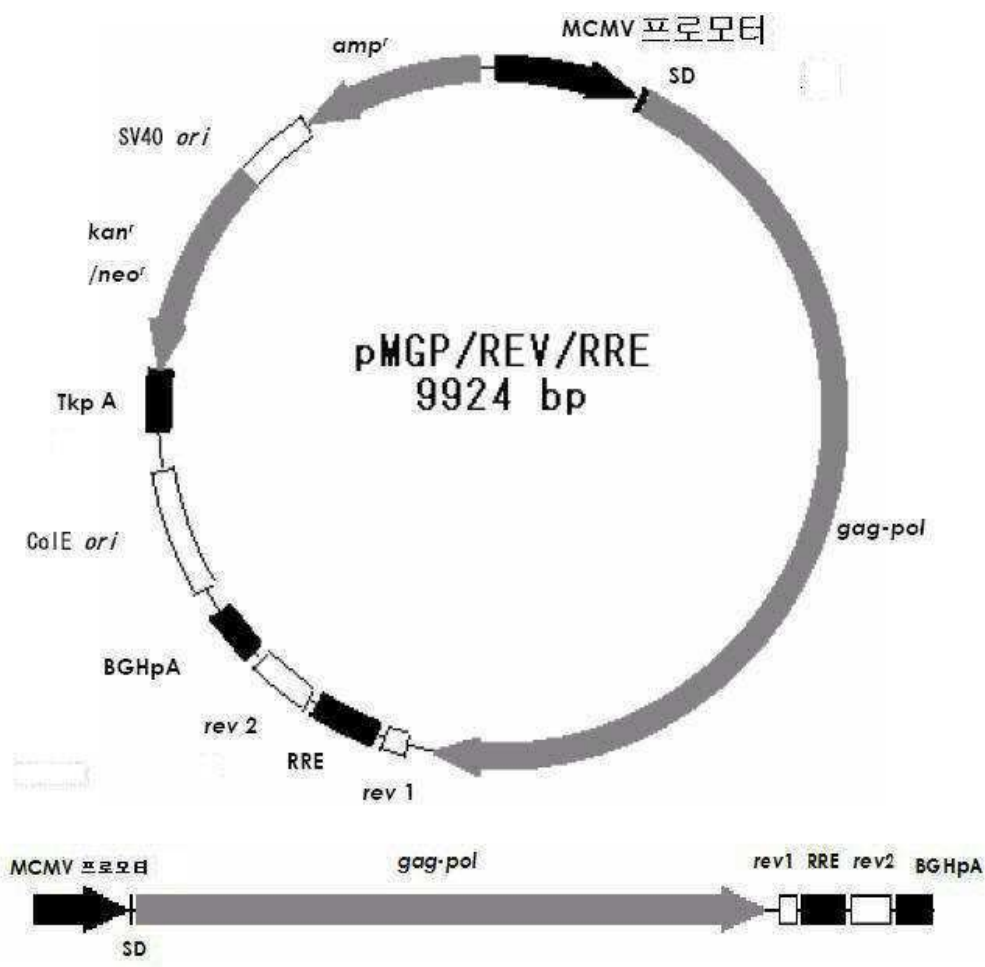
CAEV 유전체 RNA의 뉴클레오티드 서열 [전장크기 9189 bp]

R : 1-85
 U5 : 86-163
 PBS : 164-180
 SD : 346
 Gag : 512-1858
 Pol : 1729-5046
 Vif : 5006-5695
 Tat : 5688-5951
 Env : 6012-8846
 SD : 6123
 SA : 8513
 Rev : 6012-6123, 8514-8803
 RRE : 7850-8150
 PPT : 8796-8818
 U3 : 8819-9104
 R : 9105-9189

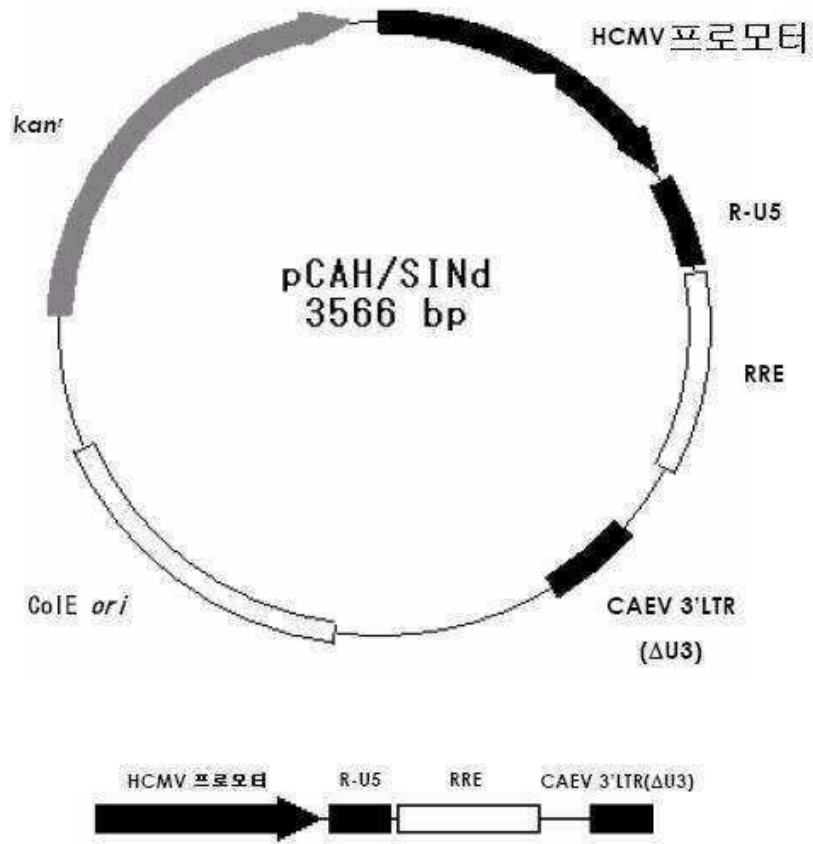
도면2a



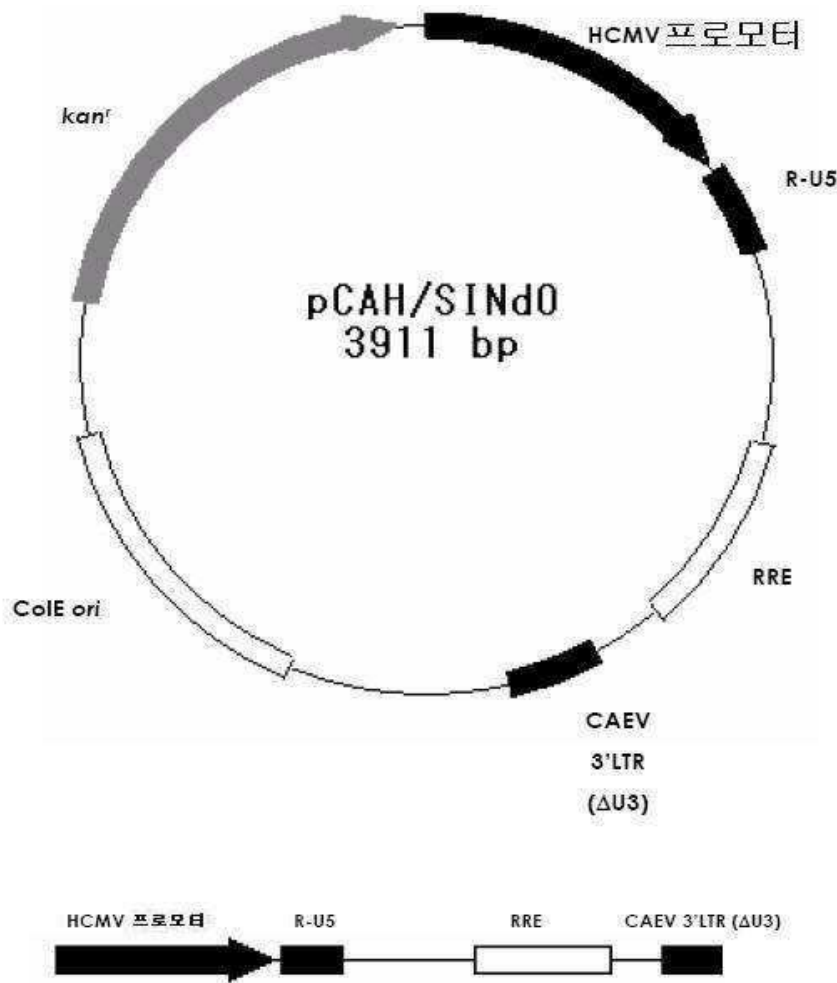
도면2b



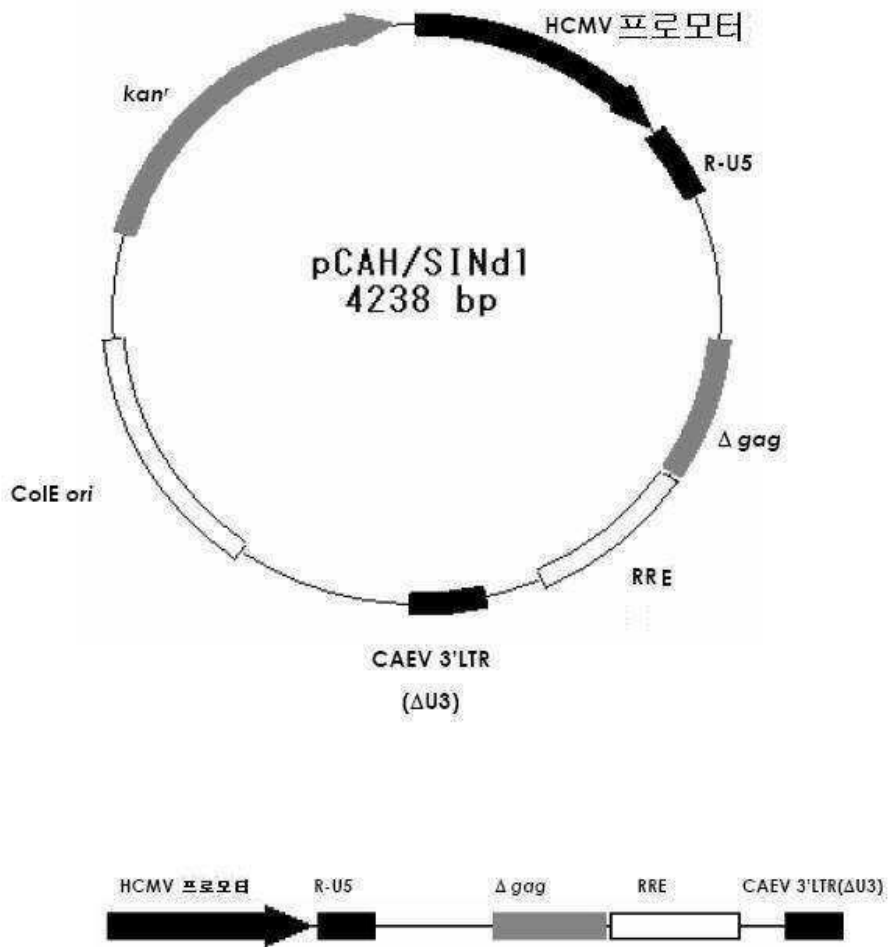
도면3a



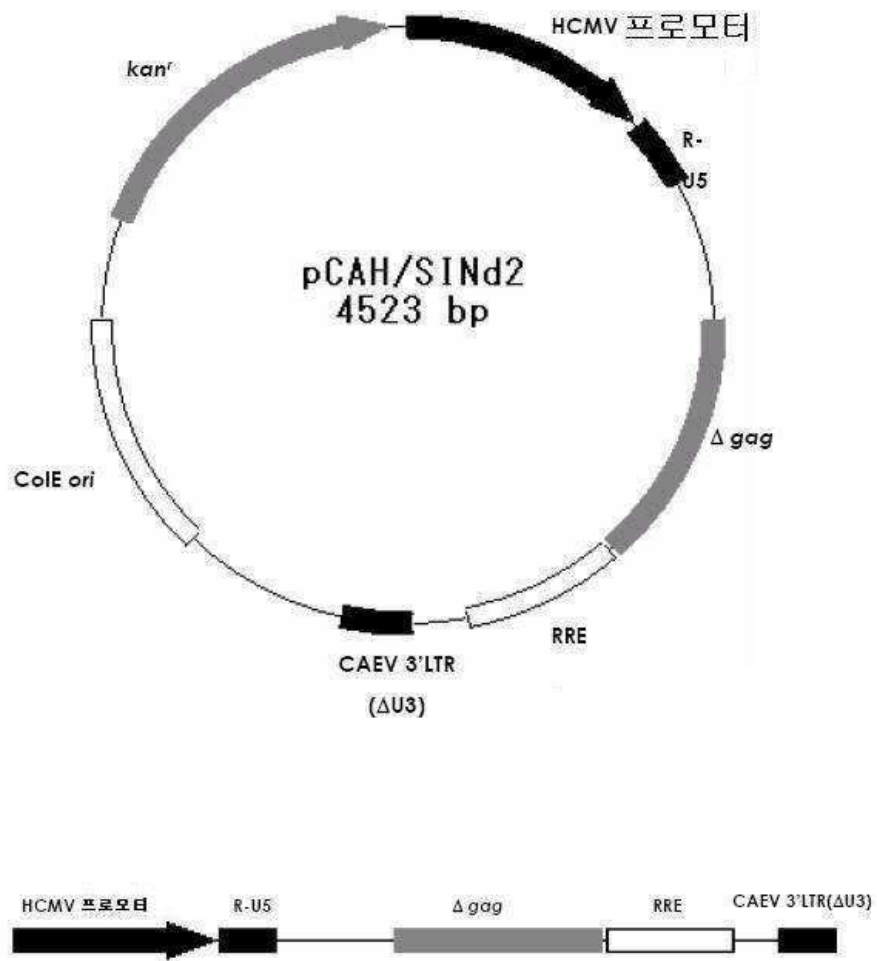
도면3b



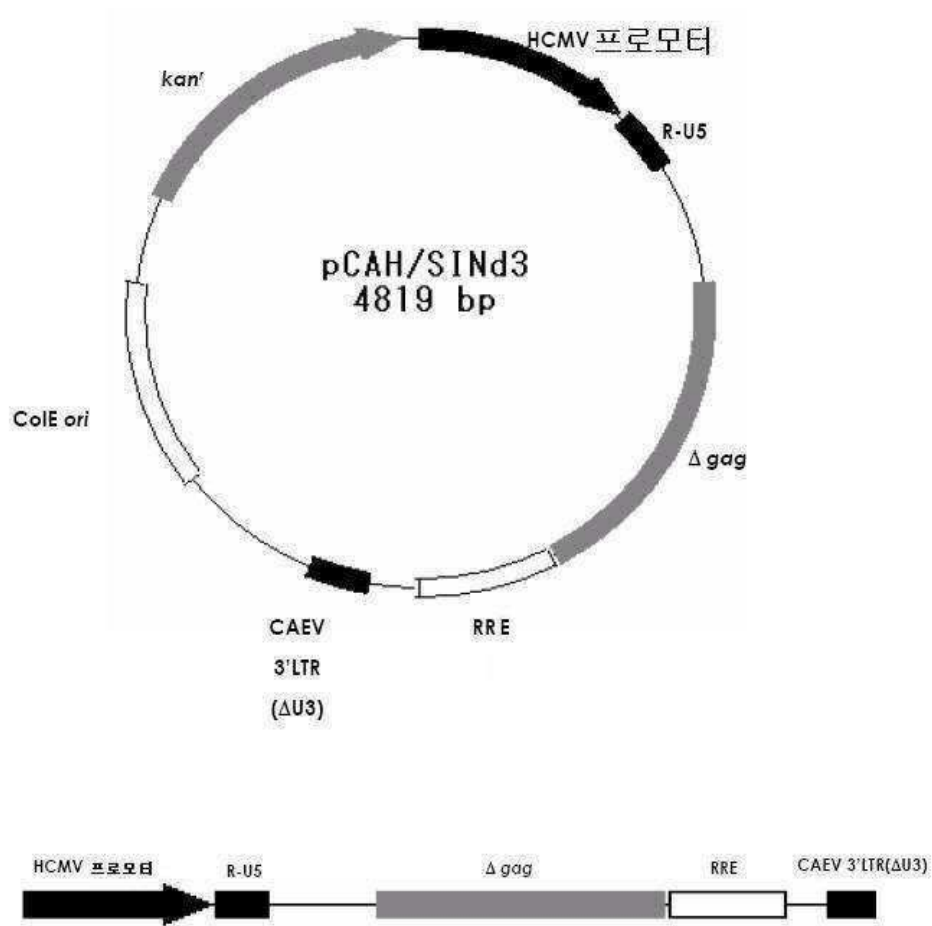
도면3c



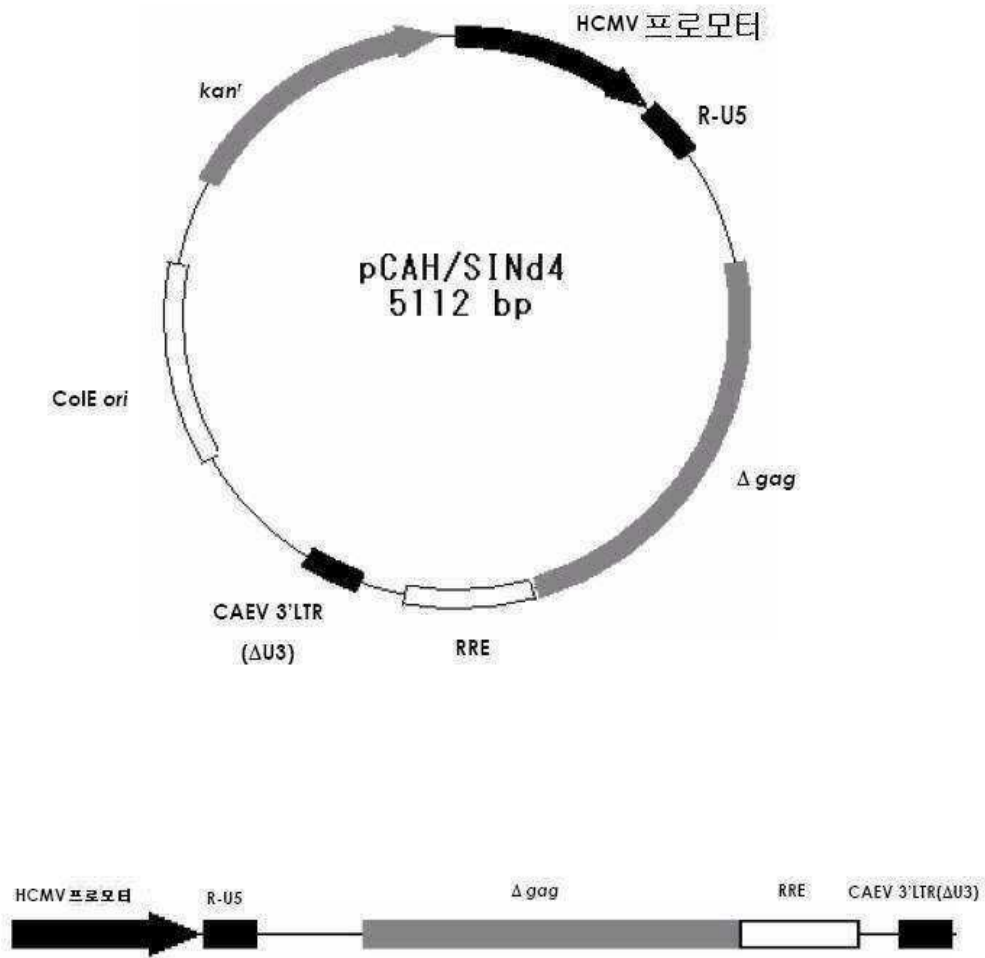
도면3d



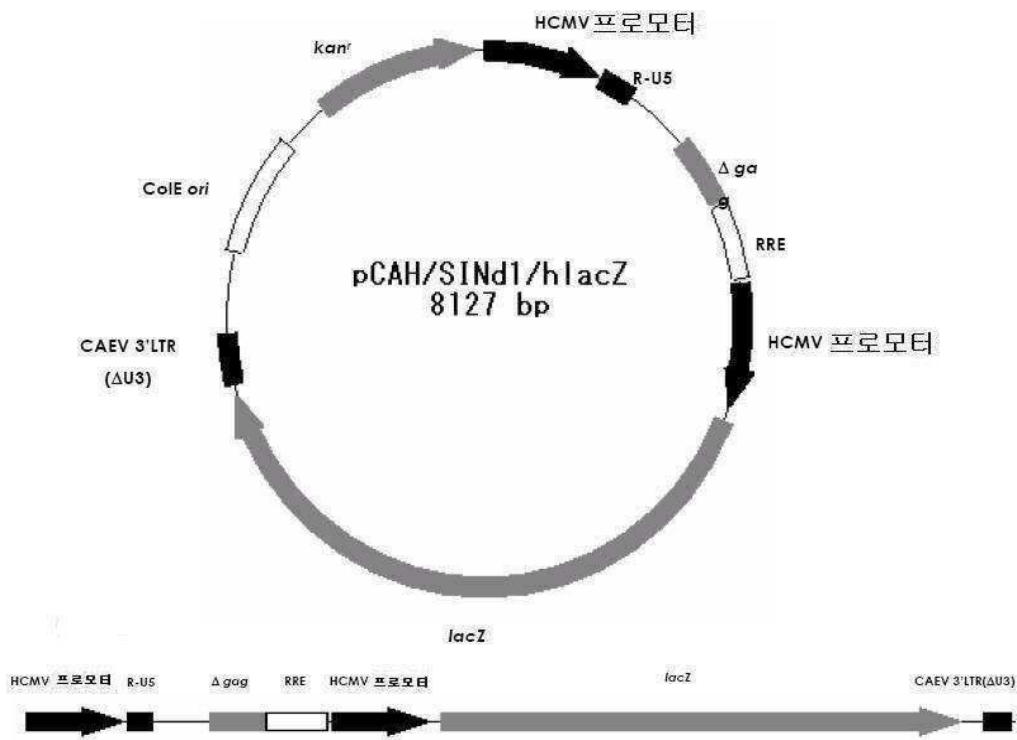
도면3e



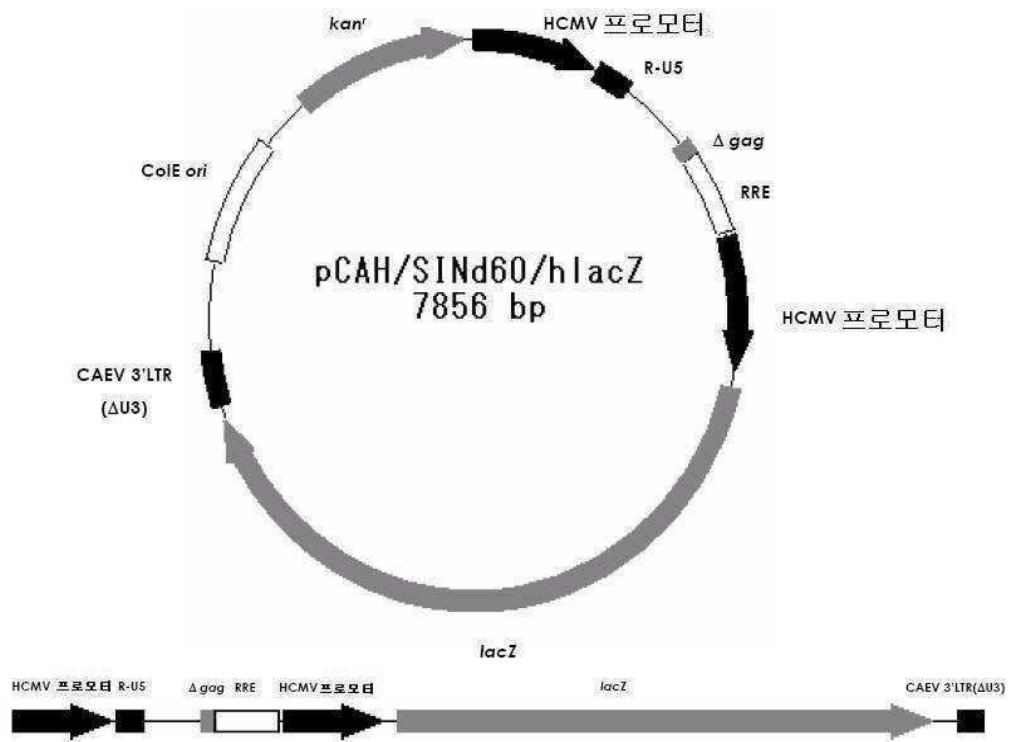
도면3f



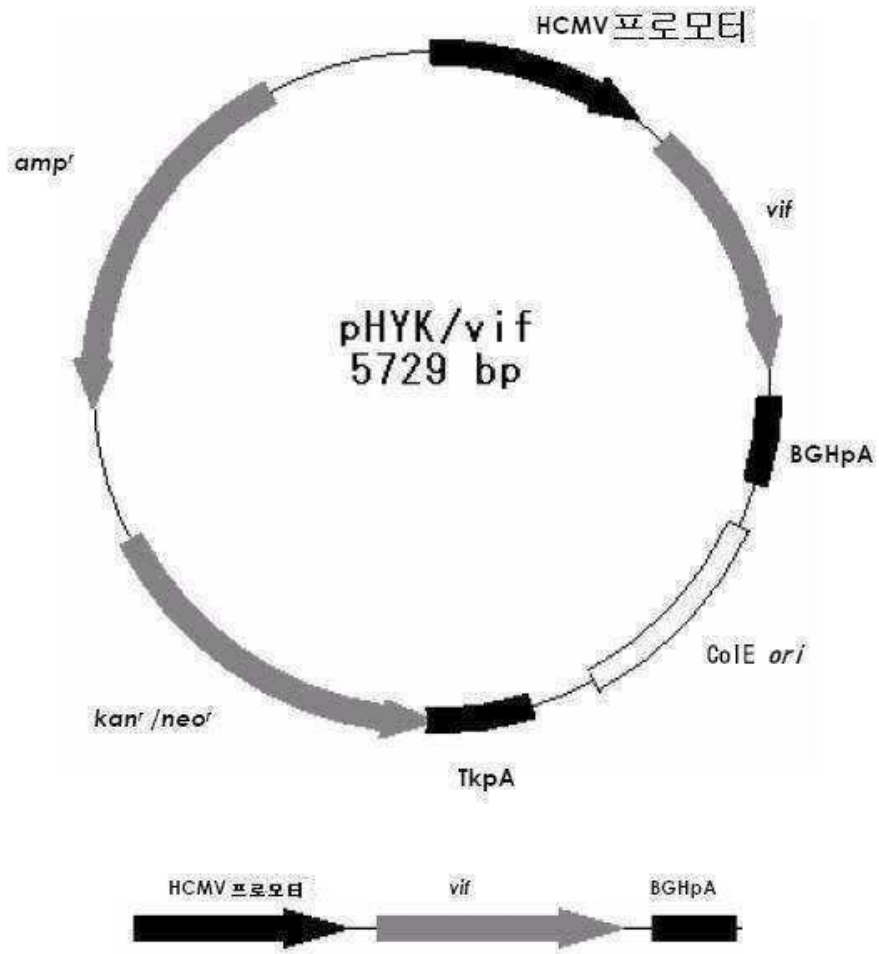
도면3g



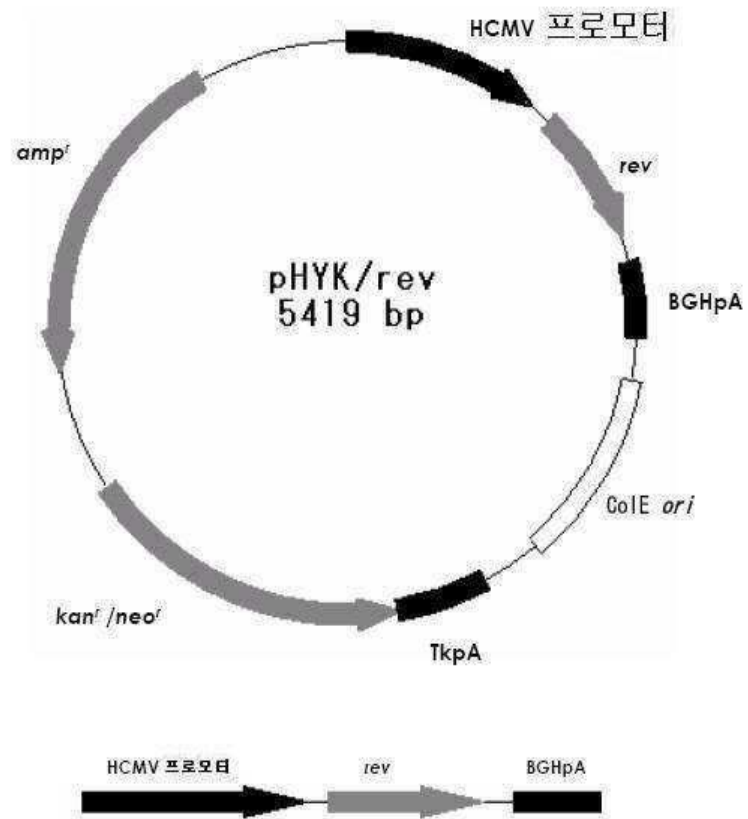
도면3h



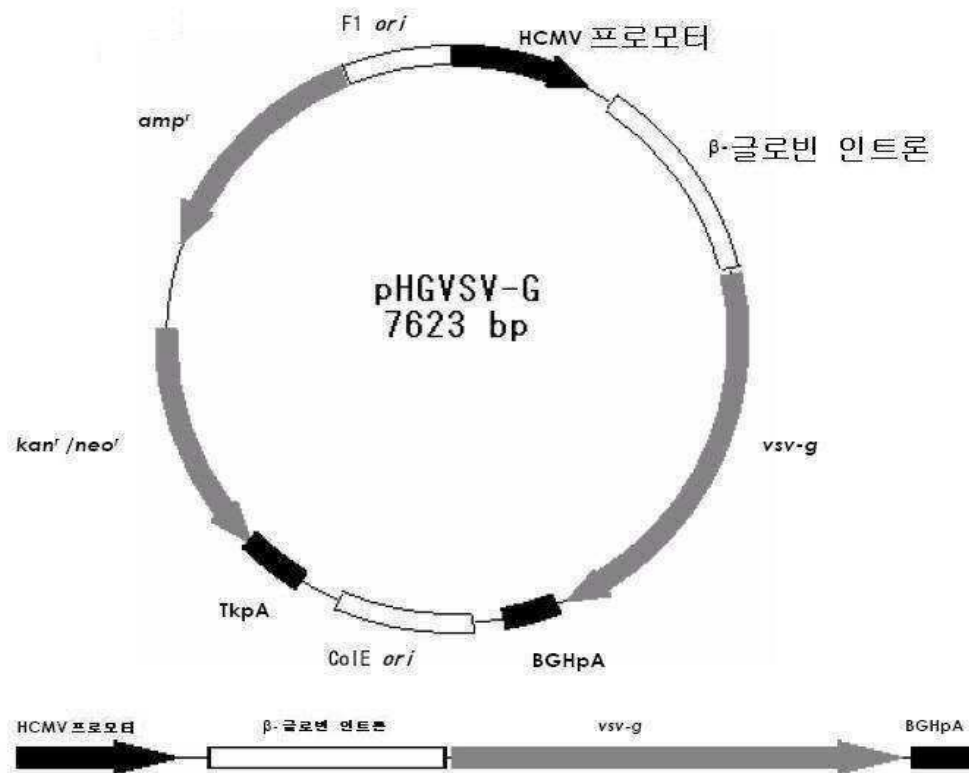
도면4



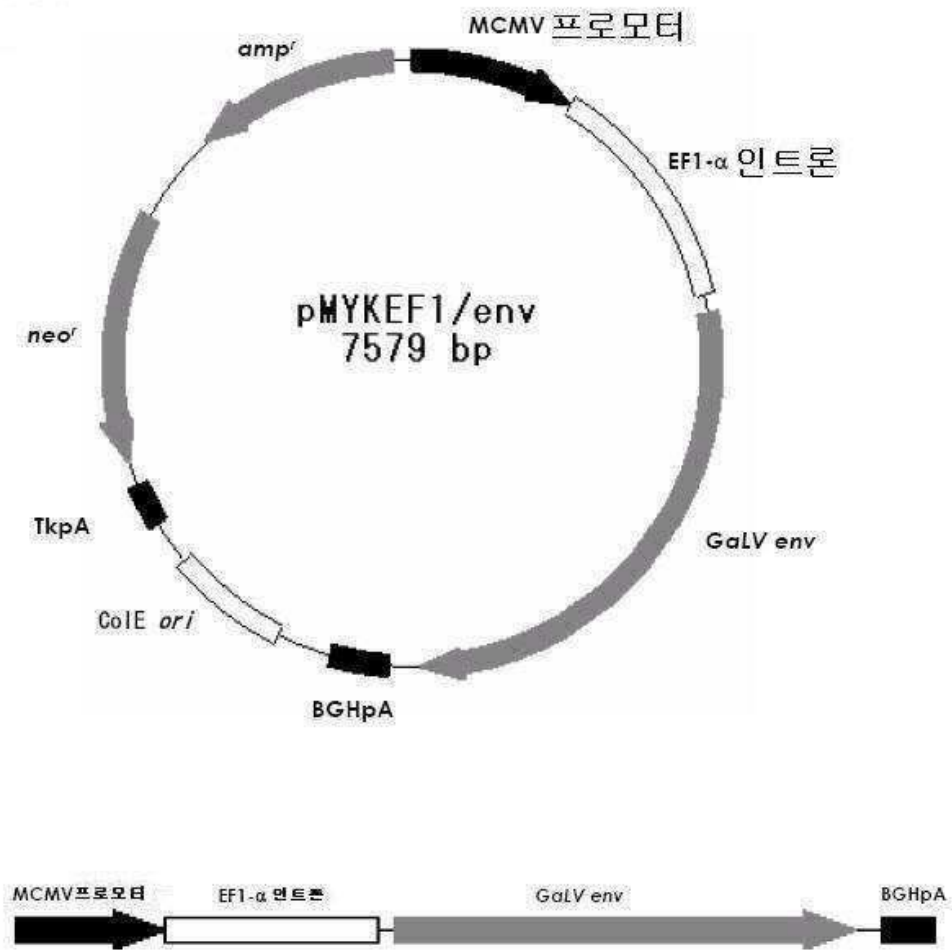
도면5



도면6a



도면6b

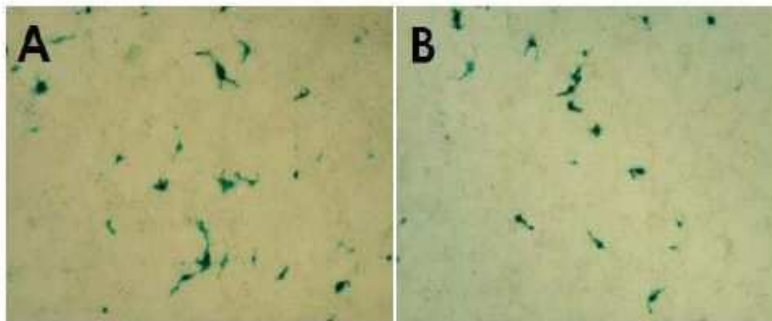


도면7



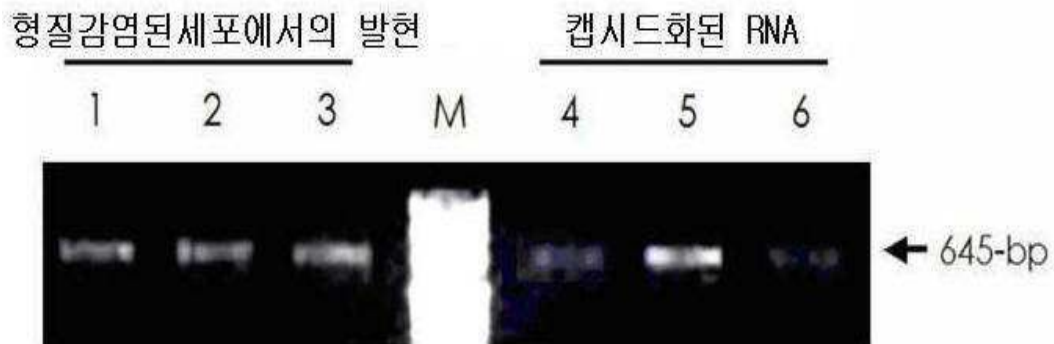
1. pCAH/SINd
2. pCAH/SINd0
3. pCAH/SINd1
4. pCAH/SINd2
5. pCAH/SINd3
6. pCAH/SINd4
7. pHR'/lacZ

도면8



A. pMGP/RRE, pCAH/SINd60/hlacZ, pHYK/rev, pHYK/vif, pHGVSV-G
 B. pEQPAM, pMFG/lacZ/puro, pHGVSV-G

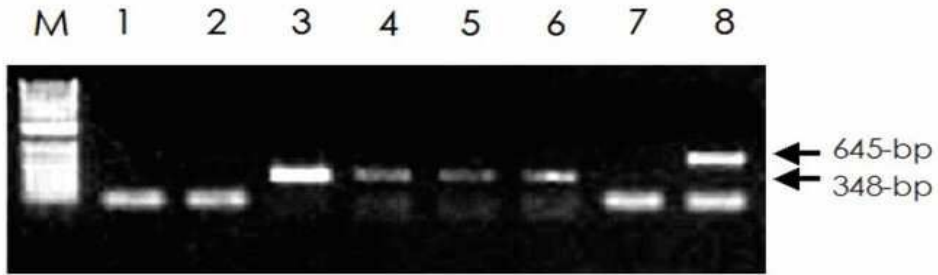
도면9



M. 사이즈 마커

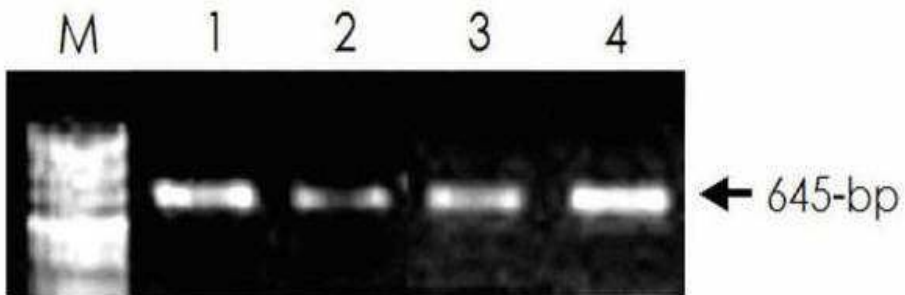
- 1, 4. pMGP/RRE, pCAH/SINd1hlacZ, pHGVSV-G
- 2, 5. pMGP/RRE, pCAH/SINd1hlacZ, pHGVSV-G, pHYK/rev
- 3, 6. pMGP/RRE, pCAH/SINd1hlacZ, pHGVSV-G, pHYK/rev, pHYK/vif

도면10



1. pMGP/RRE, pCAH/SINd, pHGVSV-G, pHYK/rev, pHYK/vif
2. pMGP/RRE, pCAH/SINd0, pHGVSV-G, pHYK/rev, pHYK/vif
3. pMGP/RRE, pCAH/SINd1, pHGVSV-G, pHYK/rev, pHYK/vif
4. pMGP/RRE, pCAH/SINd2, pHGVSV-G, pHYK/rev, pHYK/vif
5. pMGP/RRE, pCAH/SINd3, pHGVSV-G, pHYK/rev, pHYK/vif
6. pMGP/RRE, pCAH/SINd4, pHGVSV-G, pHYK/rev, pHYK/vif
7. pCAM/lacZ(L)
8. pCMVΔR8-2, pHR' /lacZ, pHGVSV-G

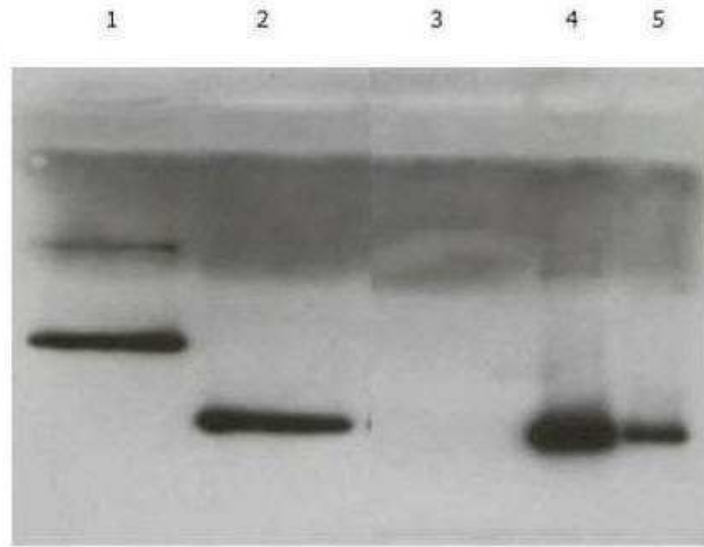
도면11



M. 사이즈 마커

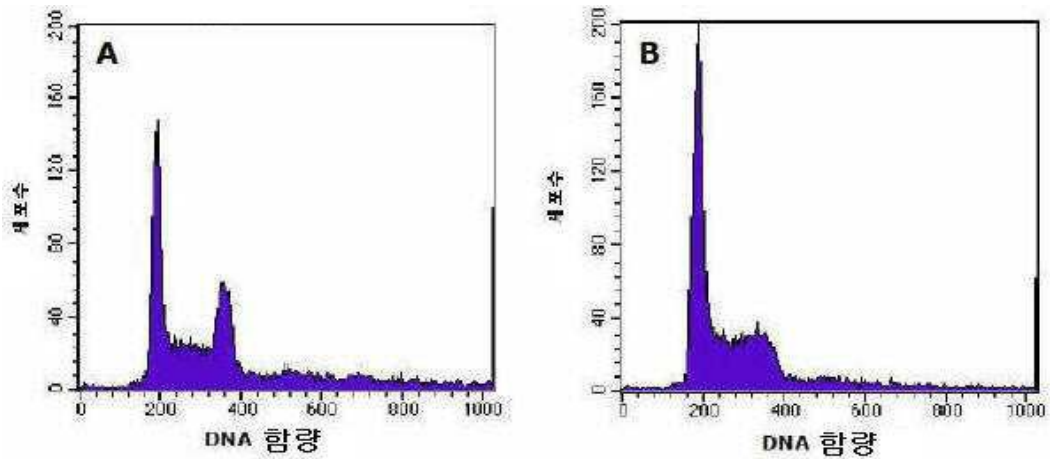
1. pMGP/RRE, pCAH/SINd1lacZ, pHGVSV-G, pHYK/rev, pHYK/vif
2. pMGP/RRE, pCAH/SINd1lacZ, pMYKEF1/env, pHYK/rev, pHYK/vif
3. pEQPAM3, pMFG/lacZpuro, pHGVSV-G
4. pCMVΔR8-2, pHR' /lacZ, pHGVSV-G

도면12

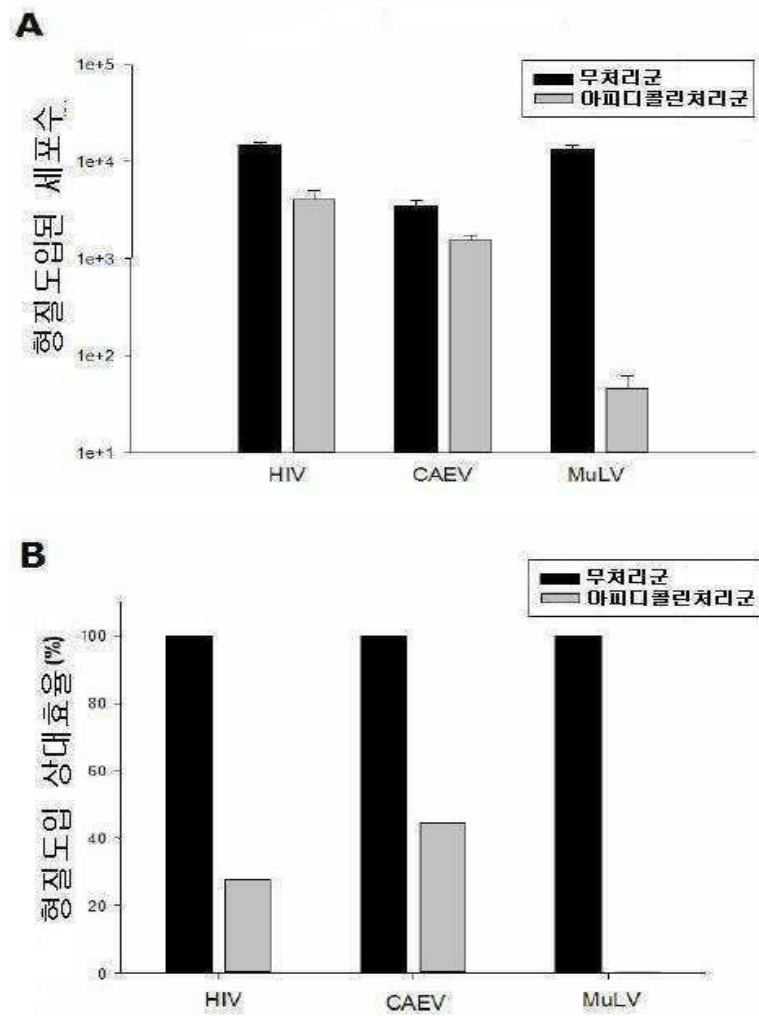


1. HIV-1-계 전이벡터에 대한 3.15kb BamH I-Kpn I 절편
2. CAEV-계 전이벡터에 대한 1.35kb Hind III-Ssp I 절편
3. 음성대조군
4. 양성대조군(3ng)
5. 양성대조군(0.3ng)

도면13



도면14



서열목록

SEQUENCE LISTING

<110> MACROGEN CO., LTD

<120> CAEV-BASED VECTOR SYSTEMS

<160> 79

<170> KopatentIn 1.71

<210> 1

<211> 9189

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 1

gagttctagg agagtcctc ctagtctctc ctctccgagg aggtaccgag acctcaaaat 60

aaaggagtga ttgccttact gccgagtga gagtgattac tgagcggccg gtgtatcggg 120

agtcgtccct taatctgtgc aataccagag cggctctcgc agctggcgcc caacgtgggg 180

cccgaggaga agaaaagaaa gcggccctga gaactcggct tctgaaaaag aggaagagga 240

caagttgcta tagcaacaag agagaagaag tagagcaaag gtccagtggc tcggaaaaag 300

aggaactgaa acttcgggga cgcctgaagg agtaaggtaa gtgactctgc tgtacgctgg 360

gcgaggcaga ggtttccttc taaattgaaa gagaagtgtt gctgcgagag gtcttggtgg 420

tcgagaatcc tgiacaaaa aaaggagga tctcggtcag gaccaggacc cctgggagta 480

atacaacagc aacaccgtaa gaaaatccgc catgggtgagt ctagatagag acatggcgag 540

gcaagtctcc ggggggaaaa gagattatcc tgagctcga aatgtatca agcatgcatg 600

caagataaaa gttcgaactca gaggggagca cttgacagaa ggaaattgtt tatggtcct 660

taaaacatta gattacatgt ttgaggacca taaagaggaa ccttggacaa aagtaaatt 720

taggacaata tggcagaagg tgaagaatct aactcctgag gagagtaaca aaaaagactt 780

tatgtctttg caggccacat tagcgggtct aatgtgttc caaatgggga tgagacctga 840

gacattgcaa gatgcaatgg ctacagtaat catgaaagat gggttactgg aacaagagga 900

aaagaaggaa gacaaaagag aaaaggaaga gagtgtcttc ccaatagtag tgcaagcagc 960

aggagggaga agctggaaag cagtagattc tghtaatgttc cagcaactgc aaacagtagc 1020

aatgcagcat ggctctgtgt ctgaggactt tgaaaggcag ttggcatatt atgctactac 1080

ctggacaagt aaagacatac tagaagtatt ggccatgatg cctggaaata gagctcaaaa 1140

ggagttaatt caagggaat taaatgaaga agcagaaagg tggagaagga ataatccacc 1200

acctccagca ggaggaggat taacagtgga tcaaattatg ggggtaggac aaacaaatca 1260
 agcagcagca caagctaaca tggatcaggc aaggcaaata tgcctgcaat gggaataaaa 1320
 tgcattaaga gcagtaagac atatggcgca caggccaggg aatccaatgc tagtaaagca 1380
 aaaaacgaat gagccatatg aagattttgc agcaagactg ctagaagcaa tagatgcaga 1440
 gccagttaca cagcctataa aagattatct aaagctaaca ctatcttata caaatgcatc 1500
 agcagattgt cagaagcaaa tggatagaac actaggacaa agagtacaac aagctagtgt 1560
 agaagaaaa atcaagcat gtagagatgt gggatcagaa gggttcaaaa tgcaattggt 1620
 agcacaagca ttaaggccag gaaaaggaaa agggaatgga cagccacaaa ggtgttacia 1680
 ctgtgaaaa ccgggacatc aagcaaggca atgtagacaa ggaatcatat gtcacaactg 1740
 tggaaagaga ggacatatgc aaaaagaatg cagaggaaag agagacataa ggggaaaaca 1800
 gcagggaaac gggaggaggg ggatcgtgt ggtgccgtcc gctcctccta tggataaact 1860
 tcagcaccac ctatggttca ggtccgcata ggttcccagc agaggaactt gttatttgat 1920
 accggggcgg accgaactat agttagatgg catgagggtc cgggaaaccc agccggaagg 1980
 ataaaactgc aaggaatagg aggaatagta gaaggagaaa aatggaataa ttagaatta 2040
 gaataaaag gagaacaag aaaggaaca atagtagtgt taccacaaag tccagtagaa 2100
 gtattaggac gagataacat ggcccgattt ggaataaaga taataatgac aaatttagag 2160
 gaaaaaagaa tccaattac aaaagtaaaa ttgaaagagg gatgtacggg tccacatgac 2220
 ccacaatggc cattaacaga agagaaatta aaaggtctaa cagaaatcat agataaatta 2280
 gtggaagaag gaaaactagg aaagccacc ccacattgga catgtaatac tccaatcttt 2340

tgcataaaaa agaaatcagg gaagtggaga atgttaatag atttcagaga attgaacaaa 2400
 cagacagaag atttaacaga agcgcagtta ggactcccgc atccgggagg actacaaaag 2460
 aaaaaacatg ttacaatatt ggacatagga gatgcatatt ttactatacc cctatatgaa 2520
 ccatatcgag agtiacacatg ttttactctia ttaagtccta ataatctagg accatgtaaa 2580
 agatactatt ggaaagtgct gccacaaggt tggaaattga gtccatctgt atatcaattt 2640
 actatgcagg agatcttaga ggattggata cagcagcatc cagaaattca atttggcata 2700
 tatatggatg atatttcat aggaagtgat ttagaaatta aaaagcatag agaaatagtg 2760
 aaagatttag ccaattatat tgcccaatat ggattcactc tgccagaaga gaagagacaa 2820
 aagggatatc cagcaaaatg gctaggattt gaactacacc cgagacctg gaaatttcag 2880
 aagcatacat tacctgaatt aacaaagga acaataacat taaataaatt acagaaatta 2940
 gtaggagaat tagtatggag acaatccata attgggaaaa gcattcctaa cattctgaaa 3000
 ttaatggaag gagatagaga attacaaagt gaaagaaaaa ttgaagaagt acatgtgaaa 3060
 gaatgggaag catgtaggaa aaaattagaa gaaatggaag gaaattatta taataaagac 3120
 aaagatgtct atggacaatt ggcttgggga gacaaagcta tagaatatat agtgtatcag 3180
 gagaaagga aaccattatg ggtaaatgtg gttcacaata taaagaacct aagcatcccc 3240
 caacaggtta ttaaagcagc gcaaaaatta acccaagaag tcatcattag gacaggaaaa 3300
 ataccatgga tattgttgcc agggaaagaa gaagattgga gactagaatt gcaattaggg 3360
 aacatcacat ggatgcaaaa attttgggcc tgttatcgag gacatacaag atggagaaaa 3420
 agaaatataa tagaagaagt agtagaaggg cctacatatt atacagatgg aggaaaaaag 3480
 aataaagtag gaagtctagg gttcatagta tcaacagggg aaaaatttag aaagcatgaa 3540

gagggcacia accagcaact agaattaaga gccatagagg aagctctaaa acaagggcct 3600

caaacaatga atttagtaac agatagtaga tatgcatttg aatTTTTatt aagaaattgg 3660

gatgaagaag taataaagaa tccaattcaa gcaagaatta tggaaattgc ccacaagaaa 3720

gataggatag gagtgcattg ggtgccagga cataaagga ttccccaaa tgaagaaata 3780

gacaaatata tttcggaat atttcttga aaagaaggag aaggaattct cccaaaaaga 3840

gaagaggatg cagggtatga tttaatatgc ccagaagagg ttacataga gccaggacia 3900

gtgaaatgca tccccataga gctaagatta aatttaaaga aatcacaatg ggctatgatt 3960

gctacaaaaa gcagcatggc tgccaaagga gtgttcacac aaggaggaat catagactca 4020

ggatatcagg gacaaataca ggtaataatg tataatagca ataaaatagc agtagtcata 4080

ccccaaagga gaaaatttgc acaattaata ttaatggata aaaagcatgg aaaattggaa 4140

ccctgggggg aaagcagaaa aacagaaagg ggagaaaaag gatttgggtc tacaggaatg 4200

tattggatag aaaatattcc tctggcagag gaagaccaca caaatggca tcaagatgcc 4260

cgatcattgc atctagaatt tgaaattcca agaacagcag cagaagacat agtaaatcaa 4320

tgtgaaatat gcaagaagc gaggacacct gcagtaatta gaggcggaaa caaaaggggg 4380

gtaaatcatt ggcaagtgga ttataccat tatgaaata tcatactatt agtatgggta 4440

gaaacaaatt caggactaat atatgcagaa aaagtaaaag gagaatcagg gcaagaattc 4500

agaataaaag tgatgcattg gtatgatta tttggtccag agtcattgca gtcagacaat 4560

ggacctgcat ttgcagcaga gcccacacag ctgttaatgc aatacctagg agtaaaacac 4620

acaacaggca taccttggaa tccacagtct caggctatag tagaaagggc acatcaacta 4680

ttgaaaagca ctttaaagaa gttccagcca caatttgtcg ctgtagaatc agccatagca 4740
 gcagccctag tcgccataaa tataaaaaga aagggtgggc tggggacaag ccctatggat 4800
 atttttatat ataataaaga acagaaaaga ataataata aatataataa aaattctcaa 4860
 aaaattcaat tcigtattata cagaataagg aaaagaggac atcaggagag tggaaaggac 4920
 caaccagggt actgtgaaa ggggaaggag ccaattgtgg taaaggatat agaaagtga 4980
 aagtatttag taatacctta caaagatgca aaattcatcc cgccaccaac aaaagaaaag 5040
 gaataaaaaa cctggaccag aattaccctt agcactatgg atacatatag cagaaagcat 5100
 taatggggat agctcatggt acataacaat gagactgcaa cagatgatgt ggggaaaaag 5160
 aggaaataag ttacaatata agaatgaaga cagggaatat gaaaattggg aaattacatc 5220
 atggggatgg aaaaatgcacc taaggagagt gaaacaatgg atacaagaca acaggagagg 5280
 aagcccatgg cagtacaag taggaggaac atggaaaagt ataggagtgt ggttctgca 5340
 agcaggagat tacagaaagg tagacaggca cttctggtgg gcatggagga tactgatatg 5400
 ttctgcagg aaagaaaagt ttgatataag agaatttatg agaggaagac atagatggga 5460
 tttgtcaaa tcctgtgctc aaggagaagt agtaaagcat actagaaca aaagtctgga 5520
 aagactagta ctgctacaga tggtagaaca gcatgtgttt caagtattgc catttggag 5580
 agccaggaga agtagtaca cagatttccc atggtgcagg gacacaacgg gatacacgca 5640
 tgcgtggtct gtccaggagt gctggttgat ggaatatctc ttagaggatg agtgaagaac 5700
 tgctcaaag aaggagaca catccagaag aacttgaag gaacgtacgg gaaagagaaa 5760
 gggatacatg gcaatggaca agcatcagag tacctgcgga aatactgcaa agatggcttg 5820
 ctatgcttag gtcaggcaga aatagaaaaga aagtgtatag agaaatgcaa aaatggatgt 5880

ggatacatcc caagcgccct gtgattaggg cctgtggatg cagactatgt aacccggggt 5940
 ggggaacata atcaaggga taataaatgc aaataaatgt aactaacaag tagcaaaagt 6000
 gtctgtgta gatggatgct ggggccagat acatgcgctt aactgggaag gaaaactggg 6060
 ttgaagtaac catggacgga gagaaggaaa ggaaaagaga aggtttcact gcgggacagc 6120
 aaggtaahta tcaacccag gtaagtaagc aaataggga cagaaatact aacccatgct 6180
 ttgcctataa agggatattc ctatggagga taccactaac aatgtggata ttgctagga 6240
 taaatagtgt tgtcagtga gaggattaca taactaat atcagatccc tatgggttct 6300
 cacccataaa aaatgtgtct ggggtaccag tgacttgtgt aacaaaagaa ttcgcaaat 6360
 ggggatgca accactagga gcgtaccctg atccagaaat agaatacaga aatgtgagtc 6420
 aggaagtagt gaaagaagta tatcaagaga attggccatg gaatacatat cattggcctc 6480
 tctggcaaat ggagaatgtt aggtactggt taaaagaaaa tatgcaagaa aatcaacaga 6540
 gaaaaataa taaaaagag ggtatagagg aattattagc aggaactata aggggaagat 6600
 tctgtgtacc ataccattt gccttgtaa aatgcacaaa gtggtgctgg tatacagcgg 6660
 ccataaaca cgagtcagga aaagcaggaa aaataaaaat aaatgcaca gaagcaagag 6720
 cagtctctg tacagaggac atgccattag cctcaatata aagagcatat tgggatgaga 6780
 aagacagaga gagcatggcc tttatgaata tcaaagcatg tgatagcaac ctaaggtgtc 6840
 agaaaagacc tggagggtgt atggaaggat accctatccc agtaggagca gaataatcc 6900
 ctgaaagtat gaaataccta aggggagcaa agagtcagta tgggggaata aaagataaga 6960
 atggagaatt aaaattacca ttaacattaa gagtgtgggt aaaattagca aatgtgtcag 7020

aatgggtaaa tgggacaccc ccggattggc aagacagaat taacgatcc aaaggaataa 7080

atgggacgct ctggggagag cttaacagta tgcatacct aggatttgcc cttagccaga 7140

acggcaaatg gtgtaactac accggggaaa taaaattagg gcaagaaaca ttccaatc 7200

attacaagcc aaactggaac tgtaccggga attggacgca atatccggtg tggcaagtga 7260

ttagaaacct ggatatggtg gaacatatga caggagaatg tgtgcagaga ccacaaaggc 7320

acaatataac agtaggaat ggaaccataa cagggaattg cagtacaaca aactgggatg 7380

gatgtaattg ctcacgatca ggaaactacc tatataacag ctctgaggga ggattgttat 7440

taattctgtg cagacaaaac agcacctaa caaggatcct gggaacaaat acaaattgga 7500

caactatgtg gggaatatac aaaaattgtt caggatgcca gaatgcaaca ttagacaaca 7560

caggagaagg aaccttagga ggtgtagcta ataagaactg tagcttgctt cataaaaatg 7620

agagcaaca gtggacttgt gccccaagac aaagagatgg aaaaacagat tcgctataca 7680

tagcaggagg aaaaaagttt tggacacgaa ttaaggcca attcagctgt gaaagtaaca 7740

taggacaatt agatggaatg ttgcatcagc aaatactatt gcaaaaatat caagtaatta 7800

aggtaagagc ttatacatat ggggtgatag aaatgccaga aaactatgca aaaacaagaa 7860

tcataaacag gaaaaaaga gaactcagcc acaagaggaa gaagagagc gttggcttgg 7920

tcattatgct agttatcatg gcaatagtag ctgccgagg ggcttctctg ggagtcgcaa 7980

acgcgattca gcagctttac actaaggcag ctgtccagac ccttgctaat gcaactgctg 8040

cacagcagga tgtgttagag gcaacctatg ccatggtaca gcatgtggct aaaggcgtac 8100

gaatcttga agctcgagtg gctcgagtgg aagctatcac agatagaata atgctatacc 8160

aagaattgga ttgttggcac tatcatcaat actgtataac ctctacaaaa acagaagtag 8220

caaaatatat caattggacg aggtttaagg ataattgcac atggcagcag tgggagagag 8280

gattacaggg gtatgataca aacttaacaa tactgttaaa ggaatcagca gcaatgacac 8340

aactagcaga agagcaagca aggaggatac cagaagtatg ggaaagtta aaagacgtct 8400

ttgattggtc aggatggttc tcatggctaa agtatattcc tattatagta gtaggattat 8460

taggatgcat tctgataaga gctgtgatat gtgtatgtca acctcttggtg cagatatata 8520

gaactctaag taccccgaca taccaacggg tcacagtcac catggaaaca agagcagacg 8580

tgcaggaga aaatcaggat tttggcgatg gcttagagga atcagacaac agcgaacaa 8640

gcgaaagagt gacagtacag aaagcttga gccgtgctg ggagctttgg cagaactcac 8700

cctggaagga gccatgaaa aggggcctgc ttaggctgct cgtccttccg ctgacgatgg 8760

gaatctggat aaatggatgg cttggagaac accacaaaaa taaaaaaga aagggtgact 8820

gtgagacatg ggctaaagag gactaataac aagctaggcc aaattcctgt aaatcacttg 8880

gggggttata agaaaagcaa gttcactatg acaaagcaaa atgtaaagc caaattcctg 8940

taaatcactt ggggggttat aagaaaagca agttcactat gacaaagcaa aatgtaaccg 9000

caagtctga cagatgtaac agctgacata tcagctgatg cttgctcatg ctgacactgt 9060

agctctgagc tgtatataag gagaagcttg ctgcttgac ttcagagttc taggagagtc 9120

cctcctagtc tctcctctcc gaggaggtac cgagacctca aaataaagga gtgattgcct 9180

tactgccga 9189

- <210> 2
- <211> 8919
- <212> DNA
- <213> Caprine arthritis-encephalitis virus

<400> 2
 gtgagtgtc tgaggagctc gaagaaaaga gtcctcagcc tctcctctcc gaggagcttc 60
 ggctcataat aaaggagtgc ttgcttcaac agaactgagc tggtcgtggt tattatcggg 120
 gaccgaagtc ccgtgcaaca ccggggcggg tctcgcagct ggcgccaac gtggggctcg 180
 agtagcttga gaagctcgac tgagatctga atccaagagc gacatcagac agcaagaaat 240
 gagagtaatg agaccgcgag ctctgctgct gtaaaaaaga ggaagtagcg ggttgccgag 300
 gcaactgtc agaagaacca ggggaaagg cttccagcaa cctcaaaaga ggaaccgaga 360
 cttcggggac gcctgaagta aggtaagtga ctctgctgta cgcgggcgga ggcataggag 420
 atccttctat tctaggaaga gaagcgtgt tctgggaggt cttggcgacc gagaatcttg 480
 ttaaataagc caggatctcg atcaggacca agaccctca ggagagggta tagacagcgt 540
 ggtaagaaat ccgccgtggt gagtctagat agagacatgg tgaggcagc ctccggaagg 600
 ggaaaggagt accccgagct aaaagaatgt ctgaaaaagg catgcaaaat aaaagtaagg 660
 gctggggggg agcgcctgac agaaggaaat tgtctctggt gtataaaac actagagtgt 720
 atgtatgagg attgtagga ggaaccttgg accccagaaa aatgtaaca attatgaaa 780
 aagttgaagc aggtagagcc tgaggagagt agcaaagcag actataactc gttaaaagca 840
 accttggcgg ggatagtctg tgtgcaaatg ggaatgcagc ccgagacact gcaggatgcg 900
 atagcaacct taaacatgag agatgaagta aaaggaaagg aaaagccatc agaagaaaag 960
 aagggaatat atcccatatt agtgcaggca ggaggaggaa gagcatggag agcggtagag 1020
 cctgctacct ttcagcagct ccaaacagtg gcaatgcagc atggactagt atcagaagaa 1080

ttgaaaggc agctagcata ctatgccacc acatggacaa gcaaggatat cttagaagta 1140
 ttagccatga tgccaggaata tagagcgcaa aaagaactaa tacaaggaata gttaaatgag 1200
 gaagcagaga gatggagaag gcagaatcca caacctgcgg gcgggttaac cgtggatcag 1260
 ataatggggg taggacaaac gaatcaggca gcggcacagg ctaatatgga tcaagcaaga 1320
 caaatatgcc tacaatgggt tataacagca ataagaggag ttaggcatat ggcccataga 1380
 ccaggaatc ccatgctggt aagacaaaaa ccaaatgaga actatgaaga gtttgccgca 1440
 aggttgtag aagcagtgga tgcagaacc gttaccaac ctataaaaga atatttaaag 1500
 gtaactctgt cttacacaaa tgcaaatcg gaatgtcaaa aacatatgga cagagtgttg 1560
 gggcaaagag tacagcaggc ctcaatagaa gaaaaaatgc aggcattgag ggacatcggg 1620
 ggaacagcat atcagatgca gttgcttgca caagccctcc gtggcggaaa agaagatggg 1680
 aaaaaatctg tagggaagt ttataactgt ggaaggcccg gacacagagc aaaagaatgc 1740
 agacaaggca ttatatgca caactgtgga aaaagagggc atatacagaa aaactgcaaa 1800
 cagaaaagaa gaaaggagca gggaaacatg aggagggggc tacgtgtggt gccgtccgca 1860
 cccctatgg agtaacgcaa gcaccactaa tagttagggt acaaataggg aatcaggaga 1920
 aacaattatt attgacaca ggggcagata aaacgatagt aagaatgcat gatggaacag 1980
 ggattccaaa cggaagaata aaattacaag ggataggagg aatagtagaa ggagaaaaat 2040
 ggaataaagt acccatgaca tataaggag aaacatcctg cccaagcttg gttgtgctaa 2100
 gagatagccc agtagaagta ttgggaagag ataacatgga agcattcggc gtaaccctaa 2160
 taatgcaaaa tttagaagat aagaaaattc ccacaatacc agtagaattg aaagaaggat 2220
 gtaaagggcc acatgtgcc cagtggccat taacagcaga gaaattacaa ggactaacag 2280

gaatagtaga aaaattacta caggaaggaa aattggcaga ggccccagag g gatggacgt 2340
 ggaacacgcc caicttctgc ataaaaaaga agtcaggaaa atggagaatg ttaatagatt 2400
 ttagggaatt aaataagcaa acagcagatt tagcagaagc gcagctagga ctgccacacc 2460
 caggagggtt gcaaaggaaa aagaatgtaa caattctgga cataggagat gcatatttca 2520
 caattccctt atacgagccc tatcagaaat atacatgctt cacactccta agtcctaaca 2580
 atttgggacc atgtaaaagg tattattgga aagtattacc ccagggatgg aaattgagcc 2640
 cagctgtata tcaattcacc atgcaaaggt tgttaaaagg atggatacaa cagcataaaa 2700
 acatacaatt tggaatatat atggatgata tctatatgg aagtatccta acgatagccc 2760
 aacataggaa gataatagaa gaattagcct catttataga acaatttggg ttacattac 2820
 cagaagataa gagacaagag ggctatccag caaaatggct aggattcgag ctacatccag 2880
 aaaaatggaa atatcaaaag cataaattgc cggaattaca agaggggta ataacctga 2940
 acaaattaca gaagatagta ggggaattag tgggagaca atccttgata ggaaagagca 3000
 tccccaatat cataaaatta atggaaggag atcgcgcatt acaaagtga aggaaaatag 3060
 aaagaataca tgtacaagaa tgggaagcat gtcaaaagaa attagatgaa atggtaggaa 3120
 attattacag agaagaagaa gatattatg gacaaataac ttgggggat aaggcaataa 3180
 aatacatagt attccaaagg aaaggggaac ccctatgggt aatgtagta catgacataa 3240
 aaaatttgag tctcccacag caagtataa aagcagcaca gaaattaacc caggaagtaa 3300
 tcataagaac aggaaaaatc ccatggctgc tactaccagg aagagaagaa gactggagat 3360
 tagaactgca ggtagggaac atcacgtgga tgccatcatt ttggtcatgt tatcgaggag 3420

cacccaagtg gaaaagaagg aacatagtgg cagcagtggg agatggaccg acatattata 3480

cagatggggg aaagaaaaac gcacagggaa gctttggctt catctccca acaggagaaa 3540

agttcagaag gcatgaagat ggaactaatc aggtattaga attaaggca atagaagatc 3600

catgtaaca aggacctgaa agcatgaaca ttgtaactga cagcaggtat gcttatgaat 3660

tcatgctccg aaactgggat gaacaggtca taagaaacc cattcaggca agaatcatgg 3720

cagaagtgca caagaaaaag caggtaggaa tacactgggt gccaggcat aaaggaatac 3780

ctcagaatga agagatagac cagtacatat cagaagtatt cttagcacga gaaggaacag 3840

ggatatgtga aaaaaggaag gaagatgctg gatatgattt attatgcccg catgaggtaa 3900

tacttaaacc ccaagaagta aaacggatcc caatagacct aaaattaaaa ttgaaagaaa 3960

agcaatgggc catgataagt gggaaaagta gcgttcgagc aaaaggaata ttigtacaag 4020

gaggcataat agattcaggg tatcaggac aagtacaagt catcctatat aatagtaata 4080

agatagaggt caaaatacca caaggcagga aatttgcca attaataatta atgaacttac 4140

aacatgaaga attagaagaa tgggaaagg aaagaaaaac agaaagagga acaaaaggat 4200

ttgggtctac aggagcattt tggatagaga atattccca agcagaggaa gaacattaca 4260

aatggcatca agatgctaga tctctgcagc tagaattcaa gatactaga gcagcagcag 4320

aagacattat acagcactgt gaggtatgtc aagaaggcaa acccgagcg atcacgagag 4380

ggggaataa aagaggaata gatcattggc aggtagacta tacacattac aaagaacaca 4440

taatattagt atgggtagag actaattcag gattaatatt tgcagagaaa gtaaaaggag 4500

aatcaggaca agaatttagg atgcagacat tgaatggta tgctttgttt caacaaaat 4560

cagtgaate agataatggg acagccttca cagctgagc tacgcagcat ctaatgaagt 4620

atttaggat tcagcacact acgggtattc cgtggaacce ccagtcacaa agtttagtag 4680
 aaagagctca tcaaacatta aaacacatgt tagaaaaatt agaaccacaa tttgtggccc 4740
 tacagtctgc catcgagcc actctagttg cgctcaatat aaaaagaaag ggtggactag 4800
 gggcaagccc tatggatatt tacatatata ataaggagca acaaagacaa caagataata 4860
 gtaataaatt aaitcagaaa aaatittgtt attacaggat cagaaaaaga ggccatccag 4920
 gagagtggaa cggcccaact gaggtactgt gggaagggga aggagccata gtagttaag 4980
 acaaagaaag tgatagatat ctagtcatcc catataaaga tgcaaaatit attccgccac 5040
 cgtcggaaaca gaagggatag aagaataggt ccagaattgc ctttatcttt atggacttat 5100
 acagcataca gcataaataa agatcccga tggatataca ccctaagact gcagcaaatg 5160
 atgtggcata ggaggggaaa taaattgaca tatgtcaggg aaaatgcaca gtacgaggag 5220
 tgggaaatga cctcgtatga gtggaggata agaatgagaa gggacaaaac aaaaagtcat 5280
 ccaagagggc atacttcgcc atggcaatat cggagacagg atggatggaa ggatgtggga 5340
 acgtggttcc tacagccagg ggactataga aaggcggatc agcagttctg gttcgcttgg 5400
 agaatagtgt cgtgttcatg taaaaggaa ggatttaaca taagagaatt tatgctaggt 5460
 acccatagat gggatttgtg taagtcgtgt tgccagggtg aagtagtaa gagaacacaa 5520
 ccctacacct tgcaaaggct cacgtggctt aaattaacag aagacatgt atttcaagta 5580
 atgcccttgt ggagagctcg caaagggatt accatagact ttccttggg caggacaca 5640
 aaaggattcc tggagccgtg gacaacgcaa gagtgttggc aaatagagta tcccttggag 5700
 gatgagtgag gaaaccccag caggaagaga accgactgca gaggaatat ttgagcaaga 5760

agcagaaagt tggaagagaa caagcgtgcg agtcccaaat gacatattac aaagatggct 5820
 agcaatgctt aggcaaagag gaaatagaaa gaaagtgctt agggaaatgc aaaaatgggc 5880
 atggaggaat cccacggcgc gggtgattcg gccgtgtgga tgtcggctat gtaaccccg 5940
 ctgggggagt aattaatcat aataaagcaa attgtaacat gctgtgtcag gtgtcttgca 6000
 ggaatggcgg agataagaaa agaagcaaag gagccactaa tccagggtaa gtataaaaa 6060
 caggtaahta gaataactat agttatatta ctaacagtaa gagcagcact aggagcagaa 6120
 tacataacca taatatcaga cccatatggg ttctctcccg tgagaaatgt gtcaggagta 6180
 cctgtaactt gtgtgacaaa agaatttagt aagtggggat gtcagccaat aggagcctac 6240
 ccagaccag acttagaata cagaaatata agtaaagaaa tattagagga agtatatcaa 6300
 caagactggc cgtggaatac ttatcattgg ccattatggc aaatggataa ttagtagaaa 6360
 tgggcaaggc aaaatttaca ggataaccgc aaggaaaaaa gggacctggc agacctatta 6420
 gcaggaaaaa taaggggaag attctgtgia ccctacccat ttgcgctcct ggagtgcatt 6480
 gaatggtgct ggtgggttaa gaacactaat gcaggggggt atggagaagc agacataaga 6540
 ataaattgct caagggcaag agcagtgagc tgcacaagtg aaatgccctt agcatccta 6600
 cagagggtat attgggaaaa ggaggaacga aaaacatgg agaaaatgac catcaaact 6660
 tgcaataaaa atttggatg caagaacaga aggggatgcg cagaagggtta tccagtacct 6720
 cccaaggcag agttattccc tccagcgttt caggatttac agccaaaagg gtacgcatat 6780
 ggggcactta gagggaacag caaatttcca caaagagtgt cgctaagaac atgggtgaaa 6840
 atagctaacc tgacaggatg ggaaaaagga aagccagcag aatggtggaa taccagccaa 6900
 caggttcatt ggtttgatac cacgccacaa tatcatttag gatatgtatt atcccagcgc 6960

cctgagaaca ggagttgtaa tttcacaggg gaaatacгаа tagggcaaca tcagtttgag 7020

tataattaca ccctgacaaa gaattgcaca aaggagaagt ggaaagagta ccccatgtgg 7080

catgtctgga ggcatttaga tcaaatgag cacttatcta gcatatgttt caaaagaccg 7140

agaagaaatg caacacaaat agggaacagt aactgcaag ggcaatgtaa tagaagtaat 7200

tggacaggat gccactgcaa tgagacaggg ataacacaa catggagaat aatggcaca 7260

aaggagcctt atctcttaaa tagcactaat gaaacatca tggcttgtt atgctggaac 7320

acaacagtgg caggggtata tgagagtcag ctaaagtgga atgagagtct taaagacgga 7380

gactatgggc tctgttttaa ttcaacaaac aggaattgta ctagaatgg agctcggcac 7440

tatgtaaaca agagagtgat aaaaaacgac acagcagatc ataattgtga tagcagcata 7500

tcagcaatag atggaatggt acatcaacaa atattactgc aaaggatca agtaattaga 7560

gtaagagcct acacatacgg agtgattgat atgccagaca attatgagac cctaccagga 7620

aggagaagga gagatctcgc aaaggccagg aaaaagaggg gcgtggcct ggtcatcatg 7680

ttagctatca tggccatagt ggctgctgca ggagcatctc tgggagtcgc gaacgcgatt 7740

cagcagtcct acaccagga cgctgtccag actcttgcta acgcgactgc tgtgcaacag 7800

caggtgttag aggcgtccta tgccatgata cagcatgtgg ctaaggaat acgcatcctt 7860

gaagcacgcg tggcgagaat ggaagttag atggatagaa tgatgttata tcaggaagta 7920

gactgctggc attatcacca atattgtgta acctctacaa gagcagacat agtgaattac 7980

attaattgga caaggtttaa agataattgc acatggcaag agtgggaaag ggagataagt 8040

gcgcatgaag gaaacatcac tatattactc aaagaatcag caaggataac acaattagca 8100

caacaaaagg tacaaagaat accagatgtg tggacagcac taaggagatc actaggatgg 8160

acacaatggc tgcttggat aaaatacctt cccataatag tagtagggat attaggatgc 8220

ataatcataa gaataatggt gtgtgtagta caaccagttc ttcagattta cagaaccttg 8280

actcagacca ggiatcaaca agtcaacttg gtgatggaga cccgggtgca actagaagaa 8340

gaagaagaag aagacggaag ggatggtgga gatggctcag agagatgcag cgatcccgc 8400

aacaaaggaa ttatgaacgc ctggaggaga gcttgggtga cttggagaaa ctcaccttgg 8460

cagaacacat ggaagaatgt ggtggtggcg ccgttgggtga ttccgctgac aatcagaatt 8520

tggctccttg gagagaatgg agagaacccc taaaagaaaa ataaaaaggg tggactgtga 8580

ggactgtgag gcctaggagc gagatagaaa cttataggcc tctcttcccg gaaagctaac 8640

tcactgtgag aggaatagca agtcacagtg aactgctaa ttgtaccgc aacctgaga 8700

tcattgcaa cacaatcctg agattatgct gacatgtgta acagctgatg cctcagctga 8760

tgcttgctca tgctgacaat gtaactagga gctctatata aacagagccc tagagcttgc 8820

tacttcagag tgctctgagg agctcgaagg aaagagtcct cagcctctcc tctccgagga 8880

gcttcggctc ataataaagg agtgcttgct tcaacagaa 8919

<210> 3
 <211> 720
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 3
 atggtgagtc tagatagaga catggcgagg caagtctccg ggggaaaag agattatcct 60

gagctcgaaa aatgtatcaa gcatcatgc aagataaaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatgtt tgaggacat 180

aaagaggaac cttggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240
 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300
 atgtgttgcc aaatggggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360
 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420
 agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480
 gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540
 gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600
 gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660
 gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

<210> 4
 <211> 720
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 4
 atggtgaggc aggcctccgg aaggggaaag gagtaccccg agctaaaaga atgtctgaaa 60
 aaggcatgca aaataaaagt aagggtctggg ggggagcgcc tgacagaagg aaattgtctc 120
 tggtgtataa aaacactaga gtgtatgtat gaggattgta gggaggaacc ttggaccca 180
 gaaaaatgta aacaattatg gaaaaagttg aagcaggtag agcctgagga gagtagcaaa 240
 gcagactata actcgttaaa agcaaccttg gcgggatag tctgtgtgca aatgggaatg 300
 cagcccgaga cactgcagga tgcgatagca accttaaca tgagagatga agtaaaagga 360
 aaggaagagc catcagaaga aaagaaggga atatatcca tattagtgca ggcaggagga 420

ggaagagcat ggagagcggg agagcctgct acctttcagc agctccaaac agtggcaatg 480

cagcatggac tagtatcaga agaatttgaa aggcagctag catactatgc caccacatgg 540

acaagcaagg atatcttaga agtattagcc atgatgccag gaaatagagc gcaaaaagaa 600

ctaatacaag gaaagttaaa tgaggaagca gagagatgga gaaggcagaa tccacaacct 660

gcgggcgggt taaccgtgga tcagataatg ggggtaggac aaacgaatca ggcagcggca 720

<210> 5

<211> 1347

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 5

atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60

gagctcgaag aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180

aaagaggaac ctiggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

atgtgttgcc aaatggggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360

atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540

gaaagcagc tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660

gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

caaattatgg gggtaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 780
 aggcaaatat gcctgcaatg ggtaataaat gcattaagag cagtaagaca tatggcgcac 840
 aggccagga atccaatgct agtaaagcaa aaaacgaatg agccatatga agattttgca 900
 gcaagactgc tagaagcaat agatgcagag ccagttacac agcctataaa agattatcta 960
 aagctaacac tatcttatac aatgcatca gcagattgtc agaagcaaat ggatagaaca 1020
 ctaggacaaa gagtacaaca agctagtgtga gaagaaaaaa tgcaagcatg tagagatgtg 1080
 ggatcagaag ggttcaaaat gcaattgtta gcacaagcat taaggccagg aaaaggaaaa 1140
 gggaatggac agccacaaag gtgttacaac tgtggaaaac cgggacatca agcaaggcaa 1200
 tgtagacaag gaatcatatg tcacaactgt ggaaagagag gacatatgca aaaagaatgc 1260
 agaggaaaga gagacataag gggaaaacag cagggaacg ggaggagggg gatacgtgtg 1320
 gtgccgtccg ctctcctat ggaataa 1347

<210> 6
 <211> 1299
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 6
 atggtgagge aggcctccgg aaggggaaag gagtaccccg agctaaaaga atgtctgaaa 60
 aaggcatgca aaataaaagt aagggtggg ggggagcgcc tgacagaagg aaattgtctc 120
 tgggtgataa aaacactaga gtgtatgtat gaggattgta gggaggaacc ttggacccca 180
 gaaaaatgta aacaattatg gaaaaagttg aagcaggtag agcctgagga gagtagcaaa 240
 gcagactata actcgttaaa agcaaccttg gcgggatag tctgtgtgca aatgggaatg 300

cagcccgaga cactgcagga tgcgatagca accttaaca tgagagatga agtaaaagga 360

aaggaaaagc catcagaaga aaagaaggga atatatcca tattagtgc ggcaggagga 420

ggaagagcat ggagagcggg agagcctgct acctttcagc agtccaaac agtggcaatg 480

cagcatggac tagtatcaga agaattttaa aggcagctag catactatgc caccacatgg 540

acaagcaagg atatcttaga agtattagcc atgatgccag gaaatagagc gcaaaaagaa 600

ctaatacaag gaaagttaa tgaggaagca gagagatgga gaaggcagaa tccacaacct 660

gctggcgggt taacctgga tcagataatg gggtaggac aaacgaatca ggcagcgga 720

caggctaata tggatcaagc aagacaaata tgcctacaat gggttataac agcaataaga 780

ggagttaggc atatggccca tagaccagga aatcccatgc tggtaagaca aaaaccaaat 840

gagaactatg aagagtttgc cgcaaggttg ttagaagcag tggatgcaga acccgttacc 900

caacctataa aagaatattt aaagtaact ctgtcttaca caaatgcaa ttcggaatgt 960

caaaaacata tggacagagt gttggggcaa agagtacagc aggcctcaat agaagaaaaa 1020

atgcaggcat gcagggacat cgggggaaca gcatatcaga tgcagttgct tgcacaagcc 1080

ctccgtggcg gaaaagaaga tgggaaaaa tctgtaggga agtgttataa ctgtggaagg 1140

cccggacaca gagcaaaaga atgcagacaa ggcatatat gtcacaactg tggaaaaga 1200

gggcatatac agaaaaactg caaacagaaa agaagaaagg agcagggaaa catgaggagg 1260

gggctactg tggtgccgtc cgcaccccct atggagtaa 1299

- <210> 7
- <211> 511
- <212> DNA
- <213> Caprine arthritis-encephalitis virus

<400> 7

gagttctagg agagtcctc ctagtctctc ctctccgagg aggtaccgag acctcaaat 60

aaaggagtga ttgccttact gccgagtga gagtgattac tgagcggccg gtgtatcggg 120

agtcgtccct taatctgtgc aataccagag cggctctcgc agctggcgcc caacgtgggg 180

cccgaggaga agaaaagaaa gcggcctga gaactcggct tctgaaaaag aggaagagga 240

caagttgcta tagcaacaag agagaagaag tagagcaaag gtccagtggc tcggaaaaag 300

aggaactgaa acttcgggga cgcctgaagg agtaaggtaa gtgactctgc tgtacgctgg 360

gcgaggcaga ggtttccttc taaattgaaa gagaagtgtt gctgcgagag gtcttggtgg 420

tcgagaatcc tgiacaaaaa aaaggagga tctcggtcag gaccaggacc cctgggagta 480

atacaacagc aacaccgtaa gaaaatccgc c 511

<210> 8
 <211> 576
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 8
 gtgagtgtc tgaggagctc gaaggaaaga gtctcagcc tctcctctcc gaggagcttc 60

ggctcataat aaaggagtgc ttgcttcaac agaactgagc tggtcgtggt tattatcggg 120

gaccgaagtc ccgtgcaaca ccggggcggt tctcgcagct ggcgccaac gtggggctcg 180

agtagcttga gaagctcgac tgagatctga atccaagagc gacatcagac agcaagaaat 240

gagagtaatg agaccgcgag ctctgctgct gtaaaaaaga ggaagtagcg gtttgccgag 300

gcaactgctc agaagaacca ggggaaaggg ctccagcaa cctcaaaaga ggaaccgaga 360

cttcggggac gcctgaagta aggtaagtga ctctgctgta cgcggggcga ggcataggag 420

atccttctat tctaggaaga gaagcgtgt tctgggaggt cttggcgacc gagaatcttg 480

ttaaataagc caggatctcg atcaggacca agaccctca ggagaggta tagacagcgt 540

ggtaagaaat ccgccgtggt gagtctagat agagac 576

<210> 9
 <211> 3318
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 9

atgtcacaac tgtggaaaga gaggacatat gcaaaaagaa tgcagaggaa agagagacat 60

aaggggaaaa cagcaggaa acgggaggag ggggatacgt gtggtgccgt ccgtcctcc 120

tatggaataa cttcagcacc acctatggtt caggtccgca taggttccca gcagaggaac 180

ttgttatttg ataccggggc ggaccgaact atagttagat ggcatgaggg ctcgggaaac 240

ccagccggaa ggataaaact gcaaggaata ggaggaatag tagaaggaga aaaatggaat 300

aatgtagaat tagaatataa aggagaaaca agaaaaggaa caatagtagt gttaccacaa 360

agtccagtag aagtattagg acgagataac atggcccgat ttggaataaa gataataatg 420

gcaaatttag aggaaaaaag aatccaatt acaaaagtaa aattgaaaga gggatgtacg 480

ggtccacatg tcccacaatg gccattaaca gaagagaaat taaaaggtct aacagaaatc 540

atagataaat tagtggaaaga aggaaaacta ggaaaggcac cccacattg gacatgtaat 600

actccaatct ttgcataaa aaagaaatca gggaagtgga gaatgttaat agatttcaga 660

gaattgaaca aacagacaga agatttaaca gaagcgcagt taggactccc gcatccggga 720

ggactacaaa agaaaaaaca tgttacaata ttggacatag gagatgcata ttttactata 780

cccctatatg aaccatatcg agagtacaca tgttttactc tattaagtcc taataatcta 840

ggacatgta aaagatacta ttggaagtg ctgccacaag gttggaatt gaggccatct 900
 gtatatcaat ttactatgca ggagatctta gaggattgga tacagcagca tccagaaatt 960
 caatttgca tatatatgga tgatatttac ataggaagtg atttagaat taaaaagcat 1020
 agagaaatag tgaagattt agccaattat attgccaat atggattcac tctgccagaa 1080
 gagaagagac aaaaggata tccagcaaaa tggctaggat ttgaactaca cccgagacc 1140
 tggaaattc agaagcatac attacctgaa ttaacaagg gaacaataac attaaataa 1200
 ttacagaaat tagtaggaga attagtatgg agacaatcca taattggga aagcattcct 1260
 aacattctga aattaatgga aggagataga gaattacaaa gtgaagaaa aattgaagaa 1320
 gtacatgta aagaatggga agcatgtagg aaaaaatag aagaatgga aggaaattat 1380
 tataataaag acaagatgt ctatggacaa ttggcttggg gagacaaagc tatagaatat 1440
 atagtgtatc aggagaaagg gaaaccatta tgggtaaatg tggttcaca tataaagaac 1500
 ctaagcatcc cgcaacaggt tattaagca gcgcaaaaat taaccaaga agtcatcatt 1560
 aggacaggaa aaatacatg gatattgttg ccagggaag aagaagattg gagactagaa 1620
 ttgcaattag ggaacatcac atggatgcca aaattttggt cctgttatcg aggacataca 1680
 agatggagaa aaagaaat ataatagaaga gtagtagaag ggcctacata ttatacagat 1740
 ggaggaaaaa agaataaagt aggaagtcta gggttcatag tatcaacagg gaaaaattt 1800
 agaaagcatg aagaggcac aaaccagcaa ctagaattaa gagccataga ggaagctcta 1860
 aaacaaggc ctcaacaat gaatttagta acagatagta gatatgcatt tgaattttta 1920
 ttaagaaatt gggatgaaga agtaataaag aatccaattc aagcaagaat tatggaatt 1980
 gccacaaga aagataggat aggagtgc atgggtgccag gacataaagg gattccca 2040

aatgaagaaa tagacaaata tatttcggaa atatttcttg caaaagaagg agaaggaatt 2100
ctcccaaaaa gagaagagga tgcagggtat gatttaatat gcccagaaga ggttaccata 2160
gagccaggac aagtgaaatg catcccata gagctaagat taaatttaa gaaatcacia 2220
tgggctatga ttgctacaaa aagcagcatg gctgccaaag gagtgttcac acaaggagga 2280
atcatagact caggatatca gggacaaata caggtaataa tgtataatag caataaata 2340
gcagtagtca taccccaagg gagaaaattt gcacaattaa tattaatgga taaaaagcat 2400
ggaaaattgg aacctgggg ggaaagcaga aaaacagaaa ggggagaaaa aggatttggg 2460
tctacaggaa tgtattggat agaaaatatt cctctggcag aggaagacca cacaaaatgg 2520
catcaagatg cccgatcatt gcatctagaa ttgaaattc caagaacagc agcagaagac 2580
atagtaaate aatgtgaaat atgcaaagaa gcgaggacac ctgcagtaat tagaggcggg 2640
aacaaaaggg gggtaaatca ttggcaagtg gattataccc attatgaaaa tatcatacta 2700
ttagtatggg tagaaacaaa ttcaggacta atatatgcag aaaaagtaa aggagaaatca 2760
gggcaagaat tcagaataaa agtgatgcat tggatgcat tatttggcc agagtattg 2820
cagtcagaca atggacctgc atttcagca gagccacac agctgttaat gcaataccta 2880
ggagtaaaac acacaacagg cataccttgg aatccacagt ctcaggctat agtagaaagg 2940
gcacatcaac tattgaaaag cactttaaag aagttccagc cacaatttgt cgctgtagaa 3000
tcagccatag cagcagccct agtcgccata aatataaaaa gaaagggtgg gctggggaca 3060
agccctatgg atattttat atataataaa gaacagaaaa gaataataa taaatataat 3120
aaaaattctc aaaaaattca attctgttat tacagaataa ggaaaagagg acatcaggag 3180

agtggaaagg accaaccag gtactgtgga aaggggaagg agccaattgt ggtaaaggat 3240
 atagaaagtg aaaagtattt agtaatacct tacaaagatg caaaattcat cccgccacca 3300
 acaaaagaaa aggaataa 3318

 <210> 10
 <211> 3324
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

 <400> 10
 atgcagacaa ggcattatat gtcacaactg tggaaaaaga gggcatatac agaaaaactg 60
 caaacagaaa agaagaaagg agcagggaaa catgaggagg gggctacgtg tgggtccctc 120
 cgcaccccct atggagtaac gcaagcacca ctaatagtta gggtaacaaat agggaatcag 180
 gagaaacaat tattatttga cacaggggca gataaaacga tagtaagaat gcatgatgga 240
 acagggattc caaacggaag aataaaatta caagggatag gaggaatagt agaaggagaa 300
 aatggaata aagtacccat gacatataag ggagaaacat cctgccaag cttggttgtg 360
 ctaagagata gcccagtaga agtattggga agagataaca tggaagcatt cggcgtaac 420
 ctaataatgg caaatttaga agataagaaa attcccacaa taccagtaga attgaaagaa 480
 ggatgtaaag ggccacatgt gcccagtg cattaacag cagagaaatt acaaggacta 540
 acaggaatag tagaaaaatt actacaggaa ggaaaattgg cagaggcccc agagggatgg 600
 acgtggaaca cgcccatctt ctgcataaaa aagaagtcag gaaaatggag aatgtaata 660
 gatttaggg aattaaata gcaaacagca gatttagcag aagcgcagct aggactgcca 720
 caccaggag ggttgcaaag gaaaaagaat gtaacaattc tggacatagg agatgcatat 780
 ttcacaatte ccttatacga gccctatcag aaatatacat gcttcacact cetaagtct 840

aacaatttgg gaccatgtaa aaggtattat tggaaagtat taccccaggg atggaaattg 900

agcccagctg tatatcaatt caccatgcaa aggttgtaa aaggatggat acaacagcat 960

aaaaacatac aatttggat atatatggat gatctata ttggaagtga tctaacgata 1020

gcccaacata ggaagataat agaagaatta gcctcatta tagaacaatt tgggtttaca 1080

ttaccagaag ataagagaca agagggctat ccagcaaat ggctaggatt cgagctacat 1140

ccagaaaaat gaaataatca aaagcataaa ttgccggaat tacaagaggg gtaataacc 1200

ctgaacaaat tacagaagat agtaggggaa ttagtggtga gacaatcctt gatagaaag 1260

agcatcccca atatcataaa attaatggaa ggagatcgcg cattacaaag tgaaaggaaa 1320

atagaaagaa tacatgtaca agaatgggaa gcatgtcaaa agaaattaga tgaaatgta 1380

ggaaattatt acagagaaga agaagataatc tatggacaaa taacttgggg ggataaggca 1440

ataaaataca tagtattcca aaggaaaggg gaaccctat gggtaaatgt agtacctgac 1500

ataaaaaatt tgagtctccc acagcaagtg ataaaagcag cacagaaatt aaccaggaa 1560

gtaatcataa gaacaggaaa aatcccatgg ctgctactac caggaagaga agaagactgg 1620

agattagaac tgcaggtagg gaacatcacg tggatgcat cattttggtc atgttatcga 1680

ggagcaccca agtggaaaag aaggaacata gtggcagcag tggtagatgg accgacatat 1740

tatacagatg ggggaaagaa aaacgcacag ggaagctttg gcttcatctc cccaacagga 1800

gaaaagtcca gaaggcatga agatggaact aatcaggtat tagaattaag ggcaatagaa 1860

gatccatgta aacaaggacc tgaaagcatg aacattgtaa ctgacagcag gtatgcttat 1920

gaattcatgc tccgaaactg ggatgaacag gtcataagaa acccattca ggcaagaatc 1980

atggcagaag tgcacaagaa aaagcaggta ggaatacact gggtgccagg gcataaagga 2040

ataacctcaga atgaagagat agaccagtac atatcagaag tattcttagc acgagaagga 2100

acagggatat gtgaaaaaag gaaggaagat gctggatatg atttattatg cccgcatgag 2160

gtaatactta aacccaaga agtaaacgg atccaatag acctaaaatt aaaattgaaa 2220

gaaaagcaat gggccatgat aagtgggaaa agtagcgttg cagcaaaagg aatatttgta 2280

caaggaggca taatagattc agggatcag ggacaagtac aagtcacct atataatagt 2340

aataagatag aggtcaaat accacaaggc aggaaattg cccaattaat attaatgaac 2400

ttacaacatg aagaattaga agaatgggga aaggaaagaa aaacagaaag aggaacaaaa 2460

ggatttgggt ctacaggagc attttgata gagaatattc cccaagcaga ggaagaacat 2520

tacaaatggc atcaagatgc tagatctctg cagctagaat tcaagatacc tagagcagca 2580

gcagaagaca ttatacagca ctgtgaggta tgtcaagaag gcaaaccgc agcgatcacg 2640

agaggggaa ataaaagg aatagatcat tggcaggtag actatacaca ttacaaagaa 2700

cacataatat tagtatgggt agagactaat tcaggattaa tatttgcaga gaaagtaaaa 2760

ggagaatcag gacaagaatt taggatgcag acattgaaat ggtatgcttt gtttcaacca 2820

aatcagtgc aatcagataa tgggacagcc ttcacagctg aggctacgca gcatctaatg 2880

aagtatttag ggattcagca cactacgggt attccgtgga accccagtc acaaagtta 2940

gtagaaagag ctcatcaaac attaaaacac atgttagaaa aattagaacc acaatttgtg 3000

gcctacagt ctgcatcgc agccactcta gttgcgctca atataaaaag aaagggtgga 3060

ctaggggcaa gcctatgga tatttacata tataataagg agcaacaag acaacaagat 3120

aatagtaata aattaattca gaaaaattt tgttattaca ggatcagaaa aagaggccat 3180

ccaggagagt ggaacggccc aactgaggta ctgtgggaag gggaaggagc catagtagtt 3240
 aaagacaaag aaagtgatag atatctagtc atcccatata aagatgcaaa atttattccg 3300
 ccaccgtcgg aacagaaggg atag 3324

<210> 11
 <211> 402
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 11
 atggatgctg gggccagata catgcgctta actgggaagg aaaactgggt tgaagtaacc 60
 atggacggag agaaggaaag gaaaagagaa ggtttcactg cgggacagca agatatacag 120
 aactetaagt accccgacat accaacgggt cacagtcate atggaacaa gagcagacgt 180
 cgcaggagaa aatcaggatt ttggcgatgg cttagaggaa tcagacaaca gcgaaacaag 240
 cgaaagagt acagtacaga aagcttggag ccgtgcctgg gagctttggc agaactcacc 300
 ctggaaggag ccatgaaaa ggggcctgct gaggctgctc gtccttccgc tgacgatggg 360
 aatctggata aatggatggc ttggagaaca ccacaaaaat aa 402

<210> 12
 <211> 321
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 12
 atggcggaga taagaaaaga agcaaaggag ccactaatcc aggaccaggt atcaacaagt 60
 caacttgggt atggagacc gggtgcaact agaagaagaa gaagaagaag acggaaggga 120
 tggatggat ggctcagaga gatgcagcga tcccgacaac aaaggaatta tgaacgctg 180

gaggagagct tgggtgactt ggagaaactc accttggcag aacacatgga agaattggt 240

ggtggcgccg ttggtgattc cgctgacaat cagaatttgg ctcttggag agaattggaga 300

gaaccctaa aagaaaaata a 321

<210> 13
 <211> 690
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 13
 atgcaaaatt catccccca ccaacaaaag aaaaggaata aaaaacctgg accagaatta 60

cccttagcac tatggataca tatagcagaa agcattaatg gggatagctc atggtacata 120

acaatgagac tgcaacagat gatgtgggga aaaagaggaa ataagttaca atataagaat 180

gaagacaggg aatatgaaaa ttgggaaatt acatcatggg gatggaaaat gcacctaagg 240

agagtgaac aatggataca agacaacagg agaggaagcc catggcagta caaagtagga 300

ggaacatgga aaagtatagg agtgtgggtc ctgcaagcag gagattacag aaaggtagac 360

aggcacttct ggtgggcatg gaggatactg atatgttctt gcaggaaaga aaagtttgat 420

ataagagaat ttatgagagg aagacataga tgggatttgt gcaaatcctg tgctcaagga 480

gaagtagtaa agcatactag aacaaaaagt ctggaaagac tagtactgct acagatgta 540

gaacagcatg tgtttcaagt attgccattg tggagagcca ggagaagtag tacaacagat 600

ttcccatggt gcagggacac aacgggatac acgcatgctt ggtctgtcca ggagtctgg 660

ttgatggaat atctcttaga ggatgagtga 690

<210> 14
 <211> 490
 <212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 14

atgcaaaatt tattccgcca ccgtcggaac agaaggata gaagaatagg caaaaagtca 60

tccaagaggg catacttcgc catggcaata tcggagacag gatggatgga aggatgtggg 120

aacgtggttc ctacagccag gggactatag aaaggcggat cagcagttct ggttcgcttg 180

gagaatagtg tcgtgttcat gtaaaaagga aggatttaac ataagagaat ttatgctagg 240

tacccataga tgggatttgt gtaagtcgtg ttgccagggt gaagtagtaa agagaacaca 300

accctacacc ttgcaaaggc tcacgtggct taaattaaca gaagaccatg tatttcaagt 360

aatgcccttg tggagagctc gcaaagggat taccatagac tttccctggt gcagggacac 420

aaaaggattc ctggagccgt ggacaacgca agagtgttgg caaatagagt atcccttga 480

ggatgagtga 490

<210> 15

<211> 720

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 15

atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60

gagctcgaag aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180

aaagaggaac cttggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

atgtgttgcc aaatggggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360

atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgic tgaggacttt 540

gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaatt aatgaagaa 660

gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

<210> 16

<211> 591

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 16

gctgtagact ctgtaatgtt ccaacaaatg caaacagtag caatgcagca tggcctcgtg 60

tccgaggatt ttgaaagaca gttagcatat tatgctacta cctggacaag taaagacata 120

ctagaagtat tggccatgat gcctgggaat agggtcaga aagaacttat tcaagggaaa 180

ttgaatgaag aagcagacag gtggagaagg aacaatccac caggaggatt aacagtggat 240

caaattatgg gggtaggaca aacaatcaa gcagcagcac aagctaacat ggatcaggca 300

agacaaatat gcctacaatg ggtaataaac gccttaagag cagtaaggca tatggctcat 360

aggccaggga atccaatgct agtaaagcaa aaaacaaatg agccatatga agaatttgca 420

gcaagactgc tagaagcaat agatgcagaa gcggttacac agccataaa agagtatcta 480

aagctaacat tatcctatac aaatgcagcc tcagattgic aaaagcaaat ggagagagtg 540

ctaggacaaa gagtacaaca ggctagtgta gaaaaaaaa tgcaagcatg t 591

<210> 17

<211> 1347
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 17
 atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60
 gagctcgaaa aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120
 ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180
 aaagaggaac ctggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240
 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300
 atgtgttgcc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360
 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420
 agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480
 gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540
 gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600
 gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660
 gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720
 caaattatgg ggttaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 780
 aggcaaatat gcctgcaatg ggtaataaat gcattaagag cagtaagaca tatggcgcac 840
 aggccagga atccaatgct agtaaagcaa aaaacgaatg agccatatga agattttgca 900
 gcaagactgc tagaagcaat agatgcagag ccagttacac agcctataaa agattatcta 960
 aagctaacac tatcttatac aatgcatca gcagattgtc agaagcaaat ggatagaaca 1020

ctaggacaaa gagtacaaca agctagtgtga gaagaaaaaa tgcaagcatg tagagatgtg 1080

ggatcagaag ggttcaaaat gcaattgtta gcacaagcat taaggccagg aaaaggaaaa 1140

gggaatggac agccacaaaag gtgttacaac tgtggaaaac cgggacatca agcaaggcaa 1200

tgtagacaag gaatcatatg tcacaactgt ggaaagagag gacatatgca aaaagaatgc 1260

agaggaaaga gagacataag gggaaaacag cagggaacg ggaggagggg gatacgtgtg 1320

gtgccgtccg ctctcctat ggaataa 1347

<210> 18

<211> 591

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 18

gctgtagact ctgtaatgtt ccaacaaatg caaacagtag caatgcagca tggcctcgtg 60

tccgaggatt ttgaaagaca gttagcatat tatgctacta cctggacaag taaagacata 120

ctagaagtat tggccatgat gcctgggaat agggctcaga aagaacttat tcaagggaaa 180

ttgaatgaag aagcagacag gtggagaagg aacaatccac caggaggatt aacagtggat 240

caaattatgg gggtaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 300

agacaaatat gcctacaatg ggtaataaac gccttaagag cagtaaggca tatggctcat 360

agcccaggga atccaatgct agtaaagcaa aaaacaaatg agccatatga agaatttgca 420

gcaagactgc tagaagcaat agatgcagaa gcggttacac agcccataaa agagtatcta 480

aagctaacat tatcctatac aaatgcagcc tcagattgtc aaaagcaaat ggagagagtg 540

ctaggacaaa gagtacaaca ggctagtgtga gaaaaaaaa tgcaagcatg t 591

<210> 19

<211> 720
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 19
 atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60

 gagctcgaag aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

 ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180

 aaagaggaac ctggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

 atgtgttgc aaatgggat gagacctgag acattgcaag atgcaatgac tacagtaatc 360

 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

 agtgtcttc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

 gtaatgttc agcaactgca aacagtagca atgcagcatg gcctcgtgic tgaggacttt 540

 gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

 gccatgatgc ctggaaatag agctcaaaag gagttaaattc aagggaatt aatgaagaa 660

 gcagaaaggt ggagaaggaa taatcacca cctccagcag gaggaggatt aacagtggat 720

<210> 20
 <211> 662
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 20
 tcaagcagca ggaggagaa gctggaaagc agtagactca gtgatgttc agcaactgca 60

 aatgtagca atgcagcatg gcctcgtgic cgaggatatt gaaaggcagt tagtatatta 120

 tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaatag 180

agctcaaaaa gagttaattc aagggaatt gaatgaggaa gcagaaaggt ggagaagaaa 240

taatccacca cctcaagcag gcggaggatt aacagtggat caaattatgg gggtaggaca 300

aacaaatcaa gcagcggcac aggctaacat ggatcaggca agacaaatat gtctgcaatg 360

ggtaataaca gactaagag cagtgagaca tatggctcac aaaccagga atccaatgct 420

agtaaagcaa aaaacaaatg agtcatatga agatTTTgcc gcaagactgc tagaagcaat 480

agatgcagaa ccagttacac agcaataaa agaatatTTa aagTTaacat tatcttacac 540

aatgcatcc tcagactgtc agaaacagat ggatagagta ctaggacaga gagtgaaca 600

agctagtgtg gaagaaaaaa tgcaagcatg cagagatgtg ggatcagaag gattcagaat 660

gc 662

<210> 21
 <211> 662
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 21

gcaagcagca ggagggagaa gctggaaagc agtagactca gtgatgttcc agcaactgca 60

aatgtagca atgcagcatg gcctcgtgtc cgaggatTTT gaaaggcagt tagcatatta 120

tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaatag 180

agctcaaaaa gagttaattc aagggaatt gaatgaggaa gcagaaaggt ggagaaggaa 240

taatccacca cctcaagcag gcggaggatt aacagtggat caaattatgg gggtaggaca 300

aacaaatcaa gcagcggcac aggctaacat ggatcaggca agacaaatat gcctgcaatg 360

ggtaataaca gactaagag cagtgagaca tatggctcac aaaccagga atccaatgct 420

agtaaagcaa aagacaaatg agtcatatga agatTTTgcc gcaagactgc tagaagcaat 480

agatgcagaa ccagttacac agcaataaa agaatatTTa aagTTaacaT tatctTacac 540

aaatgcattcc tcagactgtc aaaaacaaat ggatagaata ctaggacaga gaggTcaaca 600

agctagTgtg gaagaaaaaa Tgcaagcatg cagagatgtg ggatcagaag ggtTcagaat 660

gc 662

<210> 22

<211> 651

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 22

gcaagcagca ggagggagaa gctggaaagc agtagactca gtgatgtTcc agcaactgca 60

aaatgtagca atgcagcatg gcctcgtgtc cgaggatTTT gaaaggcagt Tggcatatta 120

Tgctactacc Tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaacag 180

agctcaaaaa gaggTaatc aggggaaatt gaataaggaa gcagaaaggT ggagaaggaa 240

Taatccacca cctcaagcac aaggaggatt aacagTggat caaattatgg gggtaggaca 300

aacaaatcag gcagcggcac aggctaacaT ggatcaggca agacaaatT gcctgcaatg 360

ggtaataaca gcaactaagag caggTgagaca TaTggctcac aaaccaggga atccaatgct 420

agtaaagcaa aagacaaatg agtcatatga agatTTTgcc gcaagactgc tagaagcaat 480

agatgcagaa ccagttacac agcaataaa agaatatTTa aagTTaacaT tatctTacac 540

aaatgcattcc tcagactgtc agaaacaaat ggatagagta ctaggacaga gaggTcaaca 600

agctagTgtg gaagaaaaaa Tgcaagcatg cagagatgtg ggatcagaag g 651

<210> 23

<211> 520
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 23
 gcaagcagca ggagggagaa gctggaaagc agtagactca gtgatgttcc agcaactgca 60
 aaatgtagca atgcagcatg gcctcgtgtc cgaggatfff gaaaggcagt tagcatatta 120
 tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaatag 180
 agctcaaaaa gagttaattc aagggaatt gaatgaggaa gcagaaaggt ggagaaggaa 240
 taatccacca cctcaagcag gcggaggatt aacagtggat caaattatgg ggtaggaca 300
 aacaaatcaa gcagcggcac aggctaacat ggatcaggca agacaaatat gcctgcaatg 360
 ggtaataaca gcaactaagag cagttagaca tatggctcac aaaccaggga atccaatgct 420
 agtaaagcaa aagacaaatg agtcatatga agatfffgcc gcaagactgc tagaagcaat 480
 agatgcagaa ccagttacac agcaaataaa gaatatttaa 520

<210> 24
 <211> 663
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 24
 gcaagcagca ggagggagaa gctggaaagc agtagactca gtgatgttcc agcaactgca 60
 aaatgtagca atgcagcatg gcctcgtgtc cgaggatfff gaaaggcagt tagcatatta 120
 tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaatag 180
 agctcaaaaa gagttaattc aagggaatt gaatgaggaa gcagaaaggt ggagaaggaa 240
 taatccacca cctcaagcag gcggaggatt aacagtggat caaattatgg ggtaggaca 300
 aacaaatcaa gcagcggcac aggctaacat ggatcaggca agacaaatat gcctgcaatg 360

ggtaataaca gcgactaaga gcagtgagac atatggetca caaaccaggg aatccaatgc 420
 tagtaaagca aaagacaaat gagtcatatg aaaaattttc agcaagactc ctagaagcaa 480
 tagatgcaga accagttaca cagcctataa aagaatattt aaagttaaca ttatcttaca 540
 caaatgcatc ctcagactgt caaaaacaaa tggatagagt actaggacag agagtgaac 600
 aagctagtgt ggaagaaaa atgcaagcat gcagagatgt gggatcagaa ggattcagaa 660
 tgc 663

 <210> 25
 <211> 1347
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

 <400> 25
 atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60
 gagctcgaag aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120
 ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180
 aaagaggaac ctiggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240
 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300
 atgtgttgcc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360
 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420
 agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480
 gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540
 gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660

gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

caaattatgg gggtaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 780

aggcaaatat gcctgcaatg ggtaataaat gcattaagag cagtaagaca tatggcgcac 840

aggccagga atccaatgct agtaaagca aaaacgaatg agccatatga agattttgca 900

gcaagactgc tagaagcaat agatgcagag ccagttacac agcctataaa agattatcta 960

aagctaacac tatcttatac aatgcatca gcagattgtc agaagcaaat ggatagaaca 1020

ctaggacaaa gagtacaaca agctagtgtg gaagaaaaaa tgcaagcatg tagagatgtg 1080

ggatcagaag ggttcaaaat gcaattgtta gcacaagcat taaggccagg aaaaggaaaa 1140

gggaatggac agccacaaag gtgttacaac tgtggaaaac cgggacatca agcaaggcaa 1200

tgtagacaag gaatcatatg tcacaactgt ggaaagagag gacatatgca aaaagaatgc 1260

agaggaaaga gagacataag gggaaaacag cagggaaacg ggaggagggg gatacgtgtg 1320

gtgccgtccg ctctcctat ggaataa 1347

<210> 26
 <211> 662
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 26
 tcaagcagca ggaggagaa gctggaaagc agtagactca gtgatgttcc agcaactgca 60

aatgtagca atgcagcatg gcctcgtgtc cgaggatittt gaaaggcagt tagtatatta 120

tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaatag 180

agctcaaaaa gagttaattc aagggaattt gaatgaggaa gcagaaaggt ggagaagaaa 240

taatccacca cctcaagcag gcggaggatt aacagtggat caaattatgg gggtaggaca 300
 aacaaatcaa gcagcggcac aggctaacat ggatcaggca agacaaatat gtctgcaatg 360
 ggtaataaca gactaagag cagtgagaca tatggctcac aaaccaggga atccaatgct 420
 agtaaagcaa aaaacaaatg agtcatatga agatTTTgcc gcaagactgc tagaagcaat 480
 agatgcagaa ccagttacac agcaaataaa agaatatTTa aagTTaacaT tatcttacac 540
 aatgcatcc tcagactgtc agaaacagat ggatagagta ctaggacaga gagtgcaaca 600
 agctagtgtg gaagaaaaaa tgcaagcatg cagagatgtg ggatcagaag gattcagaat 660
 gc 662

<210> 27
 <211> 662
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 27
 gcaagcagca ggagggagaa gctggaaagc agtagactca gtgatgttcc agcaactgca 60
 aatgtagca atgcagcatg gcctcgtgtc cgaggatTTT gaaaggcagt tagcatatta 120
 tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaatag 180
 agctcaaaaa gagTTaattc aagggaatt gaatgaggaa gcagaaaggt ggagaaggaa 240
 taatccacca cctcaagcag gcggaggatt aacagtggat caaattatgg gggtaggaca 300
 aacaaatcaa gcagcggcac aggctaacat ggatcaggca agacaaatat gcctgcaatg 360
 ggtaataaca gactaagag cagtgagaca tatggctcac aaaccaggga atccaatgct 420
 agtaaagcaa aagacaaatg agtcatatga agatTTTgcc gcaagactgc tagaagcaat 480

agatgcagaa ccagttacac agcaaataaa agaataatta aagttaacat tatcttacac 540

aaatgcatcc tcagactgtc aaaaacaaat ggatagaata ctaggacaga gagtgaaca 600

agctagtgtg gaagaaaaaa tgcaagcatg cagagatgtg ggatcagaag ggttcagaat 660

gc 662

<210> 28
 <211> 651
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 28
 gcaagcagca ggagggagaa gctggaaagc agtagactca gtgatgttcc agcaactgca 60

aaatgtagca atgcagcatg gcctcgtgtc cgaggatitt gaaaggcagt tggcatatta 120

tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaacag 180

agctcaaaaa gagttaattc aggggaaatt gaataaggaa gcagaaaggt ggagaaggaa 240

taatccacca cctcaagcac aaggaggatt aacagtggat caaattatgg gggtaggaca 300

aacaaatcag gcagcggcac aggctaacat ggatcaggca agacaaatat gcctgcaatg 360

ggtaataaca gctaagag cagtgagaca tatggctcac aaaccagga atccaatgct 420

agtaaagcaa aagacaaatg agtcatatga agatittgcc gcaagactgc tagaagcaat 480

agatgcagaa ccagttacac agcaaataaa agaataatta aagttaacat tatcttacac 540

aaatgcatcc tcagactgtc agaaacaaat ggatagagta ctaggacaga gagtgaaca 600

agctagtgtg gaagaaaaaa tgcaagcatg cagagatgtg ggatcagaag g 651

<210> 29
 <211> 520
 <212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 29

gcaagcagca ggagggagaa gctggaaagc agtagactca gtgatgttcc agcaactgca 60

aaatgtagca atgcagcatg gcctcgtgtc cgaggatfff gaaaggcagt tagcatatta 120

tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaatag 180

agctcaaaaa gagttaattc aagggaatt gaatgaggaa gcagaaaggt ggagaaggaa 240

taatccacca cctcaagcag gcggaggatt aacagtggat caaattatgg ggtaggaca 300

aacaaatcaa gcagcggcac aggctaacat ggatcaggca agacaaatat gcctgcaatg 360

ggtaataaca gactaagag cagttagaca tatggctcac aaaccagga atccaatgct 420

agtaaagcaa aagacaaatg agtcatatga agatfffgcc gcaagactgc tagaagcaat 480

agatgcagaa ccagttacac agcaataaa gaatatttaa 520

<210> 30

<211> 662

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 30

gcaagcagca ggagggagaa gctggaaagc agtagactca gtgatgttcc agcaactgca 60

aaatgtagca atgcagcatg gcctcgtgtc cgaggatfff gaaaggcagt tagcatatta 120

tgctactacc tggacaagta aagatatatt agaagtattg gccatgatgc ctggaaatag 180

agctcaaaaa gagttaattc aagggaatt gaatgaggaa gcagaaaggt ggagaaggaa 240

taatccacca cctcaagcag gcggaggatt aacagtggat caaattatgg ggtaggaca 300

aacaaatcaa gcagcggcac aggctaacat ggatcaggca agacaaatat gcctgcaatg 360

ggtaataaca gcactaagag cagtgagaca tatggctcac aaaccagga atccaatgct 420
 agtaaagcaa aagacaaatg agtcatatga aaaattttca gcaagactcc tagaagcaat 480
 agatgcagaa ccagttacac agcctataaa agaataatta aagttaacat tatcttacac 540
 aatgcatcc tcagactgtc aaaaacaaat ggatagagta ctaggacaga gagtgcaaca 600
 agctagtgtg gaagaaaaaa tgcaagcatg cagagatgtg ggatcagaag gattcagaat 660

 gc 662

 <210> 31
 <211> 720
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

 <400> 31
 atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60

 gagctcgaaa aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

 ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180

 aaagaggaac ctiggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

 atgtgtgcc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360

 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

 agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

 gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540

 gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

 gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660

gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

<210> 32

<211> 597

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 32

gcagtcgatg ctgtaatggt ccagcaaatg caaacagtag ccatgcagca tggctctgtg 60

tctgaggact ttgaaaggca gttagcatat tgtgctacta cctggacaag taaagatata 120

ttagaagtat tggccatgat gcctggaaat agagctcaaa aagagttaat tcaaggaaaa 180

ttaaacgagg aagcagaaaag gtggagaagg aataatccac cgcctccaca aggaggggga 240

ttaacagtgg atcaaattat ggggatagga caaacaatc aagcagcagc acaagctaac 300

atggatcagg caagacacat atgcctgcaa tgggtaataa cagcattaag agcagtaaga 360

catatggctc acagaccagg gaatccaatg ctcgtaaac aaaaaacaaa tgagccatat 420

gaagagtttg cagcaaaact attagaagca atagatgcag aaccagtaac acagcccata 480

aaagactatc taaagttaac attatcttat acaaatgcgt cctcagactg tcaaaagcaa 540

atggatagag tgctgggaca aagagtgcaa caagctagtg tagacgagaa aatgcaa 597

<210> 33

<211> 597

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 33

gcagtagact cagtaatggt ccagcaactg caaacagtag caatgcagca tggcctcgtg 60

tccgaggatt ttgaaaggca gttggcatat tatgctacta cctggacgag taaagacata 120

ctagaagtat tggccatgat gcctggaaac agagctcaaa aggagttaat tcaagggaaa 180
 ttaaatgaag aggcagaaaag gtggagaaga cataatccac cccctccggc gggaggagga 240
 ttaacagtgg atcaaattat ggggtagga caaacaatc aagcagcagc acaagctaac 300
 atggatcagg caagacaaat atgcctgcaa tgggtaataa cagcattaag agcagtgagg 360
 tatatgactc acaaccagg gaatccaatg ctagtaaac aaaaaacaaa tgaagcatat 420
 gaagagtta cagcgagact gctagaagca atagatgcag agccagtaac acagcccaca 480
 aaagaatata taaaactaac attatcttat acaaatgcat cctcagactg tcaaaagcaa 540
 atggatagag tactaggaca aagagtgcaa caagctagtg tagaagaaaa aatgcaa 597

 <210> 34
 <211> 597
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

 <400> 34
 gcagtcgatg ctgtaatggt ccagcaaatg caaacagtag ccatgcagca tggctctgtg 60
 tctgaggact ttgaaaggca gttagcatat tatgctacta cctggacaag taaagatata 120
 ttagaagtat tggccatgat gcctggaaat agagctcaaa aagagttaat tcaaggaaaa 180
 ttaaatgagg aagcagaaaag gtggagaagg aataatccac cgctccaca gggaggggga 240
 ttaacagtgg atcaaattat ggggatagga caaacaatc aagcagcagc acaagctaac 300
 atggatcagg caagacacat atgcctgcaa tgggtaataa cagcattaag agcagtaaga 360
 catatggctc acagaccagg gaatccaatg ctgtaaac aaaaaacaaa tgagccatat 420
 gaagagtttg cagcaaaact attagaagca atagatgcag aaccagtaac acagctcata 480
 aaagactata taaagttaac attatcttat acaaatgcgt cctcagactg tcaaaagcaa 540

atggatagag tgctgggaca aagagtgcaa caagctagtg tagacgagaa gatgcaa 597

<210> 35

<211> 1347

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 35

atggtgagtc tagatagaga catggcgagg caagtctccg ggggaaaag agattatcct 60

gagctcgaag aatgtatcaa gcatgcatgc aagataaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180

aaagaggaac ctggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

atgtgttgc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360

atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540

gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660

gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

caaatatgg gggtaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 780

aggcaaatat gcctgcaatg ggtaataaat gcattaagag cagtaagaca tatggcgcac 840

aggccaggga atccaatgct agtaaagcaa aaaacgaatg agccatatga agatattgca 900

gcaagactgc tagaagcaat agatgcagag ccagttacac agcctataaa agattatcta 960

aagctaacac taicttatac aaatgcatca gcagattgtc agaagcaaat ggatagaaca 1020

ctaggacaaa gagtacaaca agctagtgtg gaagaaaaaa tgcaagcatg tagagatgtg 1080

ggatcagaag ggttcaaaat gcaattgtta gcacaagcat taaggccagg aaaaggaaaa 1140

gggaatggac agccacaaag gtgttacaac tgtggaaaac cgggacatca agcaaggcaa 1200

tgtagacaag gaatcatatg tcacaactgt ggaaagagag gacatatgca aaaagaatgc 1260

agaggaaaga gagacataag gggaaaacag cagggaaacg ggaggagggg gatacgtgtg 1320

gtgccgtccg ctctctctat ggaataa 1347

<210> 36

<211> 597

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 36

gcagtcgatg ctgtaatgtt ccagcaaatg caaacagtag ccatgcagca tggctctgtg 60

tctgaggact ttgaaaggca gttagcatat tgtgctacta cctggacaag taaagatata 120

ttagaagtat tggccatgat gcctggaaat agagctcaaa aagagttaat tcaaggaaaa 180

ttaaacgagg aagcagaaaag gtggagaagg aataatccac cgcctccaca aggaggggga 240

ttaacagtgg atcaaatat ggggatagga caaacaatc aagcagcagc acaagctaac 300

atggatcagg caagacacat atgcctgcaa tgggtaataa cagcattaag agcagtaaga 360

catatggctc acagaccagg gaatccaatg ctcgtaaaac aaaaaacaaa tgagccatat 420

gaagagtttg cagcaaaact attagaagca atagatgcag aaccagtaac acagcccata 480

aaagactatc taaagttaac attatcttat acaaatgcgt cctcagactg tcaaaagcaa 540

atggatagag tgctgggaca aagagtgcaa caagctagtg tagacgagaa aatgcaa 597

- <210> 37
- <211> 597
- <212> DNA
- <213> Caprine arthritis-encephalitis virus

<400> 37
gcagtagact cagtaatggt ccagcaactg caaacagtag caatgcagca tggcctcgtg 60

tccgaggatt ttgaaaggca gttggcatat tatgctacta cctggacgag taaagacata 120

ctagaagtat tggccatgat gcttggaaac agagctcaaa aggagttaat tcaagggaaa 180

ttaaatagaag aggcagaaaag gtggagaaga cataatccac cccctccggc gggaggagga 240

ttaacagtgg atcaaattat ggggtagga caaacaaatc aagcagcagc acaagctaac 300

atggatcagg caagacaaat atgcctgcaa tgggtaataa cagcattaag agcagtgagg 360

tatatgactc acaaaccagg gaatccaatg ctagtaaac aaaaaacaaa tgaagcatat 420

gaagagtta cagcgagact gctagaagca atagatgcag agccagtaac acagcccaca 480

aaagaatatc taaaactaac attatcttat acaaatgcat cctcagactg tcaaaagcaa 540

atggatagag tactaggaca aagagtgcaa caagctagtg tagaagaaaa aatgcaa 597

- <210> 38
- <211> 597
- <212> DNA
- <213> Caprine arthritis-encephalitis virus

<400> 38
gcagtcgatg ctgtaatggt ccagcaaatg caaacagtag ccatgcagca tggctctgtg 60

tctgaggact ttgaaaggca gttagcatat tatgctacta cctggacaag taaagatata 120

ttagaagtat tggccatgat gcctggaaat agagctcaaa aagagttaat tcaaggaaaa 180
 ttaaatgagg aagcagaaaag gtggagaagg aataatccac cgctccaca gggaggggga 240
 ttaacagtgg atcaaattat ggggatagga caaacaatc aagcagcagc acaagctaac 300
 atggatcagg caagacacat atgcctgcaa tgggtaataa cagcattaag agcagtaaga 360
 catatggctc acagaccagg gaatccaatg ctcgtaaac aaaaaacaaa tgagccatat 420
 gaagagtttg cagcaaaact attagaagca atagatgcag aaccagtaac acagctcata 480
 aaagactatc taaagttaac attatcttat acaaatgcgt cctcagactg tcaaaagcaa 540
 atggatagag tgctgggaca aagagtgcaa caagctagtg tagacgagaa gatgcaa 597

 <210> 39
 <211> 720
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

 <400> 39
 atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60
 gagctgaaa aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120
 ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180
 aaagaggaac ctggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240
 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300
 atgtgtgcc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360
 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420
 agtgtcttcc caatagtagt gcaagcagca ggagggagaa gctggaaagc agtagattct 480
 gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540

gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaatt aatgaagaa 660

gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

<210> 40

<211> 524

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 40

taaagatata ttagaagtat tggccatgat gcctggaaat agagctcaa aagagttaat 60

tcaagggaaa ttgaatgaag aagcagaaag gtggagaagg aataatccac cacctcaagc 120

aggcggagga ttaacagtgg atcaaattat ggggtagga caaacaaatc aagcagcggc 180

acaggctaac atggatcagg caagacaaat atgcctgcaa tgggtaataa cagcactaag 240

agcagtgaga catatggctc acaaaccagg gaatccgatg ctagttaaagc aaaaaacaaa 300

tgagtcatat gaagattttg ccgcaagact gctagaagca atagatgcag aaccagttac 360

aaagcaaata aaagaatatt taaagttaac attatcttac acaaatgcat cctcagactg 420

taagaaacag atggatagag tactaggaca gagagtgcaa caagctagtg tggaagaaaa 480

aatgcaagca tgcagagatg tgggatcaga aggattcaga atgc 524

<210> 41

<211> 1347

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 41

atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60

gagctcgaaa aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atgggtgcctt aaaacattag attacatggt tgaggacat 180

aaagaggaac ctiggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

atgtgttgc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360

atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

glaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540

gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660

gcagaaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

caaattatgg gggtaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 780

aggcaaatat gcctgcaatg ggtaataaat gcattaagag cagtaagaca tatggcgcac 840

aggccaggga atccaatgct agtaaagcaa aaaacgaatg agccatatga agattttgca 900

gcaagactgc tagaagcaat agatgcagag ccagttacac agcctataaa agattatcta 960

aagctaacac tatcttatac aatgcatca gcagattgtc agaagcaaat ggatagaaca 1020

ctaggacaaa gagtacaaca agctagtgtg gaagaaaaaa tgcaagcatg tagagatgtg 1080

ggatcagaag ggttcaaaat gcaattgtta gcacaagcat taaggccagg aaaaggaaaa 1140

gggaatggac agccacaaag gtgttacaac tgtggaaaac cgggacatca agcaaggcaa 1200

tgtagacaag gaatcatatg tcacaactgt ggaaagagag gacatatgca aaaagaatgc 1260

agaggaaaga gagacataag gggaaaacag cagggaacg ggaggaggg gatacgtgtg 1320

gtgccgtccg ctctcctat ggaataa 1347

<210> 42

<211> 524

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 42

taaagatata ttagaagtat tggccatgat gcctggaat agagctcaa aagagttaat 60

tcaagggaaa ttgaatgaag aagcagaaag gtggagaagg aataatccac cacctcaagc 120

aggcggagga ttaacagtgg atcaaattat ggggtagga caaacaaatc aagcagcggc 180

acaggetaac atggatcagg caagacaaat atgcctgcaa tgggtaataa cagcactaag 240

agcagtgaga catatggctc acaaaccagg gaatccgatg ctagtaaagc aaaaaacaaa 300

tgagtcatat gaagattttg ccgcaagact gctagaagca atagatgcag aaccagttac 360

aaagcaaata aaagaatatt taaagttaac attatcttac acaaatgcat cctcagactg 420

taagaaacag atggatagag tactaggaca gagagtgcaa caagctagtg tggaagaaaa 480

aatgcaagca tgcagagatg tgggatcaga aggattcaga atgc 524

<210> 43

<211> 720

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 43

atggtgagtc tagatagaga catggcgagg caagtctccg ggggaaaag agattatcct 60

gagctcgaat aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atgggtcctt aaaacattag attacatggt tgaggacat 180
 aaagaggaac ctggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240
 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300
 atgtgttgc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360
 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420
 agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480
 gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540
 gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600
 gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660
 gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

 <210> 44
 <211> 593
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

 <400> 44
 tgccgtagac tctgtgatgt tccaccagct gcatacagta gcaatgccgc atggcctcgt 60
 gtctgaggac ttgaaaggc agttggcata ttatgctact acctggacaa gtaaagatat 120
 actggaagta ttggccatga tgcctgggaa tagagctcaa aaagaattaa ttcaaggaaa 180
 attaaatgaa gaagcagaaa ggtggagaag gaataatcca ccacctcaag caggcggagg 240
 attaacagtg gatcaaatta tggggtagg acaaacaaat caagcagctg cacaagctaa 300
 catggatcag gcaagacaaa tatgectgca atgggtaata tcagccttaa gagcagtgag 360
 acatatgtct cataaaccag ggaatccgct gctagtaaag caaaaaacaa atgagtcata 420

tgaagatfff gcagctagac tgctagaagc aatagatcca gccccagtag cacatcctat 480

aaaagattat ttaaagttaa cactatctta tacgaatgca tcatcagatt gtcaaaagca 540

aatgggtaga atgctaggat cgagagtcca tcaagccagt gtgggcaaaa aaa 593

<210> 45
 <211> 1347
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 45
 atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60

gagctcgaaa aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180

aaagaggaac ctiggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

atgtgttgc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360

atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540

gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660

gcagaaaagt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

caaattatgg gggtaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 780

aggcaaatat gcctgcaatg ggtaataaat gcattaagag cagtaagaca tatggcgcac 840

aggccagga atccaatgct agtaaagcaa aaaacgaatg agccatatga agattttgca 900

gcaagactgc tagaagcaat agatgcagag ccagttacac agcctataaa agattatcta 960

aagctaacac tatcttatac aaatgcatca gcagattgtc agaagcaaat ggatagaaca 1020

ctaggacaaa gagtacaaca agctagtgtg gaagaaaaaa tgcaagcatg tagagatgtg 1080

ggatcagaag ggttcaaat gcaattgtta gcacaagcat taaggccagg aaaaggaaaa 1140

gggaatggac agccacaag gtgttacaac tgtggaaaac cgggacatca agcaaggcaa 1200

tgtagacaag gaatcatatg tcacaactgt ggaaagagag gacatatgca aaaagaatgc 1260

agaggaaaga gagacataag gggaaaacag cagggaaacg ggaggagggg gatacgtgtg 1320

gtgccgtccg ctctctctat ggaataa 1347

<210> 46
 <211> 593
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 46

tgccgtagac tctgtgatgt tccaccagct gcatacagta gcaatgccgc atggcctcgt 60

gtctgaggac ttgaaaggc agttggcata ttatgtact acctggacaa gtaaagatat 120

actggaagta ttggccatga tgcctgggaa tagagctcaa aaagaattaa ttcaaggaaa 180

attaaatgaa gaagcagaaa ggtggagaag gaataatcca ccacctcaag caggcggagg 240

attaacagtg gatcaaatta tggggtagg acaaacaaat caagcagctg cacaagctaa 300

catggatcag gcaagacaaa tatgectgca atgggtaata tcagccttaa gagcagtgag 360

acatatgtct cataaaccag ggaatccgct gctagtaaag caaaaaacaa atgagtcata 420

tgaagatfff gcagctagac tgctagaagc aatagatcca gccccagtag cacatcctat 480
 aaaagattat ttaaagttaa cactatctta tacgaatgca tcatcagatt gtcaaaagca 540
 aatgggtaga atgctaggat cgagagtcca tcaagccagt gtgggcaaaa aaa 593

<210> 47
 <211> 720
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 47
 atggtgagtc tagatagaga catggcgagg caagtctccg ggggaaaag agattatcct 60
 gagctcgaaa aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120
 ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180
 aaagaggaac ctiggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240
 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300
 atgtgttgc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360
 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420
 agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480
 gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540
 gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600
 gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660
 gcagaaaagt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

<210> 48
 <211> 593

<212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 48
 agcagtagat tctgtaatgt tccagcaact gcaaacagta gcaatgcagc atggactcgt 60

 gtatgaagac ttgaaaggc tgctggcata ttatgctact acctggacaa gtaaagatat 120

 actggaagta ttggccatga tgcctgggaa tagagctcaa aaagaattaa ttcaaggaaa 180

 attaaatgaa gaagcagaaa ggtggagaag gaataatcca ccacctcaag caggcggagg 240

 attaacagt gatcaaatta tggggtagg acaaacaaat caagcagctg cacaagctaa 300

 catggatcag gcaagacaaa tatgcctgca atgggtaata tcagccttaa gagcagtgag 360

 acatatgtct cataaaccag ggaatccgct gctagtaaag caaaaaacaa atgagtcata 420

 tgaagatfff gcagcaagac tgctagaagc aatagatgca gagccagtag cacatcctat 480

 aaaagaatac ttaaagttaa cactatctta tacgaatgca tcatcagatt gtcaaaagca 540

 aatggataga atgctggaat caagagtaca acaagctagt gtagaacaaa aaa 593

<210> 49
 <211> 593
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 49
 agccgtagat tctgtaatgt tccagcagct gcaaacagta gcaatgcagc atggcctcgt 60

 gtcagaggac ttgaaaggc ttccagcata tcatgctact acctgggcaa gtaaagatat 120

 cttagaagta ctggccatga tgcctggaaa tagagctcaa aaagagttaa ttcaaggaa 180

 attaaatgaa gaagcagaga ggtggagaag gaataatcca ccacctccag caggaggagg 240

 gttaacagt gatcaaatta tgggtagtagg acaaacaaat caggcagcgg cacaagcaaa 300

catggatcag gcaagacaaa tatgcctaca atgggtgata tcagcactaa gagcagtaag 360
 gcataatggct cacaagccag ggaatccaat gttagtaaag caaaaagcaa atgagccata 420
 tgaagaattt gcagcaaggc tgctggaagc aatagatgcc gagccagtta atcagcccat 480
 aaaagaatat ctaaaactaa cgttgtctta tacgaatgca tcctcagatt gtcagaagca 540
 aatggataga acactaggac aaagagtcaa acaagctagt gtagaacaaa aaa 593

<210> 50
 <211> 1347
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 50
 atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60
 gagctcgaaa aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120
 ttgacagaag gaaattgttt atgggtgcctt aaaacattag attacatggt tgaggacat 180
 aaagaggaac ctiggacaaa agtaaaattt aggacaatat ggcaagaagt gaagaatcta 240
 actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300
 atgtgttgcc aaatggggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360
 atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420
 agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480
 gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540
 gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600
 gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660
 gcagaaaagt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

caaattatgg gggtaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 780
 aggcaaatat gcctgcaatg ggtaataaat gcattaagag cagtaagaca tatggcgcac 840
 aggccagga atccaatgct agtaaagcaa aaaacgaatg agccatatga agattttgca 900
 gcaagactgc tagaagcaat agatgcagag ccagttacac agcctataaa agattatcta 960
 aagctaacac tatcttatac aatgcatca gcagattgtc agaagcaaat ggatagaaca 1020
 ctaggacaaa gagtacaaca agctagtgtga gaagaaaaaa tgcaagcatg tagagatgtg 1080
 ggatcagaag ggttcaaaat gcaattgtta gcacaagcat taaggccagg aaaaggaaaa 1140
 gggaatggac agccacaaag gtgttacaac tgtggaaaac cgggacatca agcaaggcaa 1200
 tgtagacaag gaatcatatg tcacaactgt ggaaagagag gacatatgca aaaagaatgc 1260
 agaggaaaga gagacataag gggaaaacag cagggaacg ggaggagggg gatacgtgtg 1320
 gtgccgtccg ctctcctat ggaataa 1347

<210> 51
 <211> 593
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 51
 agcagtagat tctgtaatgt tccagcaact gcaaacagta gcaatgcagc atggactcgt 60
 gtatgaagac tttgaaagc tgtcggcata ttatgctact acctggacaa gtaaagatat 120
 actggaagta ttggccatga tgctgggaa tagagctcaa aaagaattaa ttcaaggaaa 180
 attaaatgaa gaagcagaaa ggtggagaag gaataatcca ccacctcaag caggcggagg 240
 attaacagtg gatcaaatta tggggtagg acaacaaat caagcagctg cacaagctaa 300

catggatcag gcaagacaaa tatgcctgca atgggtaata tcagccttaa gagcagtgag 360
 acatatgtct cataaaccag ggaatccgct gctagtaaag caaaaaacaa atgagtcata 420
 tgaagatfff gcagcaagac tgctagaagc aatagatgca gagccagtag cacatcctat 480
 aaaagaatat ctaaaactaa cgttgtctta tacgaatgca tcctcagatt gtcagaagca 540
 aatggataga aactaggac aaagagtcaa acaagctagt gtagaacaaa aaa 593

<210> 52
 <211> 593
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 52
 agccctagat tctgtaatgt tccagcagct gcaaacagta gcaatgcagc atggcctcgt 60
 gtcagaggac ttgaaaggc ttccagcata tcatgctact acctgggcaa gtaaagatat 120
 cttagaagta ctggccatga tgcctggaaa tagagctcaa aaagagttaa ttcaagggaa 180
 attaatgaa gaagcagaga ggtggagaag gaataatcca ccacctccag caggaggagg 240
 gttaacagtg gatcaaatta tgggagtagg acaaacaaat caggcagcgg cacaagcaaa 300
 catggatcag gcaagacaaa tatgcctaca atgggtgata tcagcactaa gagcagtaag 360
 gcataatggct cacaagccag ggaatccaat gttagtaaag caaaaagcaa atgagccata 420
 tgaagaatff gcagcaaggc tgctggaagc aatagatgcc gagccagtta atcagcccat 480
 aaaagaatat ctaaaactaa cgttgtctta tacgaatgca tcctcagatt gtcagaagca 540
 aatggataga aactaggac aaagagtcaa acaagctagt gtagaacaaa aaa 593

<210> 53
 <211> 720
 <212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 53

atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60

gagctcgaaa aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atggtgcctt attacatgtt tgaggacat aaagaggaac 180

cttgacaaa aaaacattag agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

atgtgttgcc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360

atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

gtaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540

gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660

gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

<210> 54

<211> 129

<212> DNA

<213> Caprine arthritis-encephalitis virus

<400> 54

cagcatggcc tcgtgtccga ggactttgaa aggcagttgg catattatgc tactacctgg 60

acaagtaagg acatattaga agtattggcc atgatgccag gaaatagagc tcaaaaggag 120

ctaattcaa 129

<210> 55
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 55
 cagcatggaa tagtatcaga agagtttgag aggcaactgt cttattatgc taccacttgg 60

 acaagcaagg atatcttaga ggtactagcc atgatgctg gcaatagagc attaaaagag 120

 ctaatacaa 129

<210> 56
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 56
 caacatggga tagtatcaga ggaatttgag agacaaatgt cttattatgc taccacatgg 60

 acaagtaagg atatcttaga agtactagca atgatgcccg ggaacagagc attaaaggag 120

 ctgatatacaa 129

<210> 57
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 57
 cagcatggac tagtatcaga agaatttgaa aggcagctag catactatgc cacaacttgg 60

 acaagcaaag acatactaga ggtgttagcc atgatgccag ggaatagagc acaaaaagaa 120

 ctaatacaa 129

<210> 58
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 58
cagcatggac ttgtgtccga agattttgag aggcaattgg catattatgc tacaacctgg 60

actagtgaag atatattaga agtattggct atgatgcctg ggaatagagc acagaaagaa 120

ttaatacaa 129

<210> 59
<211> 1347
<212> DNA
<213> Caprine arthritis-encephalitis virus

<400> 59
atggtgagtc tagatagaga catggcgagg caagtctccg gggggaaaag agattatcct 60

gagctcgaag aatgtatcaa gcatgcatgc aagataaaag ttcgactcag aggggagcac 120

ttgacagaag gaaattgttt atggtgcctt aaaacattag attacatggt tgaggacat 180

aaagaggaac ctggacaaa agtaaaattt aggacaatat ggcagaaggt gaagaatcta 240

actcctgagg agagtaacaa aaaagacttt atgtctttgc aggccacatt agcgggtcta 300

atgtgttgcc aaatgggat gagacctgag acattgcaag atgcaatggc tacagtaatc 360

atgaaagatg ggttactgga acaagaggaa aagaaggaag acaaaagaga aaaggaagag 420

agtgtcttcc caatagtagt gcaagcagca ggaggagaa gctggaaagc agtagattct 480

glaatgttcc agcaactgca aacagtagca atgcagcatg gcctcgtgtc tgaggacttt 540

gaaaggcagt tggcatatta tgctactacc tggacaagta aagacatact agaagtattg 600

gccatgatgc ctggaaatag agctcaaaag gagttaattc aagggaattt aatgaagaa 660

gcagaaaggt ggagaaggaa taatccacca cctccagcag gaggaggatt aacagtggat 720

caaattatgg gggtaggaca aacaaatcaa gcagcagcac aagctaacat ggatcaggca 780

aggcaaatat gcctgcaatg ggtaataaat gcattaagag cagtaagaca tatggcgcac 840

aggccagga atccaatgct agtaaagcaa aaaacgaatg agccatatga agattttgca 900

gcaagactgc tagaagcaat agatgcagag ccagttacac agcctataaa agattatcta 960

aagctaacac tatcttatac aaatgcatca gcagattgtc agaagcaaat ggatagaaca 1020

ctaggacaaa gagtacaaca agctagtgtg gaagaaaaaa tgcaagcatg tagagatgtg 1080

ggatcagaag ggttcaaat gcaattgtta gcacaagcat taaggccagg aaaaggaaaa 1140

gggaatggac agccacaaag gtgttacaac tgtggaaaac cgggacatca agcaaggcaa 1200

tgtagacaag gaatcatatg tcacaactgt ggaagagag gacatatgca aaaagaatgc 1260

agaggaaaga gagacataag gggaaaacag cagggaaacg ggaggagggg gatacgtgtg 1320

gtgccgtccg ctctctctat ggaataa 1347

<210> 60
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 60
 cagcatggcc tcgtgtccga ggactttgaa aggcagttgg catattatgc tactacctgg 60

acaagtaagg acatattaga agtattggcc atgatgccag gaaatagagc tcaaaaggag 120

ctaattcaa 129

<210> 61
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 61
 cagcatggaa tagtatcaga agagtttgag aggcaactgt cttattatgc taccacttgg 60

acaagcaagg atatcttaga ggtactagcc atgatgcctg gcaatagagc attaaaagag 120

ctaatacaa 129

<210> 62
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 62
 caacatggga tagtatcaga ggaatttgag agacaaatgt cttattatgc taccacatgg 60

acaagtaagg atattttaga agtactagca atgatgcccg ggaacagagc attaaaggag 120

ctgatacaa 129

<210> 63
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 63
 cagcatggac tagtatcaga agaatttgaa aggcagctag catactatgc cacaacgtgg 60

acaagcaaag acatactaga ggtgttagcc atgatgccag ggaatagagc acaaaaagaa 120

ctaatacaa 129

<210> 64
 <211> 129
 <212> DNA
 <213> Caprine arthritis-encephalitis virus

<400> 64
 cagcatggac ttgtgtccga agattttgag aggcaattgg catattatgc tacaacctgg 60

actagtgaag atatattaga agtattggct atgatgcctg ggaatagagc acagaaagaa 120

ttaatacaa 129

<210> 65
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer used to construct dig-labeled probe

<400> 65
 ctggcgtaat agcgaagagg 20

<210> 66
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer used to construct dig-labeled probe

<400> 66
 aactcgccgc acatctgaac 20

<210> 67
 <211> 3911
 <212> DNA
 <213> Artificial

<220>
 <223> pCAH/SINd0

<400> 67
 atccagcaca gtggcggccg ctagcacaaa aataaaaaaa gaaaggggtgg actgtgagac 60

atgggctaaa gaggagcggc cgctcgagtc tagaactagt ggatcagctt tgctgcttgc 120

acttcagagt tctaggagag tccctcctag tctctctctt ccgaggaggt accgagacct 180

caaaataaag gagtgattgc cttactgccg agtggagagt gattactgag cggccggtgt 240

atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgagcc gacctcgagg 300

gggggccta ttctatagtg tcacctaaat gctagagctc gctgatcagc ctcgactgtg 360
 ccttctagtt gccagccatc tgttgtttgc cctcccccg tgccttcctt gaccctggaa 420
 ggtgccactc ccaactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 480
 aggtgtcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggga ggattgggaa 540
 gacaatagca ggcattgctg ggatgcggtg ggctctatgg cttctgagc ggaaagaacc 600
 agtggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 660
 aaaagccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 720
 gctccgccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaacc 780
 gacaggacta taaagatacc aggcgtttcc cctggaagc tccctcgtgc gctctcctgt 840
 tccgacctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgt 900
 ttctcatage tcacgtgta ggtatctcag ttcggtgtag gtcgttcgt ccaagctggg 960
 ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc ttatccgta actatcgtct 1020
 tgagtccaac ccgtaagac acgacttatc gccactggca gcagccactg gtaacaggat 1080
 tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtggggc ctaactacgg 1140
 ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagta ccttcggaaa 1200
 aagagtggg agctcttgat ccggcaaaca aaccaccgct ggtagcggg gttttttgt 1260
 ttgcaagcag cagattacgc gcagaaaaa aggatctcaa gaagatcctt tgatctttc 1320
 tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttg tcatgagatt 1380
 atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 1440

aagtatatat gagtaacctg atcaggactc ttccttttca tgaacaataa aactgtctgc 1500
ttacataaac agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc 1560
taggccgcga ttaaattcca acatggatgc tgatttatat gggataaat gggctcgcga 1620
taatgtcggg caatcaggtg cgacaatcia tcgattgtat ggaagcccg atgcgccaga 1680
gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag 1740
actaaactgg ctgacggaat ttatgcctct tccgaccatc aagcatttta tccgtactcc 1800
tgatgatgca tggttactca ccactgcgat ccccgggaaa acagcattcc aggtattaga 1860
agaatatacct gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt 1920
gcattcgatt cctgtttgta attgtccttt taacagcgat cgcgtatttc gtctcgtca 1980
ggcgcaatca cgaatgaata acggtttggt tgatgcgagt gattttgatg acgagcgtaa 2040
tggtcggcct gtgaacaag tctggaaaga aatgcataaa cttttgccat tctcacggga 2100
ttcagtcgtc acicatggtg atttctcact tgataacctt atttttgacg aggggaaatt 2160
aatagttgtt attgatgttg gacgagtcgg aatcgagac cgataccagg atcttgccat 2220
cctatggaac tgcctcggtg agttttctcc ttcattacag aaacggcttt tcaaaaata 2280
tggattgat aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt 2340
ctaagaattc gcgcaattaa ccctactaa agggaacaaa agctgggtac cgggcccggtt 2400
gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 2460
catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgcca 2520
acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagga 2580
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 2640

aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggccccct 2700

ggcattatgc ccagttacatg accttatggg actttcctac ttggcagtac atctacgtat 2760

tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 2820

ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgttt 2880

ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa 2940

tggcggttag gcgtgtacgg tgggaggtct atataagcag agctcaagct tgctgcttgc 3000

acttcagagt tctaggagag tcctccttag tctctcctct ccgaggaggt accgagacct 3060

caaaataaag gagtgattgc cttactgccg agtggagagt gattactgag cggccggtgt 3120

atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgcagct ggcgccaac 3180

gtggggcccc aggagaagaa aagaaaaggc cctgagaac tcggcttctg aaaaagagga 3240

agaggacaag ttgctatagc aacaagagag aagaagtaga gcaaaggtcc agtggctcgg 3300

aaaaagagga actgaaactt cggggacgcc tgaaggagta aggttaagtga ctctgctgta 3360

cgcgggcgga ggcagaggtt tccttctaaa ttgaaagaga agtgttgctg cgagaggtct 3420

tggtgtcga gaatcctgta caaaaaaag gagggatctc ggtcaggacc aggaccctg 3480

ggagtaatac aacagcaaca ccgtaagaaa atccgcctag ggaattcgat tctagaggtg 3540

atagaaatgc cagaaaacta tgcaaaaaca agaatacataa acaggaaaaa aagagaactc 3600

agccacaaga ggaagaagag aggcgttggc ttggtcatta tgctagtat catggcaata 3660

gtagctgccg caggggcttc tctgggagtc gcaaacgca ttcagcagtc ttacactaag 3720

gcagctgtcc agacccttgc taatgcaact getgcacagc aggatgtgtt agaggcaacc 3780

tatgccatgg tacagcatgt ggctaaaggc gtacgaatct tggaaagctcg agtggctcga 3840

gtggaagcta tcacagatag aataatgcta taccaagaat tggattgttg gcactaggat 3900

ccatcgccac c 3911

<210> 68
 <211> 4238
 <212> DNA
 <213> Artificial

<220>
 <223> pCAH/SINd1

<400> 68
 atccagcaca gtggcggccg ctagcacaaa aataaaaaaa gaaaggggtg actgtgagac 60

atgggctaaa gaggagcggc cgctcgagtc tagaactagt ggatcagctt tgctgcttgc 120

acttcagagt tctaggagag tccctcctag tctctctct ccgaggaggt accgagacct 180

caaaataaag gagtgattgc cttactgccg agtggagagt gattactgag cggccggtgt 240

atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgagcc gacctcgagg 300

gggggccta ttctatagtg tcacctaaat gctagagctc gctgatcagc ctcgactgtg 360

ccttctagtt gccagccatc tgttgtttgc cctcccccg tgccttcctt gaccctggaa 420

ggtgccactc cactgtcct ttctaataa aatgaggaaa ttgcatcgca ttgtctgagt 480

aggtgtcatt ctattctggg ggggtggggtg gggcaggaca gcaaggggga ggattgggaa 540

gacaatagca ggcattgctg ggatgcggtg ggctctatgg cttctgagc ggaaagaacc 600

agtggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 660

aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt tttccatag 720

gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 780
 gcacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 840
 tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 900
 ttctcatagc tcacgctgta ggtatctcag ttcgggtgtag gtcgttcgct ccaagctggg 960
 ctgtgtgcac gaaccccccg ttcagccccga ccgctgcgcc ttatccgta actatcgtct 1020
 tgagtccaac ccgtaagac acgacttatc gccactggca gcagccactg gtaacaggat 1080
 tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtgc ctaactacgg 1140
 ctacactaga aggacagtat ttggtatctg cgtctctctg aagccagtta ccttcggaaa 1200
 aagagtgggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gttttttgt 1260
 ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 1320
 tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt 1380
 atcaaaaagg atcttcacct agatcctttt aaattaataaa tgaagtttta aatcaatcta 1440
 aagtatatat gagtaacctg atcaggactic ttccctttca tgaacaataa aactgictgc 1500
 ttacataaac agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc 1560
 taggcccgga ttaaattcca acatggatgc tgatttatat gggataaat gggctcgca 1620
 taatgctggg caatcaggtg cgacaatcta tcgattgtat ggaagccccg atgcgccaga 1680
 gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag 1740
 actaaactgg ctgacggaat ttatgectct tccgaccatc aagcatttta tccgtactcc 1800
 tgatgatgca tggttactca ccaactcgat ccccgggaaa acagattcc aggtattaga 1860
 agaatacctt gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt 1920

gcattcgatt cctgtttgta attgtccttt taacagegat cgcgtatttc gtctcgtca 1980

ggcgcaatca cgaatgaata acggtttggt tgatgcgagt gattttgatg acgagcgtaa 2040

tggctggcct gtigaacaag tctggaaaga aatgcataaa cttttgcat tctcacgga 2100

ttcagtcgtc acatcatggtg atttctcact tgataacctt atttttgacg aggggaaatt 2160

aatagttgt attgatgtg gacgagtcgg aatcgagac cgataccagg atcttgcat 2220

cctatggaac tgctcggtg agttttctcc tcattacag aaacggcttt ttcaaaaata 2280

tgtattgat aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagtttt 2340

ctaagaattc gcgcaattaa cctcactaa agggaacaaa agctgggtac cgggcccgtt 2400

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 2460

catatatgga gttccgctt acataactta cggtaaatgg cccgcctggc tgaccgcca 2520

acgacccccg cccattgacg tcaataatga cgtatgtcc catagtaacg ccaatagga 2580

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 2640

aagtgtatca tatccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct 2700

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 2760

tagtcatgc tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 2820

ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt 2880

ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa 2940

tggcggtag gcgtgtacgg tgggaggtct atataagcag agctcaagct tgctgcttgc 3000

acttcagagt tctaggagag tcctcctag tctctcctct ccgaggaggt accgagacct 3060

caaaataaag gagtattgc cttactgccg agtggagagt gattactgag cggccggtgt 3120

atcgggagtc gtcccttaat ctgtgcaata ccagagcggc tctcgcagct ggcgccaac 3180

gtggggcccg aggagaagaa aagaaagcgg ccctgagaac tggcttctg aaaaagagga 3240

agaggacaag ttgctatagc aacaagagag aagaagtaga gcaaaggtcc agtggctcgg 3300

aaaaagagga actgaaactt cggggacgcc tgaaggagta aggtaagtga ctctgctgta 3360

cgcgggcgca ggcagaggtt tccttctaaa ttgaaagaga agtgttgctg cgagaggtct 3420

tggtgtcga gaatcctgta caaaaaaag gagggatctc ggtcaggacc aggaccctg 3480

ggagtatac aacagcaaca ccgtaagaaa atccgcctag gtgagtctag atagagacta 3540

ggcgaggcaa gtctccgggg ggaaaagaga ttatcctgag ctcgaaaaat gtatcaagca 3600

tgcatgcaag ataaaagttc gactcagagg ggagcacttg acagaaggaa atgtttatg 3660

gtgccttaa acattagatt acatgtttga ggaccataaa gaggaacctt ggacaaaagt 3720

aaaatttagg acaatatggc agaagtgaa gaatctaact cctgaggaga gtaacaaaa 3780

agactttatg tctttgcagg ccacattagc ggttctaag tgttgccaaa tgggatgag 3840

acctgcagga attcgattct agaggtgata gaaatgccag aaaactatgc aaaaacaaga 3900

atcataaaca ggaaaaaag agaactcagc cacaagagga agaagagagg cgttgcttg 3960

gtcattatgc tagttatcat ggcaatagta gctgccgcag gggcttctct gggagtcgca 4020

aacgcgattc agcagtctta cactaaggca gctgtccaga ccttgctaa tgcaactgct 4080

gcacagcagg atgtgttaga ggcaacctat gccatgttac agcatgtgac taaaggcgt 4140

cgaatcttg aagctcgagt ggctcgagt gaagctatca cagatagaat aatgctatac 4200

caagaattgg atgttgca ctaggatcca tcgccacc 4238

<210> 69
 <211> 4523
 <212> DNA
 <213> Artificial

<220>
 <223> pCAH/SINd2

<400> 69
 atccagcaca gtggcggccg ctacacaaaa aataaaaaaa gaaagggtgg actgtgagac 60

 atgggctaaa gaggagcggc cgctcgagtc tagaactagt ggatcagctt tgctgcttgc 120

 acttcagagt tctaggagag tcctcctag tctctctct ccgaggaggt accgagacct 180

 caaaataaag gagtgattgc cttactgccg agtggagagt gattactgag cggccggtgt 240

 atcgggagtc gtccttaat ctgtcaata ccagagcggc tctcgagcc gacctcgagg 300

 gggggccta ttctatagtg tcacctaaat gctagagctc gctgatcagc ctcgactgtg 360

 ctttctagtt gccagccatc tgttgtttgc cctcccccg tgccttcctt gaccctggaa 420

 ggtgccactc cactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 480

 aggtgtcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggga ggattgggaa 540

 gacaatagca ggcattctgg ggatgcggtg ggctctatgg cttctgagcc ggaaagaacc 600

 agtggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 660

 aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 720

 gctccgccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaacct 780

 gacaggacta taaagatacc aggcgtttcc ccttgggaagc tcctctgtgc getctctgt 840

 tccgacctg ccgettaccg gataactgtc cgcctttctc ccttcgggaa gcgtggcgt 900

ttctcatage tcacgtgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 960
 ctgtgtgcaac gaaccccccg ttcagcccga ccgctgcgcc ttatccggtta actatcgtct 1020
 tgagtccaac ccggttaagac acgacttata gccactggca gcagccactg gtaacaggat 1080
 tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtgc ctaactacgg 1140
 ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 1200
 aagagtgggt agctcttgat ccggcaaaaa aaccaccgct ggtagcggtg gttttttgt 1260
 ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 1320
 tacggggtct gacgctcagt ggaacgaaaa ctacagttaa gggatttttg tcatgagatt 1380
 atcaaaaagg atcttcacct agatcctttt aaattaataaa tgaagtttta aatcaatcta 1440
 aagtatatat gtagaacctg atcaggactc ttccttttca tgaacaataa aactgtctgc 1500
 ttacataaac agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc 1560
 taggcccgga ttaaattcca acatggatgc tgatttatat gggtataaat gggtcgcga 1620
 taatgicggg caatcaggtg cgacaatcta tcgattgtat gggaagcccg atgcgccaga 1680
 gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag 1740
 actaaactgg ctgacggaat ttatgcctct tccgaccatc aagcatttta tccgtactcc 1800
 tgatgatgca tggttactca ccaactcgat ccccgggaaa acagattcc aggtattaga 1860
 agaatacct gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt 1920
 gcattcgatt cctgtttgta attgtccttt taacagcgat cgcgtatttc gtctcgtca 1980
 ggcgcaatca cgaatgaata acggtttggt tgatgcgagt gattttgatg acgagcgtaa 2040

tggctggcct gttgaacaag tctggaaaga aatgcataaa cttttgccat tctcaccgga 2100
 ttcagtcgtc aciccatggtg atttctcact tgataacctt atttttgacg aggggaaatt 2160
 aataggttgt attgatgttg gacgagtcgg aatcgacagac cgataccagg atcttgccat 2220
 cctatggaac tgccctgggtg agttttctcc ttcattacag aaacggcttt ttcaaaaata 2280
 tggattgat aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt 2340
 ctaagaattc gcgcaattaa ccctcactaa agggaacaaa agctgggtac cgggcccggt 2400
 gacattgatt attgactagt tattaatagi aatcaattac ggggtcatta gttcatagcc 2460
 catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgcca 2520
 acgacccccg cccattgacg tcaataatga cgtatgttc catagtaacg ccaatagga 2580
 ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 2640
 aagtgtatca taigccaagt acgcccccta ttgacgtcaa tgacggtaaa tggccccct 2700
 ggcattatgc ccagtacatg accttatggg actttctac ttggcagtac atctacgtat 2760
 tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 2820
 ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgttt 2880
 ggcacaaaa tcaacgggac tttcaaaat gtcgtaacaa ctccgccccca ttgacgcaaa 2940
 tggcggtag gcgtgtacgg tgggaggtct atataagcag agctcaagct tgctgcttgc 3000
 acttcagagt tctaggagag tcctcctag tctctcctct ccgaggaggt accgagacct 3060
 caaaataaag gagtattgc cttactgccg agtggagagt gattactgag cggccggtgt 3120
 atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgagct ggcgccaac 3180
 gtggggcccc aggagaagaa aagaaagcgg cctgagaac tcgcttctg aaaaagagga 3240

agaggacaag ttgctatagc aacaagagag aagaagtaga gcaaaggtcc agtggctcgg 3300
 aaaaagagga actgaaactt cggggacgcc tgaaggagta aggtaagtga ctctgctgta 3360
 cgcggggcga ggcagaggtt tccttctaaa ttgaaagaga agtgttgctg cgagaggtct 3420
 tgggtggtcga gaatcctgta caaaaaaag gagggatctc ggtcaggacc aggaccctg 3480
 ggagtaatac aacagcaaca ccgtaagaaa atccgcctag gtgagtctag atagagacta 3540
 ggcgaggcaa gtctccgggg ggaaaagaga ttatcctgag ctcgaaaaat gtatcaagca 3600
 tgcattgcaag ataaaagttc gactcagagg ggagcacttg acagaaggaa attgtttatg 3660
 gtccttaaa acattagatt acatgtttga ggaccataaa gaggaacctt ggacaaaagt 3720
 aaaatttagg acaatatggc agaagtgaa gaatctaact cctgaggaga gtaacaaaaa 3780
 agactttatg tctttgcagg ccacattagc ggttctaata tgttccaaa tggggatgag 3840
 acctgagaca ttgcaagatg caatggctac agtaatcatg aaagatgggt tactggaaca 3900
 agaggaaaag aaggaagaca aaagagaaaa ggaagagagt gtcttcccaa tagtagtgca 3960
 agcagcagga gggagaagct ggaaagcagt agattctgta atgttccagc aactgcaaac 4020
 agtagcaatg cagcatggcc tcgtgtctga ggactttgaa aggcagttgg catattatgc 4080
 tactacttg acaagtaaag acatactaga agtattggcc atgatgcctg caggaattcg 4140
 attctagagg tgatagaat gccagaaaac tatgcaaaaa caagaatcat aaacaggaaa 4200
 aaaagagaac tcagccaca gaggaagaag agaggcgttg gcttggatcat tatgctagtt 4260
 atcatggcaa tagtagctgc cgcaggggct tctctgggag tcgcaaacgc gattcagcag 4320
 tcttacacta aggcagctgt ccagaccctt gctaatgcaa ctgctgcaca gcaggatgtg 4380

ttagaggcaa cctatgccat ggtacagcat gtggctaaag gcgtacgaat cttggaagct 4440
 cgagtggctc gagtggaagc taccacagat agaataatgc tataccaaga attggattgt 4500
 tggcactagg atccatcgcc acc 4523

 <210> 70
 <211> 4819
 <212> DNA
 <213> Artificial

 <220>
 <223> pCAH/SINd3

 <400> 70
 atccagcaca gtggcggcgc ctagcacaaa aataaaaaaa gaaaggggtg actgtgagac 60

 atgggctaaa gaggagcggc cgctcgagtc tagaactagt ggatcagctt tgctgcttgc 120

 acttcagagt tctaggagag tccctcctag tctctctct ccgaggaggt accgagacct 180

 caaaataaag gagtgattgc cttactgccg agtggagagt gattactgag cggccggtgt 240

 atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgagcc gacctcgagg 300

 gggggccta ttctatagtg tcacctaaat gctagagctc gctgatcagc ctcgactgtg 360

 ctttctagtt gccagccatc tgttgtttgc cctcccccg tgccttcctt gaccctggaa 420

 ggtgccactc ccaactgtct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 480

 aggtgtcatt ctattctggg ggggtggggtg gggcaggaca gcaaggggga ggattgggaa 540

 gacaatagca ggcattgctg ggatgcggtg ggctctatgg cttctgagc ggaaagaacc 600

 agtggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 660

 aaaagccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt tttccatag 720

gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 780
 gcacaggacta taaagatacc aggcgtttcc ccttgggaagc tccctcgtgc gctctcctgt 840
 tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 900
 ttctcatagc tcacgctgta ggtatctcag ttcgggtgtag gtcgttcgct ccaagctggg 960
 ctgtgtgcac gaaccccccg ttcagcccca ccgctgcgcc ttatccgta actatcgtct 1020
 tgagtccaac ccgtaagac acgacttatc gccactggca gcagccactg gtaacaggat 1080
 tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtgggtgc ctaactacgg 1140
 ctacactaga aggacagtat ttggtatctg cgtctctctg aagccagtta ccttcggaaa 1200
 aagagtgggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gttttttgt 1260
 ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 1320
 tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt 1380
 atcaaaaagg atcttcacct agatcctttt aaattaataaa tgaagtttta aatcaatcta 1440
 aagtatatat gagtaacctg atcaggactc ttccctttca tgaacaataa aactgictgc 1500
 ttacataaac agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc 1560
 taggcccgca ttaaattcca acatggatgc tgatttatat gggataaat gggctcgca 1620
 taatgctggg caatcaggtg cgacaatcta tcgattgtat ggaagcccg atgcgccaga 1680
 gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag 1740
 actaaactgg ctgacggaat ttatgectct tccgaccatc aagcatttta tccgtactcc 1800
 tgatgatgca tggttactca ccaactcgat ccccgggaaa acagcattcc aggtattaga 1860
 agaatatcct gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt 1920

gcattcgatt cctgtttgta attgtccttt taacagegat cgcgtatttc gtctcgtca 1980

ggcgcaatca cgaatgaata acggtttggt tgatgagagt gattttgatg acgagcgtaa 2040

tggctggcct gtigaacaag tctggaaaga aatgcataaa cttttgcat tctcaccgga 2100

ttcagtcgtc acatcatggtg atttctcact tgataacctt atttttgacg aggggaaatt 2160

aatagttgt attgatgtg gacgagtcgg aatcgagac cgataccagg atcttgccat 2220

cctatggaac tgctcggtg agttttctcc ttattacag aaacggcttt ttcaaaaata 2280

tgtattgat aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt 2340

ctaagaattc gcgcaattaa cctcactaa agggaacaaa agctgggtac cgggccctt 2400

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 2460

catatatgga gttccgctt acataactta cggtaaatgg cccgcctggc tgaccgcca 2520

acgacccccg cccattgacg tcaataatga cgtatgtcc catagtaacg ccaatagga 2580

ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 2640

aagtgtatca tatccaagt acgcccccta ttgacgtcaa tgacggtaaa tggccccct 2700

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 2760

tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 2820

ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt 2880

ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa 2940

tggcggtag gcgtgtacgg tgggaggtct atataagcag agctcaagct tgctgcttgc 3000

acttcagagt tctaggagag tcctcctag tctctctct cagaggaggt accgagacct 3060

caaaataaag gagtattgc cttactgccg agtggagagt gattactgag cggccggtgt 3120

atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgcagct ggcgccaac 3180

gtggggcccg aggagaagaa aagaaagcgg ccctgagaac tggcttctg aaaaagagga 3240

agaggacaag ttgctatagc aacaagagag aagaagtaga gcaaaggtcc agtggctcgg 3300

aaaaagagga actgaaactt cggggacgcc tgaaggagta aggtaagtga ctctgctgta 3360

cgcgggcgca ggcagaggtt tccttctaaa ttgaaagaga agtgttgctg cgagaggtct 3420

tggtggtcga gaatcctgta caaaaaaag gagggatctc ggtcaggacc aggaccctg 3480

ggagtaatac aacagcaaca ccgtaagaaa atccgcctag gtgagtctag atagagacta 3540

ggcggaggca gtctccgggg ggaaaagaga ttatcctgag ctcgaaaaat gtatcaagca 3600

tgcattcaag ataaaagtcc gactcagagg ggagcacttg acagaaggaa atgtttatg 3660

gtgccttaaa acattagatt acatgtttga ggaccataaa gaggaacctt ggacaaaagt 3720

aaaatttagg acaatatggc agaagtgaa gaatctaact cctgaggaga gtaacaaaaa 3780

agactttatg tctttgcagg ccacattagc ggttctaata tgttccaaa tgggatgag 3840

acctgagaca ttgcaagatg caatggctac agtaatcatg aaagatgggt tactggaaca 3900

agaggaaaag aaggaagaca aaagagaaaa ggaagagagt gtcttccaa tagtagtgca 3960

agcagcagga gggagaagct ggaaagcagt agattctgta atgttccagc aactgcaaac 4020

agtagcaatg cagcatggcc tcgtgtctga ggactttgaa aggcagttgg catattatgc 4080

tactacctgg acaagtaaag acatactaga agtattggcc atgatgcctg gaaatagagc 4140

tcaaaaggag ttaattcaag ggaaattaa tgaagaagca gaaaggtgga gaaggaataa 4200

tccaccacct ccagcaggag gaggattaa agtggatcaa attatggggg taggacaaac 4260

aaatcaagca gcagcacaag ctaacatgga tcaggcaagg caaatatgcc tgcaatgggt 4320
 aataaatgca ttaagagcag taagacatat ggcgcacagg ccagggaatc caatgctagt 4380
 aaagcaaaaa acgaatgagc catatgaaga ttttgcagca agactgcagg aattcgattc 4440
 tagaggtgat agaaatgcca gaaaactatg caaaaacaag aatcataaac aggaaaaaaa 4500
 gagaactcag ccacaagagg aagaagagag gcgttggctt ggtcattatg ctagtatca 4560
 tggcaatagt agctgccgca ggggcttctc tgggagtgc aaacgcgatt cagcagtctt 4620
 acactaaggc agctgtccag acccttgcta atgcaactgc tgcacagcag gatgtgttag 4680
 aggcaaccta tgccatggta cagcatgtgg ctaaaggcgt acgaatcttg gaagctcgag 4740
 tggctcgagt ggaagctatc acagatagaa taatgctata ccaagaattg gattgttggc 4800
 actaggatcc atcgccacc 4819

<210> 71
 <211> 5112
 <212> DNA
 <213> Artificial

<220>
 <223> pCAH/SINd4

<400> 71
 atccagcaca gtggcggcgg ctagcacaaa aataaaaaaa gaaagggtgg actgtgagac 60
 atgggctaaa gaggagcggc cgctcgagtc tagaactagt ggatcagctt tgctgcttgc 120
 acttcagagt tctaggagag tccctcctag tctctctct cagaggaggt accgagacct 180
 caaataaag gagtgattgc cttactgccg agtggagagt gattactgag cggccggtgt 240
 atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgagcc gacctcgagg 300

gggggcccta ttctatagtg tcacctaaat gctagagctc gctgatcagc ctcgactgtg 360
 ccttctagtt gccagccatc tgttgtttgc cctcccccg tgccttcctt gaccctggaa 420
 ggtgccactc ccaactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 480
 aggtgtcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggga ggattggaa 540
 gacaatagca ggcattgctg ggatgcggtg ggctctatgg cttctgagc gaaagaacc 600
 agtggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 660
 aaaagccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt tttccatag 720
 gctccgccc cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaacc 780
 gacaggacta taaagatacc aggcgtttcc cctggaagc tccctctgac gctctcctgt 840
 tccgacctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgt 900
 ttctcatage tcacgtgta ggtatctcag ttcggtgtag gtcgttcgt ccaagctggg 960
 ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc ttatccgta actatcgtct 1020
 tgagtccaac ccgtaagac acgacttatc gccactggca gcagccactg gtaacaggat 1080
 tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtggggc ctaactacgg 1140
 ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagta ccttcgaaa 1200
 aagagtggg agctcttgat ccggcaaaca aaccaccgt ggtagcggg gttttttgt 1260
 ttgcaagcag cagattacgc gcagaaaaa aggatctcaa gaagatcctt tgatctttc 1320
 tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggatcttg tcatgagatt 1380
 atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 1440

aagtatatat gagtaacctg atcaggactc ttccttttca tgaacaataa aactgtctgc 1500
ttacataaac agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc 1560
taggccgcga ttaaattcca acatggatgc tgatttatat gggataaat gggctcgcga 1620
taatgtcggg caatcaggtg cgacaatcia tcgattgtat ggaagcccg atgcgccaga 1680
gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag 1740
actaaactgg ctgacggaat ttatgcctct tccgaccatc aagcatttta tccgtactcc 1800
tgatgatgca tggttactca ccactgcgat ccccgggaaa acagcattcc aggtattaga 1860
agaatatacct gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt 1920
gcattcgatt cctgtttgta attgtccttt taacagcgat cgcgtatttc gtctcgtca 1980
ggcgcaatca cgaatgaata acggtttggt tgatgcgagt gattttgatg acgagcgtaa 2040
tggtcggcct gtgaacaag tctggaaaga aatgcataaa cttttgccat tctcacgga 2100
ttcagtcgtc acicatggtg atttctcact tgataacctt atttttgacg aggggaaatt 2160
aatagttgt attgatgttg gacgagtcgg aatcgagac cgataccagg atcttgccat 2220
cctatggaac tgctcgggtg agttttctcc ttcattacag aaacggcttt tcaaaaata 2280
tggattgat aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt 2340
ctaagaattc gcgcaattaa ccctactaa agggaacaaa agctgggtac cgggcccggtt 2400
gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 2460
catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgcca 2520
acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagga 2580
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 2640

aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggccccct 2700

ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 2760

tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 2820

ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgttt 2880

ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa 2940

tggcggttag gcgtgtacgg tgggaggtct atataagcag agctcaagct tgctgcttgc 3000

acttcagagt tctaggagag tcctccttag tctctcctct ccgaggaggt accgagacct 3060

caaaataaag gagtgattgc cttactgccg agtggagagt gattactgag cggccggtgt 3120

atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgagct ggcgccaac 3180

gtggggcccg aggagaagaa aagaaagcgg cctgagaac tcggcttctg aaaaagagga 3240

agaggacaag ttgctatagc aacaagagag aagaagtaga gcaaaggtcc agtggctcgg 3300

aaaaagagga actgaaactt cggggacgcc tgaaggagta aggttaagtga ctctgctgta 3360

cgcgggcgga ggcagaggtt tccttctaaa ttgaaagaga agtgttgctg cgagaggtct 3420

tggtgtcga gaatcctgta caaaaaaag gagggatctc ggtcaggacc aggaccctg 3480

ggagtaatac aacagcaaca ccgtaagaaa atccgcctag gtgagtctag atagagacta 3540

ggcgaggcaa gtctccgggg ggaaaagaga ttatcctgag ctgaaaaat gtatcaagca 3600

tgcatgcaag ataaaagttc gactcagagg ggagcacttg acagaaggaa attgtttatg 3660

gtgccttaaa acattagatt acatgtttga ggaccataaa gaggaacctt ggacaaaagt 3720

aaaatttagg acaatatggc agaagtgaa gaatctaact cctgaggaga gtaacaaaaa 3780

agactttatg tctttgcagg ccacattagc ggttctaag tgttgccaaa tgggatgag 3840
 acctgagaca ttgcaagatg caatggctac agtaatcatg aaagatgggt tactggaaca 3900
 agaggaaaag aaggaagaca aaagagaaaa ggaagagagt gtcttcccaa tagtagtgca 3960
 agcagcagga gggagaagct ggaaagcagt agattctgta atgttccagc aactgcaaac 4020
 agtagcaatg cagcatggcc tcgtgtctga ggactttgaa aggcagttgg catattatgc 4080
 tactacctgg acaagtaaag acatactaga agtattggcc atgatgcctg gaaatagagc 4140
 tcaaaaggag ttaattcaag ggaaattaa tgaagaagca gaaaggtgga gaaggaataa 4200
 tccaccacct ccagcaggag gaggattaac agtggatcaa attatggggg taggacaaac 4260
 aatcaagca gcagcacaag ctaacatgga tcaggcaagg caaatatgcc tgcaatgggt 4320
 aataaatgca ttaagagcag taagacatat ggcgcacagg ccagggaatc caatgctagt 4380
 aaagcaaaaa acgaatgagc catatgaaga ttttgcagca agactgctag aagcaataga 4440
 tgcagagcca gttacacagc ctataaaaga ttatctaaag ctaacactat cttatacaaa 4500
 tgcacagca gattgtcaga agcaaatgga tagaacacta ggacaaagag tacaacaagc 4560
 tagttagaa gaaaaaatgc aagcatgtag agatgtggga tcagaagggt tcaaaatgca 4620
 attgttagca caagcattaa ggccaggaaa aggaaaaggg aatggacagc cacaaagggtg 4680
 ttacaactgt ggaaaaccgg gacatcaagc aaggcactgc aggaattcga ttctagaggt 4740
 gatagaatg ccagaaaact atgcaaaaac aagaatcata aacaggaaaa aaagagaact 4800
 cagccacaag aggaagaaga gaggcgttgg cttggtcatt atgctagtta tcatggcaat 4860
 agtagctgcc gcaggggctt ctctgggagt cgcaaacgagc attcagcagt cttactactaa 4920
 ggcagctgtc cagaccttg ctaatgcaac tgetgcacag caggatgtgt tagaggcaac 4980

ctatgccatg gtacagcatg tggctaaagg cgtacgaatc ttggaagctc gaggctctcg 5040

agtgggaagct atcacagata gaataatgct ataccaagaa ttggattgtt ggcactagga 5100

tccatcgcca cc 5112

<210> 72
 <211> 7579
 <212> DNA
 <213> Artificial

<220>
 <223> pMYKEF1/env

<220>
 <221> misc_feature
 <222> (994)..(994)
 <223> n is a, c, g, or t

<400> 72
 aacaggaaag ttccattgga gccaaagtaca ttgagtcaat agggactttc caatgggttt 60

tgcccagtac ataaggtcaa tgggaggtaa gccaatgggt ttttccatt actggcacgt 120

atactgagtc attagggact ttccaatggg ttttgcccag tacataaggt caataggggt 180

gaatcaacag gaaagtccca ttggagccaa gtacactgag tcaatagga ctttcattg 240

ggttttgccc agtaciaaaag gtcaataggg ggtgagtcaa tgggtttttc ccattattgg 300

cacgtacata aggtcaatag gggtgagtca ttgggttttt ccagccaatt taattaaac 360

gccatgtact ttccacat tgacgtcaat ggctattga aactaatgca acgtgacctt 420

taaacggtac tttccatag ctgattaatg ggaaagtacc gttctcgagc caatacacgt 480

caatgggaag tgaagggca gccaaaacgt aacaccgcc cggttttccc tggaaattcc 540

atattggcac gcattctatt ggctgagctg cgttcacgtg ggtataagag gcgcgaccag 600
 cgtcggtaacc gtcgcagctc tcggtctgac caccgtagaa cgcagagctc ctgctgcag 660
 gcatgcaagc ttgtaagt ccgtgtgtgg ttcccgcggg cctggcctct ttacgggta 720
 tggcccttgc gtgccttgaa ttacttccac gccctggct gcagtacgtg attcttgatc 780
 ccgagcttcg ggttgaagt ggggtgggaga gttcaaggcc ttgcgcttaa ggagccctt 840
 gccttttgc ttgagttgag gcctggcctg ggcgctgggg ccgccgctg caaatctggt 900
 ggcaccttcg cgctgtctc gctgcttctg ataagtctct agccatttaa aattttgat 960
 gacctgctgc gacgcttttt ttctggcaag atantcttgt aatgcgggc caagatctgc 1020
 aacctggtat ttcggttttt ggggccgcgg gcggctacgg ggcccgtgcg tcccagcgca 1080
 catgttcggc gaggaggggc ctgcgagcgc ggccaccgag aatcggacgg gggtagtctc 1140
 aagctggccg gectgctctg gtgectggcc tcgcccgcgc gtglatgcc ccgcectggg 1200
 cggaaggct ggcccgtcgc gcaccagttg cgtgagcgga aagatggccg cttcccggcc 1260
 ctgctgcagg gagctcaaaa tggaggacgc ggcgctcggg agagcgggcg ggtgagtcac 1320
 ccacacaaag gaaaagggcc tttccgtcct cagccgtcgc ttcatgtgac tccacggagt 1380
 accggcgccc gtccaggcac ctcgattagt tctcgagctt ttggagtac tcgtctttag 1440
 gttgggggga ggggttttat gcgatggagt ttcccacac tgagtgggtg gagactgaag 1500
 ttaggccagc ttggcacttg atgtaattct ccttgggaatt tgcccttttt gaggttggat 1560
 cttggttcat tctcaagcct cagacagtgg ttcaaagttt ttttcttcca tttcagggat 1620
 ccactagtaa cggcccagc tgtgctggaa ttgatcata cctggtgttg ctgactacc 1680
 cgaccgggt aaaagtcgat ggtattgctg cctgggtcca tgcttctcac ctcaaactg 1740

caccaccttc ggcaccagat gagtctctggg agctggaaaa gactgatcat cctcttaagc 1800

tgcgtattcg gcggcggcgg gacgagctcg caaaataaga acccccacca gcccatgacc 1860

ctcacttggc aggtactgtc ccaaactgga gacgttgtct gggatacaaa ggcagtcag 1920

cccccttggc ctigtgtggc cacacttaaa cctgatgtat gtgccttggc ggctagtctt 1980

gagtcctggg atatcccggg aaccgatgtc tcgtcctcta aacgagtcag acctccggac 2040

tcagactata ctgccgtta taagcaaac acctggggag ccatagggtg cagctacct 2100

cgggctagga ctagaatggc aagctctacc ttctacgtat gtccccgga tggccggacc 2160

ctttcagaag ctagaagtg cggggggcta gaatccctat actgtaaaga atgggattgt 2220

gagaccacgg ggaccggtta ttggctatct aaatcctcaa aagacctcat aactgtaaaa 2280

tgggacaaa atagcgaatg gactcaaaaa ttcaacagt gtcaccagac cggctggtgt 2340

aaccccccta aaatagattt cacagacaaa gaaaaattat ccaaggactg gataacggga 2400

aaaacctggg gattaagatt ctatgtgtct ggacatccag gcgtacagtt caccattcgc 2460

ttaaaaatca ccaacatgcc agctgtggca gtaggtcctg acctcgtcct tgtggaacaa 2520

ggacctccta gaacgtcctt cgctctccca cctcctcttc cccaaggga agcgcaccg 2580

ccatctctcc ccgactctaa ctccacagcc ctggcgacta gtcacaaac tcccacggtg 2640

agaaaaacaa ttgttacctt aaacactccg cctcccacca caggcgacag actttttgat 2700

cttgtgcagg gggccttctt aaccttaaat gctaccaacc caggggccac tgagtcttgc 2760

tggctttgtt tggccatggg ccccccttat tatgaagcaa tagcctcatic aggagaggtc 2820

gcctactcca ccgaccttga ccggtgccgc tgggggaccc aaggaaagct caccctcact 2880

gaggtctcag gacacgggtt gtgcatagga aaggtgcctt ttacccatca gcatctctgc 2940
 aatcagacce taiccatcaa ttcttccgga gaccatcagt atctgctccc ctccaacat 3000
 agctggtggg ctigcagcac tggcctcacc ccttgcctct ccacctcagt ttttaatcag 3060
 actagagatt tctgtatcca ggtccagctg attcctcgca tctattacta tctgaagaa 3120
 gttttgttac aggccatga caattctcac cccaggacta aaagagagc tgtctcactt 3180
 accctagctg ttttactggg gttgggaatc acggcgggaa taggtactgg ttcaactgcc 3240
 ttaattaaag gacctataga cctccagcaa ggcttgacaa gcctccagat cgccatagat 3300
 gctgacctcc gggccctcca agactcagtc agcaagttag aggactcact gacttcctg 3360
 tccgaggtag tgctccaaaa taggagagc cttgacttgc tgtttctaaa agaaggtggc 3420
 ctctgtgagg ccctaaagga agagtgtgt ttttacctag accactcagg tgcagtacgg 3480
 gactccatga aaaaactcaa agaaaaactg gataaaagac agttagagcg ccagaaaagc 3540
 caaaactggt atgaaggatg gttcaataac tccccttggg tcactacct gctatcaacc 3600
 atcgctgggc ccctattact cctccttctg ttgctcatcc tcgggccatg catcatcaat 3660
 aagttagttc aattcatcaa tgataggata agtgcagtta aaattctggt ccttagacaa 3720
 aaatatcagg ccctagagaa cgaaggtaac ctttaatttt gctctaagat tagagctatt 3780
 cacaagagaa atgggatca ctagtgaatt ctgcagatat ccatcacact ggcggccgct 3840
 cgagcatgca tctagaggc cctattctat agtgtcacct aaatgctaga gctcgctgat 3900
 cagcctcgac tgtgcttct agttgccagc catctgttgt ttgccctcc cccgtgctt 3960
 ccttgacct ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat 4020
 cgattgtct gagtaggtgt cattctattc tgggggtgg ggtggggcag gacagcaagg 4080

gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggcttctg 4140

aggcggaaag aaccagtggc ggtaatacgg ttatccacag aatcagggga taacgcagga 4200

aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 4260

gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag 4320

aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg aagctcctc 4380

gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgctt tctcccttcg 4440

ggaagcgtgg cgctttctca tagctcacgc ttaggtatc tcagttcggg ttaggtcgtt 4500

cgctccaagc tgggctgtgt gcacgaacc cccgttcagc ccgaccgctg cgccttatec 4560

ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc 4620

actggtaaac ggattagcag agcgaggatg ttaggcgggtg ctacagagtt cttgaagtgg 4680

tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 4740

gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 4800

ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 4860

cctttgatct ttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 4920

ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaatta aaaatgaagt 4980

tttaaatcaa tctaaagtat atatgagtaa cctgaggcta tggcagggcc tgccgccccg 5040

acgttggctg cgagccctgg gccttcacc gaacttgggg ggtggggtgg ggaaaaggaa 5100

gaaacgcggg cgtattggcc ccaatggggt ctcggtgggg tatcgacaga gtgccagccc 5160

tgggaccgaa ccccgcgttt atgaacaaac gaccaaacac cgtgcgtttt attctgtctt 5220

ttattgccg tcatagcgg ggttccttcc ggtattgtct cttccgtgt ttcagttagc 5280
 ctccccctag ggtggcgaa gaactccagc atgagatccc cgcgctggag gatcatccag 5340
 ccggcgtccc ggaaaacgat tccgaagccc aacctttcat agaaggcggc ggtggaatcg 5400
 aaatctcgtg atggcaggtt gggcgtcgtc tggtcggca tttcgaacc cagagtcccg 5460
 ctcaagaaga ctctcaaga aggcgataga aggcgatgcg ctgcgaatcg ggagcggcga 5520
 taccgtaaag cacgaggaag cggtcagccc attcggccc aagctcttca gcaatatcac 5580
 ggtagccaa cgctatgtcc tgatagcgtt cgccacacc cagccggcca cagtcgatga 5640
 atccagaaaa gcggccattt tccacatga tattcggcaa gcaggcatcg ccatgggtca 5700
 cgacgagatc ctgcccgtcg ggcgatctcg ccttgaccct ggcaaacagt tcggctggcg 5760
 cgagcccctg atgctcttga tcatctgat cgacaagacc ggcttccatc cgagtacgtg 5820
 ctgctcgtg gcatgtttc gcttgggtg cgaatgggca ggtagccgga tcaagcgtat 5880
 gcagccgccc cattgcatca gccatgatgg atactttctc ggcaggagca aggtgagatg 5940
 acaggagatc ctgccccggc acttcgcca atagcagcca gtccttccc gcttcagtga 6000
 caacgtcgag cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agcccgctg 6060
 cctcgtcttg cagttcattc agggcaccgg acaggtcggc cttgacaaaa agaaccgggc 6120
 gccctgccc tgacagccgg aacacggcgg catcagagca gccgattgtc tgttgtccc 6180
 agtcatagcc gaatagcctc tccaccaag cggccggaga acctgctgc aatccatctt 6240
 gttcaatcat gcgaaacgat cctcatctg tctcttgatc gatctttgca aaagcctagg 6300
 cctcaaaaa agcctctca ctacttctgg aatagctcag aggccgagc gcctcggcc 6360
 tctgcataaa taaaaaaaa tagtcagcca tggggcggag aatggcgga actggcgga 6420

gttagggcg ggatgggagg agttaggggc gggactatgg ttgctgacta attgagatgc 6480
 atgctttgca tacttctgcc tgctggggag cctggggact ttccacacct ggttgetgac 6540
 taattgagat gcatgctttg catacttctg cctgctgggg agcctgggga ctttccacac 6600
 cctaactgac acacattcca cagctgggtc tttccgctc aggactcttc ctttttcaat 6660
 aatcaatct aaagtatata tgagtaaact tggctgaca gttaccaatg cttaatcagt 6720
 gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgctg actccccgc 6780
 gtgtagataa ctacgatagc ggagggctta ccatctggcc ccagtgtgc aatgataccg 6840
 cgagaccac gctcaccgc tccagattta tcagcaataa accagccagc cgaaggggc 6900
 gagcgagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttggtgccg 6960
 gaagctagag taagtagtcc gccagttaat agtttgcca acgttggtgc cattgctaca 7020
 ggcacgtgg tgcacgctc gtcgtttggt atggcttcat tcagctccgg ttccaacga 7080
 tcaaggcgag ttacatgac ccccatgttg tgcaaaaaag cggttagctc cttcggtcct 7140
 cggatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg 7200
 cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca 7260
 accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata 7320
 cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 7380
 tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaaccact 7440
 cgtgcacca actgatcttc agcatctttt actttacca gcgtttctgg gtgagcaaaa 7500
 acaggaagge aaaatgccgc aaaaaaggga ataaggcgca cacggaaatg ttgaatactc 7560

atactcttcc tttttcaat 7579

<210> 73
 <211> 3566
 <212> DNA
 <213> Artificial

<220>
 <223> pCAH/SINd

<400> 73
 atccagcaca gtggcggccg ctagcacaaa aataaaaaaa gaaagggtgg actgtgagac 60
 atgggctaaa gaggagcggc cgctcgagtc tagaactagt ggatcagctt tgctgcttgc 120
 acttcagagt tctaggagag tcctcctag tctctcctct ccgaggaggt accgagacct 180
 caaaataaag gagtgattgc cttactgccg agtggagagt gattactgag cggccggtgt 240
 atcgggagtc gtccttaat ctgtgcaata ccagagcggc tctcgcagcc gacctcgagg 300
 gggggccta ttctatagtg tcacctaaat gctagagctc gctgatcagc ctcgactgtg 360
 ctttctagtt gccagccatc tgttgtttgc cctcccccg tgccttcctt gaccctggaa 420
 ggtgccactc cactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 480
 aggtgtcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggga ggattgggaa 540
 gacaatagca ggcattctgg ggatgcggtg ggctctatgg cttctgagc ggaaagaacc 600
 agtggcggtg atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 660
 aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 720
 gctccgccc cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaacc 780
 gacaggacta taaagatacc aggcgtttcc cctggaagc tcctcgtgc gctctcctgt 840

tccgaccctg ccgcttaccg gataacctgtc cgcctttctc ccttcgggaa gcgtggcgct 900

ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 960

ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct 1020

tgagtccaac ccggtaaagc acgacttatc gccactggca gcagccactg gtaacaggat 1080

tagcagagcg aggtatgtag gcggtgctac agagtctctg aagtgggtgc ctaactacgg 1140

ctacactaga aggacagtat ttggtatctg cgtctctctg aagccagtta ccttcggaaa 1200

aagagtgggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gttttttgt 1260

ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 1320

tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt 1380

atcaaaaagg atcttcacct agatcctttt aaattaataaa tgaagtttta aatcaatcta 1440

aagtatatat gagnaacctg atcaggactc ttccttttca tgaacaataa aactgtctgc 1500

ttacataaac agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc 1560

taggcccgca ttaaattcca acatggatgc tgatttatat gggtataaat gggtcgcga 1620

taatgtcggg caatcagggt gcacaatcta tcgattgtat gggaagcccg atgcgccaga 1680

gttgtttctg aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag 1740

actaaactgg ctgacggaat ttatgcctct tccgaccatc aagcatttta tccgtactcc 1800

tgatgatgca tggttactca ccaactgcgat ccccgggaaa acagcattcc aggtattaga 1860

agaatatacct gattcagggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt 1920

gcattcgatt cctgtttgta attgtccttt taacagcgat cgcgtatttc gtctcgtca 1980

ggcgcaatca cgaatgaata acggtttggt tgatgcgagt gattttgatg acgagcgtaa 2040

tggctggcct gttgaacaag tctgaaaga aatgcataaa cttttgcat tctcaccgga 2100
 ttcagtcgtc aciccatggtg atttctcact tgataacctt atttttgacg aggggaaatt 2160
 aataggttgt attgatgttg gacgagtcgg aatcgacagac cgataccagg atcttgccat 2220
 cctatggaac tgccctcgtg agttttctcc ttattacag aaacggcttt ttcaaaaata 2280
 tggattgat aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt 2340
 ctaagaattc gcgcaattaa ccctcactaa agggaacaaa agctgggtac cgggcccggtt 2400
 gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc 2460
 catatatgga gttcccggtt acataactta cggtaaatgg cccgcctggc tgaccgcca 2520
 acgacccccg cccattgacg tcaataatga cgtatgttc catagtaacg ccaatagga 2580
 ctttcattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 2640
 aagtgtatca taigccaagt acgccccctta ttgacgtcaa tgacggtaaa tggccccct 2700
 ggcattatgc ccagtacatg accttatggg actttctac ttggcagtac atctacgtat 2760
 tagtcatcgc tattaccatg gtgatcggtt ttggcagta catcaatggg cgtggatagc 2820
 ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgttt 2880
 ggcacaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa 2940
 tggcggtag gcgtgtacgg tgggaggtct atataagcag agctcaagct ttgctgcttg 3000
 cacttcagag ttctaggaga gtcctccta gtctctctc tccgaggagg taccgagacc 3060
 tcaaaataaa ggagtgattg ccttactgcc gagtggagag tgattactga gggccggtg 3120
 tatcgggagt cgtcccttaa tctgtgcaat accagagcgg ctctcgcagc tggcggaat 3180

tcgattctag aggtgataga aatgccagaa aactatgcaa aaacaagaat cataaacagg 3240

aaaaaaagag aactcagcca caagaggaag aagagaggcg ttggcttggc cattatgcta 3300

gttatcatgg caatagtagc tgccgcaggg gcttctctgg gagtgcgaaa cgcgattcag 3360

cagtcttaca ctaaggcagc tgtccagacc cttgctaata caactgctgc acagcaggat 3420

ggttagagg caacctatgc catggtacag catgtggcta aaggcgtacg aatcttggaa 3480

gctcgagtgg ctcgagtgga agctatcaca gatagaataa tgctatacca agaattggat 3540

tgttggcact aggatccatc gccacc 3566

<210> 74
 <211> 7623
 <212> DNA
 <213> Artificial

<220>
 <223> pHGVSV-G

<400> 74
 gccgcgcttg acattgatta ttgactagtt attaatagta atcaattacg gggtcattag 60

ttcatagccc atatatggag ttccgcgcta cataacttac ggtaaatggc cgccttggct 120

gaccgcccaa cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc 180

caatagggac ttccattga cgtcaatggg tggactattt acggtaaact gcccacttgg 240

cagtacatca agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat 300

ggccccctg gcattatgcc cagtacatga ctttatggga ctttctact tggcagtaca 360

tctacgtatt agtcatcgtc attaccatgg tgatgcgggt ttggcagtac atcaatgggc 420

gtggatagcg gtttgactca cggggatttc caagtcctcca cccattgac gtcaatggga 480

gtttgttttg gcaccaaagt caacgggact ttccaaaatg tcgtaacaac tccgcccatt 540

tgacgcaaat gggcggtagg cgtgtacggt gggaggctta tataagcaga gctctctggc 600

taactagaga acccactgct tactggctta tcgaaattaa tacgactcac tatagggaga 660

ccaagcttg gtaccgagct cggatccact agtaacggcc gccagtgtgc tggaaatcga 720

tgatcctgag aacttcaggg tgagtctaig ggacccttga tgtttcttt cccttcttt 780

tctatggtta agttcatgct ataggaaggg gagaagtaac aggttacagt ttagaatggg 840

aaacagacga atgattgcat cagtgtggaa gtctcaggat cgttttagtt tcttttattt 900

gctgttcata acaattgttt tcttttgttt aattcttctt ttcttttttt ttcttctccg 960

caatttttac tattatactt aatgccttaa cattgtgtat aacaaaagga aatatctctg 1020

agatacatta agtaacttaa aaaaaaactt tacacagtct gcctagtaca ttactatttg 1080

gaatatatgt gtgcttattt gcatattcat aatctcccta ctttattttc ttttattttt 1140

aattgataca taatcattat acatatttat gggttaaagt gtaatgtttt aatatgtgta 1200

cacatattga ccaaatcagg gtaattttgc atttgaatt ttaaaaaatg ctttcttctt 1260

ttaataact tttttgttta tcttatttct aatactttcc ctaatctctt tctttcaggg 1320

caataatgat acaatgtatc atgcctcttt gcaccattct aaagaataac agtgataatt 1380

tctgggttaa ggcaatagca atatttctgc atataaatat ttctgcatat aaattgtaac 1440

tgatgtaaga ggtttcatat tgctaatagc agctacaatc cagctaccat taggcccttt 1500

tgctaatcat gttcatacct cttatcttcc tcccacagct cctgggcaac gtgctgtctt 1560

ggtgctggc ccataccttt ggcaaagaat cactagttaa ttctgcagat atccatcaca 1620

ctggcggcgg ctcgaggaat tctgacacta tgaagtgcct tttgtactta gcctttttat 1680

tcattggggt gaattgcaag ttcaccatag tttttccaca caaccaaaaa ggaaactgga 1740

aaaatgttcc ttctaattac cattattgcc cgccaagctc agatttaaat tggcataatg 1800

acttaatagg cacagcctta caagtcaaaa tgcccaagag tcacaaggct attcaagcag 1860

acggttggat gtgtcatgct tccaaatggg tctactcttg tgatttccgc tggatggac 1920

cgaagtatat aacacattcc atccgatcct tctctccatc tgtagaacia tgcaaggaaa 1980

gcattgaaca aacgaaacia ggaacttggc tgaatccagg cttcctcct caaagttgtg 2040

gatatgcaac tgtgacggat gccgaagcag tgattgtcca ggtgactcct caccatgtgc 2100

tggttgatga atacacagga gaatgggttg attcacagtt catcaacgga aatgcagca 2160

attacatatg cccactgtc cataactcta caacctggca ttctgactat aaggtcaaag 2220

ggctatgtga ttctaacctc atttccatgg acatcacctt cttctcagag gacggagagc 2280

tatcatccct gggaaaggag ggcacagggt tcagaagtaa ctactttgct tatgaaactg 2340

gaggcaaggc ctgcaaaatg caaactgca agcattgggg agtcagactc ccatcaggtg 2400

tctggttcca gatggctgat aaggatctct ttgctgcagc cagattccct gaatgcccag 2460

aagggtcaag tatctctgct ccatctcaga cctcagtgga tgtaagtcta attcaggacg 2520

ttgagaggat cttggattat tccctctgcc aagaaactg gagcaaaatc agagcgggtc 2580

ttccaatctc tccagtggat ctcagctatc ttgctcctaa aaaccagga accggtcctg 2640

ctttcaccat aatcaatgg accctaaaat actttgagac cagatacatc agagtcgata 2700

ttgctgctcc aatcctctca agaatggctg gaatgatcag tggaactacc acagaaaggg 2760

aactgtggga tgaactggca ccatatgaag acgtggaaat tggaccaat ggagttctga 2820

ggaccagttc aggatataag tttcctttat acatgattgg acatggtatg ttggactcgg 2880
 atcttcatct tagctcaaag gctcaggtgt tcgaacatcc tcacattcaa gacgctgctt 2940
 cgcaacttcc tgatgatgag agtttatttt ttggtgatac tgggctatcc aaaaatccaa 3000
 tcgagcttgt agaaggttgg ttcagtagtt ggaaaagctc tattgcctct tttttcttta 3060
 tcatagggtt aatcattgga ctattcttgg ttctccgagt tggtatccat ctttgatta 3120
 aattaaagca caccaagaaa agacagattt atacagacat agagatgaac cgacttggaa 3180
 agtaactcaa atcctgcaca acagattctt catgtttgga ccaaatcaac ttgtgatacc 3240
 atgctcaaag aggctcaat tatatttgag tttttaattt ttatgaaaa aaaaaaaaaa 3300
 aacggaattc ctcgagcatg catctagagg gccctattct atagtgtcac ctaaatgcta 3360
 gagctcgctg atcagcctcg actgtgcctt ctagttgcca gccatctgtt gtttgcacct 3420
 cccccgtgcc ttcttgacc ctggaagggtg ccactccac tgctctttcc taataaatg 3480
 aggaaattgc atcgattgt ctgagtaggt gtcattctat tctgggggtt ggggtggggc 3540
 aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggat gcggtgggct 3600
 ctatggcttc tgaggcggaa agaaccagtg gcggtataac ggttatccac agaatcaggg 3660
 gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 3720
 gccgcgttgc tggcgttttt ccataggtc cgccccctg acgagcatca caaaaatcga 3780
 cgctcaagtc agaggtggcg aaaccgaca ggactataaa gataccagge gtttcccct 3840
 ggaagctccc tcgtgcctc tctgttccg acctgccc ttaccggata cctgtccgc 3900
 ttttccctt cggaagcgt ggcctttct catagctcac gctgtaggta tctcagttcg 3960
 gtgtaggtcg ttcgctcaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc 4020

tgcgccttat cegtaacta tcgtcttgag tccaaccgg taagacacga cttatcgcca 4080
 ctggcagcag ccaactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 4140
 ttcttgaagt ggiggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct 4200
 ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaacc 4260
 accgctggta gcggtggttt ttttgttgc aagcagcaga ttacgcgcag aaaaaagga 4320
 tctcaagaag atcctttgat cttttctac gggctgacg ctcagtggaa cgaaaactca 4380
 cgtaagga ttttggcat gagattatca aaaagatct tcacctagat ctttttaat 4440
 taaaaatgaa gttttaaate aatctaaagt atatatgagt aacctgagc tatggcaggg 4500
 cctgcgcccc cgactttggc tgcgagccct gggccttcac ccgaacttgg ggggtggggt 4560
 ggggaaaagg aaaaaacgcg ggcgtattgg cccaatggg gtctcggtgg ggtatcgaca 4620
 gagtgccagc cctgggaccg aaccccgct ttatgaaca acgaccaac accgtgcgtt 4680
 ttattctgtc tttttattgc cgtcatagcg cgggttcctt ccggtattgt ctcttccgt 4740
 gtttcagtta gcctcccct aggggtggcg aagaactcca gcatgagatc cccgcgctgg 4800
 aggatcatcc agccggcgtc ccgaaaacg attccgaagc ccaaccttc atagaaggcg 4860
 gcggtggaat cgaatctcg tgatggcagg ttggcgtcg cttggtcggc catttcgaac 4920
 ccagagtcc cgtcagaag aactcgtcaa gaaggcgata gaaggcgatg cgtgcgaat 4980
 cgggagcggc gataccgtaa agcacgagga agcgtcagc ccattcgccg ccaagctctt 5040
 cagcaatatc acggtagcc aacgtatgt cctgatagcg gtccgccaca ccagccggc 5100
 cacagtcgat gaatccagaa aagcggccat tttccaccat gatattcggc aagcaggcat 5160

cgccatgggt cacgacgaga tcctcgccgt cgggcatgct cgccttgagc ctggcgaaca 5220

gttcggctgg cgcgagcccc tgatgctctt gatcatcctg atcgacaaga cgggttcca 5280

tccgagtacg tgctcgctcg atgcgatggt tcgcttggg gtcgaatggg caggtagccg 5340

gatcaagcgt atgcagccgc cgcattgcat cagccatgat ggatactttc tcggcaggag 5400

caagtgaga tgacaggaga tcctgccccg gcacttcgcc caatagcagc cagtcccttc 5460

ccgcttcagt gacaacgtcg agcacagctg cgcaaggaac gcccgtcgtg gccagccacg 5520

atagccgcgc tgctcgtct tgcagttcat tcagggcacc ggacaggtcg gtcttgacaa 5580

aaagaaccgg gcgccctgc gctgacagcc ggaacacggc ggcatcagag cagccgattg 5640

tctgttgctc ccagtcatag ccgaatagcc tctccacca agcggccgga gaacctgct 5700

gcaatccatc ttgttcaatc atgcgaaacg atcctcatcc tgtctcttga tcgatctttg 5760

caaaagccta ggctccaaa aaagcctect cactacttct ggaatagctc agaggccgag 5820

gcggcctcgg ccctgcata aataaaaaaa attagtcagc catggggcgg agaattggcg 5880

gaactggcg gagttaggg cgggatgggc ggagttaggg gcgggactat ggttgctgac 5940

taattgagat gcatgctttg catacttctg cctgctgggg agcctgggga cttccacac 6000

ctggttgctg actaattgag atgcatgctt tgcatacttc tgcctgctgg ggagcctggg 6060

gactttccac accctaactg acacacattc cacagctggt tctttccgcc tcaggactct 6120

tcctttttca ataatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa 6180

tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac catagttgcc 6240

tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg cccagtgct 6300

gcaatgatac cgcgagacc acgtcaccg gctccagatt tatcagcaat aaaccagcea 6360

gccggaaggg ccgagcgcag aagtggtect gcaactttat ccgcctccat ccagctctatt 6420
 aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttggt 6480
 gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc 6540
 ggttccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa agcggttagc 6600
 tccttcggtc ctccgatcgt tgtcagaagt aagtggccg cagtgttacc actcatggtt 6660
 atggcagcac tgcataattc tcttactgtc atgcatccg taagatgctt ttctgtgact 6720
 ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc 6780
 ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt 6840
 ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg 6900
 atgtaacca ctctgcacc caactgatct tcagcatctt ttaactttcac cagcgtttct 6960
 gggtagcaaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa 7020
 tgttgaatac tcatactctt cttttttcaa tattattgaa gcatttatca gggttattgt 7080
 ctcatgagcg gatcacatatt tgaatgtatt tagaaaaata aacaatatagg ggttccgcgc 7140
 acatttccc gaaaagtgcc acctgacgcg ccctgtagcg gcgattaag cgcggcgggt 7200
 gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctccttc 7260
 gctttcttc cttcctttct cgccacgttc gccggctttc cccgtcaage tctaaatcg 7320
 gggctccctt tagggttccg atttagtgtt ttacggcacc tcgaccccaa aaaacttgat 7380
 tagggtgatg gttcacgtag tgggcatcg ccctgataga cggtttttcg cctttgacg 7440
 ttggagtcca cgttctttaa tagtggactc ttgttccaaa ctggaacaac actcaacct 7500

atctcggctct attcttttga tttaataaggg attttgccga tttcggccta ttggttaaaa 7560

aatgagctga tttacaacaaa atttaacgcg aattttaaca aaatattaac gcttacaatt 7620

tac 7623

<210> 75
 <211> 5419
 <212> DNA
 <213> Artificial

<220>
 <223> pHYK/rev

<400> 75
 gcgcgcgttg acattgatta ttgactagt attaatagta atcaattacg gggtcattag 60

ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaattggc ccgcctggct 120

gaccgcccac cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc 180

caatagggac tticcattga cgtcaatggg tggactatit acggtaaact gcccaacttg 240

cagtacatca agtgtatcat atgccaagia cgccccctat tgacgtcaat gacggtaaat 300

ggccccctg gcattatgcc cagtacatga cttatggga ctttctact tggcagtaca 360

tctacgtatt agtcatcgt attaccatgg tgatgcggtt ttggcagtac atcaatgggc 420

giggatagcg gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga 480

gtttgtttg gcacaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat 540

tgacgcaaat gggcggtagg cgtgtacggt gggaggctca tataagcaga gctctctggc 600

taactagaga acccaactgct tactggctta tcgaaattaa tacgactcac tatagggaga 660

cccaagctta gatggatgct ggggccagat acatgcgctt aactgggaag gaaaactggg 720

ttgaagtaac catggacgga gagaaggaaa ggaaaagaga aggtttcact gcgggacagc 780
 aagatataca gaactctaag taccgccgaca taccaacggg tcacagtcac catggaaaca 840
 agagcagacg tcgcaggaga aaatcaggat ttggcgatg gcttagagga atcagacaac 900
 agcgaacaa gcgaaagagt gacagtacag aaagcttggg gccgtgcctg ggagctttgg 960
 cagaactcac cciggaagga gccatggaaa aggggcctgc tgaggctgct cgtccttccg 1020
 ctgacgatgg gaatctggat aaatggatgg ctgggagaac accacaaaa taagaattct 1080
 gcagatatcc atcacactgg cgcccgctcg agcatgcac tagagggcc tattctatag 1140
 tgtcacctaa atgctagagc tcgctgatca gcctcgactg tgccttctag ttgccagcca 1200
 tctgttgttt gccctcccc cgtgccttcc ttgacctgg aaggtgccac tcccactgtc 1260
 ctttctaataaaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg 1320
 gggggtgggg tgggagcagga cagcaagggg gaggattggg aagacaatag caggcatgct 1380
 ggggatgagg tgggctctat ggcttctgag gcgaaagaa ccagtggcgg taatacggtt 1440
 atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 1500
 caggaaccgt aaaaaggccg cgttctggc gtttttccat aggtccgcc ccctgacga 1560
 gcatcacaaa aatcgacgt caagtcagag gtggcgaac ccgacaggac tataaagata 1620
 ccaggcgttt cccctggaa gctccctcgt gcctctcct gttccgacc tcccgttac 1680
 cggatactg tccgcttcc tccctcggg aagcgtggcg ctttctcata gtcacgctg 1740
 taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaacccc 1800
 cgttcagccc gaccgtgag cttatccgg taactatcgt cttgagtcca acccgtaag 1860
 acacgactta tcgcaactgg cagcagccac tgtaacagg attagcagag cgaggtatgt 1920

aggcgggtgct acagagttct tgaagtgggtg gcctaactac ggctacacta gaaggacagt 1980

attinggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 2040

atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 2100

gcgcaaaaa aaaggatctc aagaagatcc ttgatcttt tctacggggt ctgacgctca 2160

gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 2220

ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaacc 2280

tgaggctatg gcagggcctg ccgccccgac gttggctgcg agccctgggc cttcacccga 2340

acttgggggg tggggtgggg aaaaggaaga aacgcgggcg tattggcccc aatggggtct 2400

cgggtgggta tcgacagagt gccagccctg ggaccgaacc ccgctttat gaacaaacga 2460

cccaacaccg tgcgttttat tctgtctttt tattgcccgc atagcggggg ttccttccgg 2520

tattgtctcc ttcggtgitt cagttagcct cccctaggg tgggcgaaga actccagcat 2580

gagatccccg cgctggagga tcatccagcc ggcgtcccgg aaaacgattc cgaagcccaa 2640

cctttcatag aaggcggcgg tggaatcgaa atctcgtgat ggcaggttgg gcgtcgttg 2700

gtcggtcatt tcgaaccca gagtcccgt cagaagaact cgtcaagaag gcgatagaag 2760

gcgatcgcct gcgaatcggg agcggcgata ccgtaaagca cgaggaagcg gtcagcccat 2820

tcgccccaa gctcttcagc aatatcacgg gtagccaacg ctatgtcctg atagcgttcc 2880

gccacacca gccggccaca gtcgatgaat ccagaaaagc ggccattttc caccatgata 2940

ttcggcaagc aggcacgcc atgggtcacg acgagatcct cgccgtcggg catgctgcc 3000

ttgagcctgg cgaacagttc ggctggcgcg agcccctgat gctcttgatc atcctgateg 3060

acaagaccgg cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggteg 3120

aatgggcagg tagccggatc aagcgtatgc agccgccgca ttgcatcagc catgatggat 3180

actttctcgg caggagcaag gtgagatgac aggagatcct gccccggcac ttcgccaat 3240

agcagccagt cccttcccgc ttcagtgaca acgtcgagca cagctgcgca aggaacgccc 3300

gtcgtggcca gccacgatag ccgcgctgcc tcgtcttgca gttcattcag ggcaccggac 3360

aggtcggctt tgacaaaaag aaccggggcg ccttcgctg acagccggaa cacggcgga 3420

tcagagcagc cgattgtctg ttgtgccag tcatagccga atagcctctc cacccaagcg 3480

gccggagaac ctgcgtgcaa tccatcttgt tcaatcatgc gaaacgatcc tcctctgtc 3540

tcttgatcga tctttgcaaa agcctaggcc tccaaaaaag cctcctcact acttttgaa 3600

tagctcagag gccgaggcgg cctcggcctc tgcataaata aaaaaatta gtcagccatg 3660

ggcgggagaa tggcggaac tggcgggagt tagggcggg atggcgggag ttagggcggg 3720

gactatggtt gctgactaat tgagatgcat gctttgcata cttctgcctg ctggggagcc 3780

tggggacttt ccacacctgg ttgctgacta attgagatgc atgctttgca tacttctgcc 3840

tgctggggag cctggggact ttccacacc taactgacac acattccaca gctggttctt 3900

tccgcctcag gactcttct tttcaataa atcaatctaa agtatatatg agtaaacttg 3960

gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg 4020

ttcatccata gttgctgac tccccgtcgt gtagataact acgatacggg agggcttacc 4080

atctggcccc agtgtgcaa tgataccgag agaccacgc tcaccggctc cagatttacc 4140

agcaataaac cagccagccg gaagggccga ggcagaagt ggtcctgcaa ctttatccgc 4200

ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 4260

ttgcgcaac gttgttgcca ttgctacagg catcgtggg tcaagctcgt cgtttggat 4320
 ggcttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc ccatgttggtg 4380
 caaaaaagcg gttagctcct tcggctctcc gatcgttggtc agaagtaagt tggccgcagt 4440
 gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag 4500
 atgcttttct gtgactgggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg 4560
 accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt 4620
 aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct 4680
 gttgagatcc agttcgatgt aaccactcgc tgcaccaaac tgatcttcag catcttttac 4740
 tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggat 4800
 aaggcgaca cgaaatgtt gaatactcat actcttctt tttcaatatt attgaagcat 4860
 ttatcagggt taitgtctca tgagcggata catattgaa tgtatttaga aaaataaaca 4920
 aataggggtt ccgcgcacat tccccgaaa agtgccacct gacgcgcct gtagcggcgc 4980
 attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg ccagcgcct 5040
 agcgcgccgt ccttctgctt tcttccttc ctttctgcc acgttcgcc gctttccccg 5100
 tcaagctcta aatcgggggc tcccttagg gttccgattt agtgctttac ggcacctcga 5160
 ccccaaaaaa cttgattagg gtgatggttc acgtagtggg ccatcgcct gatagacggt 5220
 ttttcgccct ttgacttgg agtccactt ctttaatagt ggactcttgt tccaaactgg 5280
 aacaacactc aaccctatct cggctctatc ttttgattta taagggattt tgccgatttc 5340
 ggctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt ttaacaaaat 5400

attaacgctt acaatttac 5419

<210> 76
 <211> 5729
 <212> DNA
 <213> Artificial

<220>
 <223> pHYK/vif

<400> 76
 gcgcgcgttg acattgatta ttgactagtt attaatagta atcaattacg gggtcattag 60

ttcatagccc atatatggag ttccgcgta cataacttac ggtaaatggc cgcctggct 120

gaccgcecaa cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc 180

caatagggac tttcattga cgtcaatggg tggactatit acggtaaact gcccaacttg 240

cagtacatca agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat 300

ggccccctg gcattatgcc cagtacatga ctttatggga ctttctact tggcagtaca 360

tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc 420

gtggatagcg gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga 480

gtttgtttg gcaccaaatt caacgggact ttccaaatg tcgtaacaac tccgccccat 540

tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctctctggc 600

taactagaga acccaactgct tactggctta tcgaaattaa tacgactcac tatagggaga 660

cccaagcttg gtaccgagct cggatccgag atgcaaaatt catcccgcca ccaacaaaag 720

aaaaggaata aaaaacctgg accagaatta cccttagcac tatggatcca tatagcagaa 780

agcattaatg gggatagctc atggtacata acaatgagac tgcaacagat gatgtgggga 840

aaaagaggaa ataagttaca atataagaat gaagacaggg aatatgaaaa ttgggaaatt 900
 acatcatggg gatggaaaaat gcacctaagg agagtgaaac aatggataca agacaacagg 960
 agaggaagcc catggcagta caaagtagga ggaacatgga aaagtatagg agtgtggttc 1020
 ctgcaagcag gagattacag aaagtagac aggcacttct ggtgggcatg gaggatactg 1080
 atatgttctt gcaggaaaga aaagtttgat ataagagaat ttatgagagg aagacataga 1140
 tgggatttgt gcaaatcctg tgctcaagga gaagtagtaa agcatactag aacaaaaagt 1200
 ctggaagac tagtactgct acagatggta gaacagcatg tgtttcaagt attgccattg 1260
 tggagagcca ggagaagtag tacaacagat ttcccatggt gcagggacac aacgggatac 1320
 acgcatgctt ggtctgtcca ggagtgttgg ttgatggaat atctcttaga ggatgagtga 1380
 ccggaattct gcagataacc atcacactgg cggccgctcg agcatgcatc tagagggccc 1440
 tattctatag tgcacctaa atgctagagc tcgctgatca gcctcgactg tgccttctag 1500
 ttgccagcca tcgtttgttt gccctcccc cgtgccttcc ttgacctgga aaggtgccac 1560
 tcccactgtc ctttcctaataaaaatgagga aatgcatcg cattgtctga gtaggtgtca 1620
 ttctattctg gggggtgggg tgggacagga cagcaagggg gaggattggg aagacaatag 1680
 caggcatgct ggggatgcgg tgggctctat ggcttctgag gcggaagaa ccagtggcgg 1740
 taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 1800
 agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gttttccat aggctccgcc 1860
 ccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 1920
 tataaagata ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgacct 1980
 tgccgcttac cggataacctg tccgctttc tccttcggg aagcgtggcg ctttctcata 2040

gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc 2100
 acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 2160
 acccgtaag acacgactta tgcgccactgg cagcagccac tggtaacagg attagcagag 2220
 cgaggatgt aggcgggtgct acagagttct tgaagtgggt gcctaactac ggctacacta 2280
 gaaggacagt atttggatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 2340
 gtagctcttg atccggcaaa caaacaccg ctggtagcgg tggttttttt gtttgcaagc 2400
 agcagattac ggcagaaaa aaagatctc aagaagatcc tttgatcttt tctacggggt 2460
 ctgacgctca gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa 2520
 ggatcttcac ctatgcctt ttaaattaa aatgaagttt taaatcaatc taaagtatat 2580
 atgagtaacc tgaggctatg gcagggcctg ccgccccgac gttggctgcg agccctgggc 2640
 cttcacccga acttgggggg tggggtgggg aaaaggaaga aacgcgggcg tattggcccc 2700
 aatggggtct cggtggggta tgcacagagt gccagccctg ggaccgaacc ccgctttat 2760
 gaacaaacga cccaacaccg tgcgttttat tctgtctttt tattgccgtc atagcgggg 2820
 ttcttccgg tattgtctcc ttccgtggtt cagttagcct cccctaggg tgggcgaaga 2880
 actccagcat gagatccccg cgctggagga tcatccagcc ggcgtcccgg aaaacgattc 2940
 cgaagcccaa cctttcatag aaggcggcgg tggaatcgaa atctcgtgat ggcaggttgg 3000
 gcgtcgttg gtcggtcatt tcgaaccca gagtcccgtc cagaagaact cgtcaagaag 3060
 gcgatagaag gcgatgcgt gcgaatcggg agcggcgata ccgtaaagca cgaggaagcg 3120
 gtcagcccat tcgcccaaa gctcttcagc aatatacagg gtagccaacg ctatgtcctg 3180

atagcgggtcc gccacaccca gccggccaca gtcgatgaat ccagaaaagc ggccattttc 3240
 caccaatgata ttgggcaagc aggcacgccc atgggtcacg acgagatcct cgccgtcggg 3300
 catgctcgcc ttgagcctgg cgaacagttc ggctggcgcg agcccctgat gctcttgatc 3360
 atcctgatcg acaagaccgg cttccatccg agtacgtgct cgctcgatgc gatgtttcgc 3420
 ttggtggctg aatgggcagg tagccggatc aagcgtatgc agccgccca ttgcatcagc 3480
 catgatggat actttctcgg caggagcaag gtgagatgac aggagatcct gccccggcac 3540
 ttcgccaat agcagccagt ccttcccgc ttcagtgaca acgtcgagca cagctgcgca 3600
 aggaacgcc gtcgtggcca gccacgatag ccgcgctgcc tcgtcttgca gttcattcag 3660
 ggcaccggac aggtcggctt tgacaaaaag aaccgggcgc ccctgcgctg acagccggaa 3720
 cacggcgga tcagagcagc cgattgtctg ttgtgccag tcatagccga atagcctctc 3780
 cacccaagcg gccggagaac ctgctgcaa tccatcttgt tcaatcatgc gaaacgatec 3840
 tcacctctgc tcttgatcga tctttgcaa agcctaggcc tccaaaaaag cctcctcact 3900
 acttctggaa tagctcagag gccgaggcgg cctcggcctc tgcataaata aaaaaatta 3960
 gtcagccatg ggcggagaaa tgggcggaac tgggcggagt tagggcggg atggcgggag 4020
 ttagggcggg gactatggtt gctgactaat tgagatgat gctttgcata cttctgcctg 4080
 ctggggagcc tggggacttt ccacacctgg ttgctgacta attgagatgc atgctttgca 4140
 tacttctgcc tctggggag cctggggact ttccacacc taactgacac acattccaca 4200
 gctggttctt tccgctcag gactcttct tttcaataa atcaatctaa agtatatatg 4260
 agtaaacttg gctcgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 4320
 gctatattcg ttcatccata gttgectgac tcccctcgt gtagataact acgatacggg 4380

agggttacc atctggccc agtgetgcaa tgataccgag agaccacgc tcaccggctc 4440

cagatttacc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 4500

ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 4560

cagttaatag ttigcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt 4620

cgtttggat ggcttcattc agtccgggt cccaacgac aaggcgagtt acatgatccc 4680

ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgc agaagtaagt 4740

tggccgagc gttatcactc atggttatgg cagcaactgca taattctctt actgtcatgc 4800

catccgtaag atgcttttct gtgactgggt agtactcaac caagtcattc tgagaatagt 4860

glatgcccgc accgagttgc tcttgcccgg cgtcaatag ggataatacc gcgccacata 4920

gcagaacttt aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga 4980

tcttaccgct gttgagatcc agttcgatgt aaccactcg tgcaccaac tgatctttag 5040

catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 5100

aaaagggaat aaggcgaca cggaaatgt gaatactcat actcttcctt tttcaatatt 5160

attgaagcat ttatcagggt tattgtctca tgagcggata catattttaa tgtatttaga 5220

aaaataaaca aataggggtt ccgcccacat ttccccgaaa agtgccacct gacgcgcct 5280

gtagcggcgc attaagcgcg gcgggtgtgg tggttacgag cagcgtgacc gctacacttg 5340

ccagcgcct agcggcctt ctttctctt tcttccctt ctttctgccc acgttcgccc 5400

gctttcccg tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac 5460

ggcacctcga cccccaaaaa cttgattagg gtgatgttc acgtagtggg ccacgcctt 5520

gatagacggt ttttcgcctt ttgacgttgg agtccacggt ctttaaatagt ggactcttgt 5580
 tccaaactgg aacaacactc aacctatct cggctctattc ttttgattta taagggattt 5640
 tgccgatttc ggctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt 5700
 ttaacaaaat attaacgctt acaatttac 5729

<210> 77
 <211> 9446
 <212> DNA
 <213> Artificial

<220>
 <223> pMGP/RRE

<400> 77
 aacagaaaag ttccattgga gcccaagtaca ttgagtcaat agggactttc caatgggttt 60
 tgcccagtac ataaggtaaa tgggaggtaa gccaatgggt ttttccatt actggcacgt 120
 aactgagtc attaggact ttccaatggg ttttgcccag tacataaggt caatagggtt 180
 gaatcaacag gaaagtccca ttggagccaa gtacactgag tcaataggga ctttccattg 240
 ggttttgccc agtacaaaag gtcaataggg ggtgagtcaa tgggtttttc ccattattgg 300
 cacgtacata aggtcaatag gggtagtca tgggtttttt ccagccaatt taattaaac 360
 gccatgtact ttcccacat tgacgtcaat gggtattga aactaatgca acgtgacctt 420
 taaacgtac tttccatag ctgattaatg ggaaagtacc gttctcgagc caatacacgt 480
 caatgggaag tgaaggga gccaaaact aacaccgcc cggttttccc tggaaattcc 540
 atattggcac gcattctatt ggctgagctg cgttcacgtg ggtataagag gcgcgaccag 600
 cgteggtaac gtcgcagtct tcggctgac caccgtagaa cgcagagctc ctgctgcag 660

gcaagcttgg taccgagctc ggatcccggg gaggtaccaa aatccgcat ggtgagtcta 720
gatagagaca tggcgaggca agtctccggg gggaaaagag attatcctga gctcgaaaaa 780
tgtatcaagc atgcatgcaa gataaaagt tgcactcagag gggagcactt gacagaagga 840
aattgtttat ggtgccttaa aacattagat tacatgtttg aggaccataa agaggaacct 900
tggacaaaag taaaatttag gacaatatgg cagaaggatga agaatctaac tctgaggag 960
agtaacaaaa aagactttat gtccttgcag gccacattag cgggtctaata gtgttgccaa 1020
atgggatga gacctgagac attgcaagat gcaatggcta cagtaatcat gaaagatggg 1080
ttactggaac aagaggaaaa gaaggaagac aaaagagaaa aggaagagag tgtcttccca 1140
atagtagtgc aagcagcagg agggagaagc tggaaagcag tagattctgt aatgttcag 1200
caactgcaaa cagtagcaat gcagcatggc ctcggtctg aggactttga aaggcagttg 1260
gcatattatg ctactacctg gacaagtaaa gacatactag aagtattggc catgatgcct 1320
ggaaatagag ctcaaaagga gtttaattcaa gggaaattaa atgaagaagc agaaaggtgg 1380
agaaggaata atccaccacc tccagcagga ggaggattaa cagtggatca aattatgggg 1440
gtaggacaaa caaatcaagc agcagcacia gctaacaatgg atcaggcaag gcaaatatgc 1500
ctgcaatggg taataaatgc attaagagca gtaagacata tggcgcacag gccaggaat 1560
ccaatgctag taaagcaaaa aacgaatgag ccatatgaag atttgcagc aagactgcta 1620
gaagcaatag atgcagagcc agttacacag cctataaaag attatctaaa gctaacta 1680
tcttatacaa atgcatcagc agattgtcag aagcaaatgg atagaact aggacaaaga 1740
gtacaacaag ctagtgtaga agaaaaatg caagcatgta gagatgtggg atcagaaggg 1800
ttcaaatgc aattgttagc acaagcatta aggccaggaa aaggaaaagg gaatggacag 1860

ccacaaaggt gttacaactg tggaaaaccg ggacatcaag caaggcaatg tagacaagga 1920

atcatatgtc acaactgtgg aaagagagga catatgcaaa aagaatgcag aggaaagaga 1980

gacataaggg gaaaacagca gggaaacggg aggaggggga tacgtgtggt gccgtccgct 2040

cctcctatgg aataacttca gcaccaccia tggttcaggt ccgcataggt tcccagcaga 2100

ggaacttgtt atttgatacc ggggcggacc gaactatagt tagatggcat gagggctcgg 2160

gaaaccagc cggaaggata aaactgcaag gaataggagg aatagtagaa ggagaaaaat 2220

ggaataatgt agaattagaa tataaaggag aaacaagaaa gggaacaata gtagtgttac 2280

cacaaagtcc agtagaagta ttaggacgag ataacatggc ccgatttga ataaagataa 2340

taatggcaaa tttagaggaa aaaagaatcc caattacaaa agtaaaattg aaagagggat 2400

gtacgggtcc acatgtccca caatggccat taacagaaga gaaattaaaa ggtctaacag 2460

aatcataga taaattagtg gaagaaggaa aactaggaaa ggcaccccca cattggacat 2520

gtaatactcc aatcttttgc ataaaaaga aatcagggaa gtggagaatg ttaatagatt 2580

tcagagaatt gaacaaacag acagaagatt taacagaagc gcagttagga ctcccgcac 2640

cgggaggact acaaaagaaa aaacatgtta caatattgga cataggagat gcatatttta 2700

ctataccct atatgaacca tategagagt acacatgttt tactctatta agtcctaata 2760

atctaggacc atgtaaaaga tactattgga aagtgctgcc acaagttgg aaattgagtc 2820

catctgtata tcaatttact atgcaggaga tcttagagga ttggatacag cagcatccag 2880

aaattcaatt tggcatatat atggatgata ttacatagg aagtattta gaaattaaaa 2940

agcatagaga aatagtgaag gatttagcca attatattgc ccaatatgga ttcactctgc 3000

cagaagagaa gagacaaaag ggatatccag caaaatggct aggatttgaa ctacacccgc 3060
 agacctggaa atttcagaag catacattac ctgaattaac aaaggaaca ataacattaa 3120
 ataaattaca gaaattagta ggagaattag tatggagaca atccataatt gggaaaagca 3180
 ttccaaacat tctgaaatta atggaaggag atagagaatt acaaagttaa agaaaaattg 3240
 aagaagtaca tctgaaagaa tgggaagcat gtaggaaaaa attagaagaa atggaaggaa 3300
 attattataa taaagacaaa gatgtctatg gacaattggc ttggggagac aaagctatag 3360
 aatatatagt gtatcaggag aaagggaaac cattatgggt aaatgtggtt cacaatataa 3420
 agaacctaag catccccaa caggttatta aagcagcgc aaaattaacc caagaagtca 3480
 tcattaggac aggaaaaata ccatggatat tgttgccagg gaaagaagaa gattggagac 3540
 tagaattgca attagggaaac atcacatgga tgccaaaatt ttggtcctgt tatcaggac 3600
 atacaagatg gagaaaaaga aatataatag aagaagtagt agaaggcct acatattata 3660
 cagatggagg aaaaaagaat aaagtaggaa gtctagggtt catagtatca acaggggaaa 3720
 aatttagaaa gcatgaagag ggcacaaacc agcaactaga attaagagcc atagaggaag 3780
 ctctaaaaca aggcctcaa acaatgaatt tagtaacaga tagtagatat gcatttgaat 3840
 ttttattaag aaattgggat gaagaagtaa taaagaatcc aattcaagca agaattatgg 3900
 aaattgcca caagaaagat aggataggag tgcattgggt gccaggacat aaagggattc 3960
 cccaaaatga agaaatagac aaatatattt cggaaatatt tcttgcaaaa gaaggagaag 4020
 gaattctccc aaaaagagaa gaggatgcag ggtatgattt aatatgcca gaagaggtta 4080
 ccatagagcc aggacaagtg aaatgcatcc ccatagagct aagattaaat ttaaagaaat 4140
 cacaatgggc tatgattgct acaaaaagca gcatggctgc caaaggagtg ttcacacaag 4200

gaggaatcat agactcagga tatcaggac aaatacaggt aataatgtat aatagcaata 4260

aatagcagt agicataccc caaggagaaa aatttcaca attaataatta atggataaaa 4320

agcatgaaaa attggaaccc tgggggaaa gcagaaaaac agaaaggga gaaaaaggat 4380

ttgggtctac aggaatgtat tggatagaaa atattcctct ggcagaggaa gaccacaaa 4440

aatggcatca agatgccca tcattgcac tagaatttga aattccaaga acagcagcag 4500

aagacatagt aaatcaatgt gaaatagca aagaaggcag gacacctgca gtaattagag 4560

gcggaaacaa aaggggggta gatcattggc aagtggatta taccattat gaaaatatca 4620

tactattagt atggtagaa acaaatcag gactaatata tgcagaaaa gtaaaaggag 4680

aatcaggca agaattcaga ataaaagtga tgcaatggta tgcattattt ggtccagagt 4740

cattgcagtc agacaatgga cctgcatttg cagcagagcc cacacagctg ttaatgcaat 4800

acctaggagt aaaacacaca acaggcatac cttggaatcc acagtctcag gctatagtag 4860

aaaggcaca tcaactattg aaaagcactt taaagaagt ccagccacaa tttgtcgctg 4920

tagaatcagc catagcagca gccctagtcg ccataaatat aaaaagaaag gttgggctgg 4980

ggacaagccc tatggatatt tttatatata ataaagaaca gaaaagaata aataataat 5040

ataataaaaa ttctcaaaa attcaattct gttattacag aataaggaaa agaggacatc 5100

caggagagt gaaaggacca acccaggtac tgtggaagg ggaaggagca attgtggtaa 5160

aggatataga aagtgaaaag tatttagtaa taccttaca agatgcaaaa ttcattccgc 5220

accaacaaa agaaaaggaa taaaaaacct ggaccagaat tacccttagc actatggatc 5280

cactagtaac ggccgccagt gtgctggaat tetgcagata tccatcacac tggcggccgg 5340

gctgcaggaa ttcatagaa aagatatcaa aaacaagaat cataaacagg aaaaaagag 5400

aactcagcca caagaggaag aagagaggcg ttggcttggg cattatgcta gttatcatgg 5460

caatagtagc tgccgcaggg gcttctctgg gagtcgcaaa cgcgattcag cagtcttaca 5520

ctaaggcagc tgiccagacc cttgctaag caactgctgc acagcaggat gtgtagagg 5580

caacctatgc catggtacag catgtggcta aaggcgtacg aatcttgaa gctcgagtgg 5640

ctcgagtgga agctatcaca gatagaatag cggccgcca tcaagcttat cgataccgtc 5700

ggccgctcga gcatgcatct agagggcct attctatagt gtcacctaaa tgctagagct 5760

cgctgatcag cctcgactgt gccttctagt tgccagccat ctgttgtttg cccctcccc 5820

gtgccttctt tgacctgga aggtgccact cccactgtcc tttcctaata aatgaggaa 5880

attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtggggg ggggcaggac 5940

agcaaggggg aggattggga agacaatagc aggcattgctg gggatgcggt gggctctatg 6000

gcttctgagg cggaaagaac cagtggcggg aatacggtta tccacagaat caggggataa 6060

cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc 6120

gttgctggcg tttttccata ggctcgcgcc cctgacgag catcacaaaa atcgacgctc 6180

aagtcagagg tggcgaacc cgacaggact ataaagatac caggcgttcc ccctggaag 6240

ctccctcgtg cgctctctg ttccgacct gccgcttacc ggatacctgt ccgcctttct 6300

cccttcggga agcgtggcgc tttctcatag ctacagctgt aggtatctca gttcgggtga 6360

ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc accgctgcgc 6420

cttatecggg aactatcgtc ttgagtccaa cccgtaaga cagacttat cgccactggc 6480

agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt 6540

gaagtggtag cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct 6600

gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc 6660

tggtagcggg ggtttttttg tttgcaagca gcagattacg cgagaaaaa aaggatctca 6720

agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta 6780

agggatcttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaa 6840

atgaagtttt aaatcaatct aaagtatata tgagtaacct gaggctatgg cagggcctgc 6900

cgccccgacg ttggctgcga gccctgggccc ttaccgccaa cttggggggg ggggtgggga 6960

aaaggaagaa acgcgggcgt attggcccca atggggcttc ggtggggtat cgacagagtg 7020

ccagccctgg gaccgaacct cgcgtttatg acaaacgac ccaacaccgt gcgttttatt 7080

ctgtcttttt attgcegtca tagcgcgggt tccttcgggt attgtctctt tccgtgtttc 7140

agttagcctc ccctagggt gggcgaagaa ctccagcatg agatccccgc gctggaggat 7200

catccagccg gcgtcccga aaacgattcc gaagcccaac ctttcataga aggccgagggt 7260

ggaatcgaaa tctcgtgatg gcaggttggg cgtcgttgg tcggtcattt cgaaccccag 7320

agtcccgtc agaagaactc gtcaagaagg cgatagaagg cgatgcgctg cgaatcgga 7380

gcggcgatac cgtaaagcac gaggaagcgg tcagccatt cgccccaag ctcttcagca 7440

atatcacggg tagccaacgc tatgtcctga tagcgggtccg ccacaccag cggccacag 7500

tcgatgaatc cagaaaagcg gccatcttcc accatgatat tcggcaagca ggcacgcca 7560

tgggtcacga cgagatcctc gccgtcgggc atgctcgcct tgagcctggc gaacagttcg 7620

gctggcgcga gccctgatg ctcttgatca tctgatcga caagaccggc ttccatccga 7680

gtacgtgctc gctcgatgcg atgtttcgcct tggtaggtcga atgggcaggt agccggatca 7740

agcgtatgca gccgccgcat tgcattcagcc atgatggata ctttctcggc aggagcaagg 7800

tgagatgaca ggagatcctg ccccggaact tcgccaata gcagccagtc ccttcccgt 7860

tcagtgacaa cgtcagcac agctcgcgca ggaacgccg tcgtggccag ccacgatagc 7920

cgcgctgcct cgtcttgagc ttcatcagc gcaccggaca ggtcggcttt gacaaaaaga 7980

accggggccc cctgcgctga cagccggaac acggcgcat cagagcagcc gattgtctgt 8040

tgtgcccagt catagccgaa tagcctctcc acccaagcgg ccggagaacc tgcgtgcaat 8100

ccatcttggt caatcatgcg aaacgatcct catcctgtct cttgatgat ctttgcaaaa 8160

gcctaggcct ccaaaaaagc ctctcacta cttctggaat agctcagagg ccgaggcggc 8220

ctcggcctct gcataataa aaaaaattag tcagccatgg ggcggagaat gggcggaact 8280

ggcgaggatt agggcgggga tggcgaggat tagggcgggg actatggttg ctgactaatt 8340

gagatgatg ctttgatac ttctgcctgc tggggagcct ggggactttc cacacctggt 8400

tgctgactaa ttgagatgca tgctttgcat acttctgcct gctggggagc ctggggactt 8460

tccacacct aactgacaca cattccacag ctggttcttt ccgcctcagg actcttctt 8520

ttcaataaa tcaatctaaa gtatatatga gtaaacttg tctgacagtt accaatgctt 8580

aatcagtgag gcacctatct cagcagctg tctatttctg tcatccatag ttgcctgact 8640

ccccgctg tagataacta cgatacggga gggttacca tctggccca gtgctgcaat 8700

gataccgca gaccacgct caccggctcc agatttatca gcaataaacc agccagccgg 8760

aagggccgag cgcagaagtg gtctgcaac ttatccgcc tccatccagt ctattaattg 8820

ttccgggaa gctagagtaa gtagttgcc agttaatagt ttgcgcaac ttgttgccat 8880

tgctacagge atcgtggtgt cacgctcgtc gtttggatg gcttcattca getccggttc 8940
 ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt 9000
 cggtcctccg atcgtttgca gaagtaagtt ggccgcagtg ttatcactca tggttatggc 9060
 agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga 9120
 gtactcaacc aagtcattct gagaatagtg tatgcgcgca ccgagttgct cttgcccggc 9180
 gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgtca tcattgaaa 9240
 acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta 9300
 acccactcgt gcacccaact gatcttcagc atcttttact ttaccagcg tttctgggtg 9360
 agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg 9420
 aataactata ctcttctttt ttcaat 9446

<210> 78
 <211> 7856
 <212> DNA
 <213> Artificial

<220>
 <223> pCAH/SIND60/hlacZ

<400> 78
 gttgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata 60
 gccatataat ggagttccgc gttacataac ttacggtaaa tggcccgctt ggctgaccgc 120
 ccaacgacce ccgccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag 180
 ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac 240

atcaagtga tcatatgcca agtaccccc ctattgacgt caatgacggt aaatggcccc 300
 cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg 360
 tattagtcat cgctattacc atgggatgac ggttttggca gtacatcaat gggcgtggat 420
 agcggtttga ctacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt 480
 tttggcacca aaatcaacgg gactttcaa aatgtcgtaa caactccgcc ccattgacgc 540
 aaatgggcgg taggcgtgta cgggtgggagg tctatataag cagagctcaa gcttgctgct 600
 tgcacttcag agttctagga gagtccctcc tagtctctcc tctccgagga ggtaccgaga 660
 cctcaaaata aaggagtgat tgccttactg ccgagtggag agtgattact gagcggccgg 720
 tgtatcgga gtcgtccctt aatctgtgca ataccagagc ggctctcgca gctggcgccc 780
 aacgtggggc ccgaggagaa gaaaagaaag cggccctgag aactcgctt ctgaaaaga 840
 ggaagaggac aagtgtctat agcaacaaga gagaagaagt agagcaaagg tccagtggct 900
 cgaaaaaga ggaactgaaa cttcggggac gcctgaagga gtaagtaag tgactctgct 960
 gtacgcgggg cgaggcagag gtttccttct aaattgaaag agaagtgtt ctgagagagg 1020
 tcttggtggt cgagaatcct gtacaaaaa aaggagggat ctcggtcagg accaggaccc 1080
 ctgggagtaa tacaacagca acaccgtaag aaaatccgcc taggtgagtc tagatagaga 1140
 ctaggcgagg caagtctccg ggggaaaag agattatcct gcaggaattc gattctagag 1200
 gtgatagaaa tgccagaaaa ctatgcaaaa acaagaatca taaacaggaa aaaaagagaa 1260
 ctcagccaca agaggaagaa gagaggcgtt ggcttggca ttatgctagt tatcatggca 1320
 atagtagctg ccgacggggc ttctctggga gtcgcaaag cgattcagca gtcttacct 1380
 aaggcagctg tccagacct tgctaatgca actgctgcac agcaggatgt gttagaggca 1440

acctaigcca tggtagacga tgtggctaaa ggcgtacgaa tcttgggaagc tegagtggct 1500

cgagtggaag ctatcacaga tagaataatg ctataccaag aattggattg ttggcactag 1560

gatccatcag ccaccattaa cgcttacaat ttacgcgcgc gttgacattg attattgact 1620

agttattaat agtaatcaat tacggggtca ttagttcata gcccatatat ggagtccgc 1680

gttacataac ttacggtaaa tggcccgcct ggctgaccgc ccaacgaccc ccgccattg 1740

acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa 1800

tgggtggact atttacggta aactgccac ttggcagtac atcaagtgt tcatatgcca 1860

agtaccccc ctattgacgt caatgacggt aatggcccg cctggcatta tgcccagtac 1920

atgaccttat gggactttcc tacttggcag tacatctacg tattagtcat cgctattacc 1980

atggtgatgc ggttttgca gtacatcaat gggcgtggat agcggtttga ctacacggga 2040

ttccaagtc tccacccat tgacgtcaat gggagtttgt tttggcacca aatcaacgg 2100

gactttcaa aatgtcgtaa caactccgc ccattgacgc aatggcgg taggcgtgta 2160

cggtgggagg tctataaag cagagctctc tggctaacta gagaaccac tgcttactgg 2220

cttatcgaat ttaatacagc tcactatagg gagaccaag ctgcttacca tggggggttc 2280

tcatcatcat catcatcatg gtatggctag catgactggt ggacagcaa tgggtcggga 2340

tctgtacgac gatgacgata aggtacctaa ggatcagctt ggagtgtac ccgtcgtttt 2400

acaacgtcgt gactgggaaa accctggcgt tacccaactt aatgccttg cagcacatcc 2460

ccctttgcc agctggcgta atagcgaaga ggeccgcacc gatcgcctt cccaacagtt 2520

gcgagcctg aatggcgaat ggcctttgc ctggtttccg gcaccagaag cgggtgccgga 2580

aagctggctg gaggcgcgac ttccctgaggc cgatactgtc gtcgtccct caaactggca 2640
gatgcacggt tacgatgcgc ccatctacac caacgtaacc tatcccatta cggteaatec 2700
gccgtttgtt cccacggaga atccgacggg ttgttactcg ctacattta atgttgatga 2760
aagctggcta caggaaggcc agacgcgaat ttttttgat ggcgttaact cggcgtttca 2820
tctgtggtgc aacgggcgct gggctcggtta cggccaggac agtcgtttgc cgtctgaatt 2880
tgacctgagc gcatttttac gcgccggaga aaaccgcctc gcggtgatgg tgctgcgttg 2940
gagtgacggc agttatctgg aagatcagga tatgtggcgg atgagcggca ttttccgtga 3000
cgtctcgttg ctgcataaac cgactacaca aatcagcgat ttccatgttg ccaactcgctt 3060
taatgatgat ttcagccgcg ctgtactgga ggctgaagtt cagatgtgcg gcgagttgcg 3120
tgactaccta cgggtaacag tttctttatg gcagggtgaa acgcaggtcg ccagcggcac 3180
cgcgctttc ggcggtgaaa ttatcgatga gcgtgggtgt tatgccgatc gcgtcacact 3240
acgtctgaac gtcgaaaacc cgaaactgtg gagcgccgaa atcccgaatc tctatcgtgc 3300
gggtggtgaa ctgcacaccg ccgacggcac gctgattgaa gcagaagcct gcgatgtcgg 3360
tttcccgag gtgcggttg aaaatggtct gctgctgctg aacggcaagc cgttgctgat 3420
tcgagcggtt aaccgtcacg agcatcatcc tctgcatggt caggatcatgg atgagcagac 3480
gatggtgcag gatatcctgc tgatgaagca gaacaacttt aacgccgtgc gctgttcgca 3540
ttatccgaac catccgctgt ggtacacgct gtgcgaccgc tacggcctgt atgtggtgga 3600
tgaagccaat attgaaacc acggcatggt gccaatgaat cgtctgaccg atgatccgcg 3660
ctggctaccg gcgatgagcg aacgcgtaac gcgaatggtg cagcgcgatc gtaatcacc 3720
gagtgatgac atctggtcgc tggggaatga atcaggccac ggcgctaate acgacgcgt 3780

gtatcgctgg atcaaatctg tcatccttc ccgcccggg cagtatgaag gcggcgggagc 3840
 cgacaccacg gccaccgata ttatttgccc gatgtacgcg cgcgtggatg aagaccagcc 3900
 ctccccgct gtgccgaaat ggtccatcaa aaaatggctt tcgctacctg gagagacgcg 3960
 cccgctgac ctttgcgaat acgcccacgc gatgggtaac agtcttgcg gtttcgctaa 4020
 atactggcag gcgtttcgtc agtatccccg ttacagggc ggcttcgtct gggactgggt 4080
 ggatcagtcg ctgattaat atgatgaaaa cggcaaccg tggtcggctt acggcgggta 4140
 ttttgccgat acgccgaac atcgccagt tctgtatgaac ggtctgtct ttgccaccg 4200
 cagcccgcat ccagcgtga cggaaagcaa acaccagcag cagttttcc agttccgttt 4260
 atccgggcaa accatcgaag tgaccagcga atacctgttc cgtcatagcg ataacagct 4320
 cctgcactgg atggtggcgc tggatggtaa gccgctggca agcgggtaag tgcctctgga 4380
 tgtcgtcca caagtaaac agttgattga actgcctgaa ctaccgcagc cggagagcgc 4440
 cgggcaactc tggctcacag tacgcgtagt gcaaccgaac gcgaccgcat ggtcagaagc 4500
 cgggcacatc agcgcctggc agcagtgcg tctggcggaa aacctcagtg tgacgtccc 4560
 cgcccgctcc cacccatcc cgcctctgac caccagcga atggattttt gcatcgagct 4620
 gggtataaag cgttggcaat ttaaccgcca gtcaggcttt cttcacaga tgtgattgg 4680
 cgataaaaaa caactgctga cgccgtgcg cgtacgttc acccgtgcac cgctggataa 4740
 cgacattggc gtaagtgaag cgaccgcat tgaccctaac gcctgggtcg aacgctggaa 4800
 ggccggcgggc cattaccagg ccgaagcagc gttgttcag tgcacggcag atacacttc 4860
 tgatcgctg ctgattacga ccgctcacgc gtggcagcat caggggaaaa cttatttat 4920

cagccggaaa acctaccga ttgatgtag tggtaaatg gcgattaccg ttgatgtga 4980

agtggcgagc gatacacgc atccggcgcg gattggcctg aactgccagc tggcgcaggt 5040

agcagagcgg gtaaactggc tcggattagg gcccaagaa aactatccc accgccttac 5100

tgccctgt ttgaccct gggatctgc attgtcagac atgatacc cgtacgtctt 5160

cccagcga aacggtctgc gctgcccgc gcgcaattg aattatggc cacaccagt 5220

gcgcccgcac ttccagtca acatcagcc ctacagtcaa cagcaactga tggaaaccag 5280

ccatgccat ctgctgcag cggaagaagg cacatggctg aatatcgac gtttccatat 5340

ggggattggt ggcgacgact cctggagccc gtcagtatcg gcggaattac agctgagcgc 5400

cggtcgtac cattaccagt tggctcgtg tcaaaaataa taaagccgaa ttctgcagat 5460

atccagcaca gtggcggccg ctagcacaaa aataaaaaa gaaaggtga ctgtgagaca 5520

tgggctaaag aggagcggcc gctcgagtct agaactagt gatcagctt gctgcttga 5580

cttcagagtt ctaggagagt ccctcctagt ctctctctc cgaggaggta ccgagacctc 5640

aaaataaagg agtgattgcc ttactgccga gtggagagtg attactgagc ggccggtgta 5700

tcgggagtcg tccttaatc tgtgaatac cagagcggct ctgcagccg acctcgaggg 5760

ggggccctat tctatagtgt cacctaaatg ctgagctcg ctgatcagc tcgactgtgc 5820

cttctagttg ccagccatct gttgtttgcc cctccccgt gccttcttg acctggaag 5880

gtgccactcc cactgtcctt tctaataaa atgaggaat tgcacgcat tgtctgagta 5940

ggtgtcattc tttctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag 6000

acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctgaggcg gaaagaacca 6060

gtggcggtaa tacggttacc cacagaatca ggggataacg caggaaagaa catgtgagca 6120

aaagccagc aaaagccag gaacctaaa aagccgcgt tgcggcgtt tttccatagg 6180
 ctccgcccc ctgacgagca tcacaaaaat cgacgtcaa gtcagagggt gcgaaaccg 6240
 acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt 6300
 cgcacctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 6360
 tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc 6420
 tgtgtgcacg aacccccct tcagcccgac cgtgcgctt tatccggtaa ctatcgtctt 6480
 gagtccaacc cgtaagaca cgacttatcg cactggcag cagccactgg taacaggatt 6540
 agcagagcga ggtatgtagg cgggtctaca gatttcttga agtgggtggc taactacggc 6600
 tacactagaa ggacagtatt tggatatcgc gctctgctga agccagttac ctccgaaaa 6660
 agagtggta gctcttgatc cggcaaaaa accaccgctg gtagcgggtg ttttttgtt 6720
 tgcaagcagc agattacgag cagaaaaaaaa ggatctcaag aagatccttt gatctttct 6780
 acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttgtt catgagatta 6840
 tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa 6900
 agtatatatg agtaacctga tcaggactct tccttttcat gaacaataaa actgtctgct 6960
 tacataaaca gtaatacaag ggggtttatg agccatattc aacgggaaac gtcttgctct 7020
 agcccgcat taaattcaa catggatgct gatttatatg ggtataaatg ggctcgcgat 7080
 aatgtcgggc aatcagggtc gacaatctat cgattgtatg ggaagcccga tgcgccagag 7140
 ttgtttctga aacatggcaa aggtagcgtt gccaatgatg ttacagatga gatggtcaga 7200
 ctaaactggc tgacggaatt tatgctcttt ccgaccatca agcattttat ccgtactcct 7260

gatgatgcat ggttactcac cactgcgac cccgggaaaa cagcattcca ggtattagaa 7320
 gaatacctg attcaggatga aaatattggt gatgcgctgg cagtgttctt gcgccggttg 7380
 cattcgattc ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg tctcgctcag 7440
 gcgcaatcac gaatgaataa cggtttgggt gatgcgagtg atttgatga cgagcgaat 7500
 ggctggcctg ttgaacaagt ctggaagaa atgcataaac ttttgccatt ctaccggat 7560
 tcagtcgca ctcatggatga tttctcactt gataacctta tttttgacga ggggaaatta 7620
 ataggttga ttgatgttg acgagtcgga atcgagacc gataccagga tcttgccatc 7680
 ctatggaact gcctcggatga gttttctctt tcattacaga aacggctttt tcaaaaatat 7740
 ggtattgata atcctgatat gaataaattg cagtttcatt tgatgctcga tgagtttttc 7800
 taagaattcg cgcaattaac cctcaactaaa gggaacaaaa gctgggtacc gggccc 7856

<210> 79
 <211> 8127
 <212> DNA
 <213> Artificial

<220>
 <223> pCAH/SINd1/hlacZ

<400> 79
 gttgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata 60
 gccatataat ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc 120
 ccaacgacc ccgccattg acgtcaataa tgacgtatgt tccatagta acgccaatag 180
 ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgccac ttggcagtac 240
 atcaagtga tcatatgcca agtacgccc ctattgacgt caatgacgtt aatggcccg 300

cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg 360

tattagtcat cgctattacc atggatgatgc ggttttggca gtacatcaat gggcgtggat 420

agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt 480

tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc 540

aaatgggcgg taggcgtgta cgggtgggagg tctatataag cagagctcaa gcttgctgct 600

tgcacttcag agttctagga gagtcctcc tagtctctcc tctccgagga ggtaccgaga 660

ctcaaaata aaggagtgat tgccttactg ccgagtggag agtgattact gagcggccgg 720

tgtatcggga gtcgtccctt aatctgtgca ataccagagc ggctctcgca gctggcgccc 780

aacgtggggc ccgaggagaa gaaaagaaag cggccctgag aactcgctt ctgaaaaaga 840

ggaagaggac aagttagctat agcaacaaga gagaagaagt agagcaaagg tccagtggct 900

cggaaaaaga ggaactgaaa cttcggggac gcctgaagga gtaagtaag tgactctgct 960

gtacgcgggg cgaggcagag gtttccttct aaattgaaag agaagtgtg ctgcgagagg 1020

tcttgggtgt cgagaatcct gtacaaaaaa aaggagggat ctcggtcagg accaggaccc 1080

ctgggagtaa tacaacagca acaccgtaag aaaatccgcc taggtgagtc tagatagaga 1140

ctaggcgagg caagtctccg gggggaaaag agattatcct gagctcgaaa aatgtatcaa 1200

gcatgcatgc aagataaaag ttcgactcag aggggagcac ttgacagaag gaaattgttt 1260

atggtgcctt aaaacattag attacatggt tgaggacat aaagaggaac cttggacaaa 1320

agtaaaattt aggacaatat ggcagaaggt gaagaatcta actcctgagg agagtaacaa 1380

aaaagacttt atgtctttgc aggccacatt agcgggtcta atgtgttgc aaatggggat 1440

gagaccgggc tgcaggaatt cgattctaga ggtgatagaa atgccagaaa actatgcaaa 1500

aacaagaatc ataacagga aaaaaagaga actcagccac aagaggaaga agagaggcgt 1560
tggcttggtc attatgctag ttatcatggc aatagtagct gccgcagggg cttctctggg 1620
agtcgcaaac gcgattcagc agtcttacac taaggcagct gtccagaccc ttgctaatagc 1680
aactgctgca cagcaggatg tgtagaggc aacctatgcc atggtacagc atgtggctaa 1740
aggcgtacga atcttgggaag ctcgagtggc tcgagtggaa gctatcacag atagaataat 1800
gctataccaa gaattggatt gttggcacta ggatccatca gccaccatta acgcttacia 1860
tttacgcgcg cgttgacatt gattattgac tagttattaa tagtaatcaa ttacggggtc 1920
attagtcat ageccatata tggagtccg cgttacataa cttacggtaa atggcccgcc 1980
tggctgaccg cccaacgacc cccgccatt gacgtcaata atgacgtatg ttcccatagt 2040
aacgccaata gggactttcc attgacgtca atgggtggac tatttacggt aaactgcca 2100
cttggcagta caicaagtgt atcatatgcc aagtacgccc cctattgacg tcaatgacgg 2160
taaatggccc gcctggcatt atgcccagia catgacctta tgggactttc ctacttggca 2220
gtacatctac gtattagca tcgctattac catgggatg cggttttggc agtacatcaa 2280
tggcgtgga tagcgtttg actcacgggg atttccaagt ctccaccca ttgacgtcaa 2340
tgggagtttg ttttggcacc aaaatcaacg ggactttcca aaatgtcgta acaactccgc 2400
cccattgacg caaatggcg gtagcgtgt acggtgggag gtctatataa gcagagctct 2460
ctggctaact agagaacca ctgcttactg gcttatcgaa attaatacga ctcactatag 2520
ggagacccaa gctttaagct taccatgggg gttctcacc atcatcatca tcatggtatg 2580
gcatgactgg tggacagcaa atgggtcggg atctgtacga cgatgacgat aaggtaccta 2640

aggatcagct tggagttgat cccgtcgttt tacaacgtcg tgaactgggaa aaccctggcg 2700

ttaccaact taatcgctt gcagcacatc ccctttcgc cagctggcgt aatagecgaag 2760

aggcccgcac cgatcgccct tccaacagt tgcgcagcct gaatggcgaa tggcgctttg 2820

cctggtttcc ggcaccagaa gcggtgccgg aaagctggct ggagtgcgat cttcctgagg 2880

ccgatactgt cgtcgtcccc tcaaacggc agatgcacgg ttacgatgcg cccatctaca 2940

ccaacgtaac ctatccatt acggtcaatc cgcgcttgt tcccacggag aatccgacgg 3000

gtgttactc gtcacattt aatgttgatg aaagctggct acaggaagc cagacgcgaa 3060

ttatTTTTga tggcgtaaC tcggcgtttC atctgtggtg caacgggcgc tgggtcggtt 3120

acggccagga cagtcgtttg ccgtctgaat ttgacctgag cgcatTTTTa cgcgccggag 3180

aaaaccgct cgcggtgatg gtgctcgctt ggagtgcagg cagttatctg gaagatcagg 3240

atatgtggcg gatgagcggc atttccgtg acgtctcgtt gctgcataaa ccgactacac 3300

aaatcagcga tticcatgtt gccactcgt ttaatgatga tttcagccgc gctgtactgg 3360

aggctgaagt tcagatgtgc ggcgagttgc gtgactacct acgggtaaca gtttctttat 3420

ggcaggtga aacgcagtc gccagcggca ccgcgccttt cggcggtgaa attatcgatg 3480

agcgtggtgg ttatgccgat cgcgtcacac tacgtctgaa cgtcgaaaac ccgaaactgt 3540

ggagcggcga aatcccgaat ctctatcgtg cggtggttga actgcacacc gccgacggca 3600

cgctgattga agcagaagcc tgcgatgtcg gtttccgca ggtgcggatt gaaaatggtc 3660

tgtgtctgt gaacggcaag ccgttctga ttcgagcgt taaccgtcac gagcatcatc 3720

ctctgatgg tcaggtcatg gatgagcaga cgatggtgca ggatatactg ctgatgaagc 3780

agaacaactt taacccgtg ccgtgttcgc attatccgaa ccatccgtg tggtacacgc 3840

tgtgcgaccg ctacggcctg tatgtggtgg atgaagccaa tattgaaacc cacggcatgg 3900
 tgccaatgaa tcgtctgacc gatgatccgc gctggctacc ggcatgagc gaacgcgtaa 3960
 cgcgaatggt gcagcgcgat cgtaatcacc cgagtgtgat catctggtcg ctggggaatg 4020
 aatcaggcca cggcctaata cagcagcgc tgtatcgtg gatcaaatct gtcgatcctt 4080
 cccgcccgtt gcagtatgaa ggcggcggag ccgacaccac ggccaccgat attatttgcc 4140
 cgatgtacgc gcgcgtggat gaagaccagc cttcccggc tgtgccgaaa tggccatca 4200
 aaaaatggct ttcgctacct ggagagacgc gcccgctgat cttttgcgaa tacgcccacg 4260
 cgatgggtaa cagtcttggc ggtttcgcta aatactggca ggcgtttcgt cagtatcccc 4320
 gtttacaggc cgcttcgctc tgggactggg tggatcagtc gctgattaaa tatgatgaaa 4380
 acggcaacce gtgctggct tacggcgtg attttggcga tacgccgaac gatcggcagt 4440
 tctgtatgaa cggctctgctc tttgccgacc gcacgccgca tccagcgtg acggaagcaa 4500
 aacaccagca gcagtttttc cagttccgtt tatccgggca aaccatcgaa gtgaccagcg 4560
 aatacctgtt ccgtcatagc gataacgagc tcctgactg gatggtggcg ctggatggta 4620
 agccgctggc aagcggtgaa gtgcctctgg atgtcgctcc acaaggtaaa cagttgattg 4680
 aactgctga actaccgag ccggagagcg cgggcaact ctggctcaca gtacgcgtag 4740
 tgcaaccgaa cgcgaccgca tggtcagaag cgggcacat cagcgcctgg cagcagtggc 4800
 gtctggcgga aaacctcagt gtgacgctcc ccgccgctc ccacccatc ccgatctga 4860
 ccaccagcga aatggatfff tgcacgagc tgggtaataa gcgttgcaa tttaacgcc 4920
 agtcaggctt tctttcacag atgtggattg gcgataaaaa acaactgctg acgccgctgc 4980

gcgatcagtt caccctgca ccgctggata acgacattgg cgtaagttaa gcgaccgca 5040

ttgaccctaa cgctgggtc gaacgtgga aggcggcggg ccattaccag gccgaagcag 5100

cgttgttga gtgcacggca gatacacttg ctgatgcggt gctgattacg accgctcacg 5160

cgtggcagca tcaggggaaa accttattta tcagccggaa aacctaccgg attgatgta 5220

gtggtcaaat ggcgattacc gttgatgttg aagtggcgag cgatacccg catccggcgc 5280

ggattggcct gaactgccag ctggcgcagg tagcagagcg ggtaactgg ctcgattag 5340

ggccgcaaga aaactatccc gaccgctta ctgccgctg ttttgaccgc tgggatctgc 5400

cattgtcaga catgtatacc ccgtacgtct tcccagcga aaacggtctg cgctgcggga 5460

cgcgcaatt gaattatggc ccacaccagt ggcgcgcgga cttccagttc aacatcagcc 5520

gctacagtca acagcaactg atggaaacca gccatcgcca tctgctgcac gcggaagaag 5580

gcacatggct gaatatcgac ggtttccata tggggattgg tggcgacgac tcttgagcc 5640

cgtcagtatc ggcggaatta cagctgagcg ccggtcgcta ccattaccag ttggtctggt 5700

gtcaaaaaata ataaagccga attctgcaga tatccagcac agtggcgcc gctagcacia 5760

aaataaaaaa agaaagggtg actgtgagac atgggctaaa gaggagcggc cgctcgagtc 5820

tagaactagt ggatcagctt tgctgcttgc acttcagagt tctaggagag tcctcctag 5880

tctctctct cagaggagt accgagacct caaataaag gattgattgc ctactgccg 5940

agtggagagt gattactgag cggccggtgt atcgggagtc gtccttaat ctgtcaata 6000

ccagagcggc tctgcagcc gacctcgagg gggggcccta ttctatagt tcacctaat 6060

gctagagctc gctgatcagc ctgactgtg cttctagtt gccagccatc tgttgtttgc 6120

ccctccccg tgccttctt gacctggaa ggtgccactc ccaetgtct ttcctaataa 6180

aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg gggtaggggtg 6240
gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggatgcggtg 6300
ggctctatgg ctctgaggc ggaaagaacc agtggcggta atacggttat ccacagaatc 6360
aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 6420
aaagccgcg ttgctggcgt ttttccatag gctccgccc cctgacgagc atcacaataa 6480
tcgacgtca agtcagaggt ggcgaaacc gacaggacta taaagatacc aggcgtttcc 6540
ccctggaagc tcctctgtc gctctcctgt tccgacctg ccgcttaccg gataacctgc 6600
cgctttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag 6660
ttcgggtgtag gtcgttcgct ccaagctggg ctgtgtgcaac gaacccccg ttcagcccga 6720
ccgctgcgcc ttatccgta actatcgtct tgagtccaac ccgtaagac acgacttate 6780
gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac 6840
agagtcttg aagtggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg 6900
cgctctgctg aagccagta ccttcggaaa aagagtggg agctcttgat cggcaaaaca 6960
aaccaccgct ggtagcggtg gttttttgt ttgcaagcag cagattacgc gcagaaaaaa 7020
aggatctcaa gaagatcctt tgatctttc tacgggtct gacgctcagt ggaacgaaaa 7080
ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt 7140
aaattaaata tgaagtttta aatcaatcta aagtatatat gagnaacctg atcaggactc 7200
ttcttttca tgaacaataa aactgtctgc ttacataaac agtaatacaa ggggtgttat 7260
gagccatatt caacgggaaa cgtcttctc taggccgca ttaaattcca acatggatgc 7320

tgatttatat gggataaat gggctcgcga taatgtcggg caatcaggtg cgacaatcta 7380
 tcgattgtat gggaagcccg atgcgccaga gttgtttctg aaacatggca aaggtagcgt 7440
 tgccaatgat gttacagatg agatggtcag actaaactgg ctgacggaat ttatgcctct 7500
 tccgaccatc aagcatttta tccgiactcc tgatgatgca tggttactca cactgcat 7560
 ccccggaaaa acagcattcc aggtattaga agaatacct gattcaggtg aaaatattgt 7620
 tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt cctgtttgta attgtccttt 7680
 taacagcgat cgcgtatttc gtctcgtca ggcgcaatca cgaatgaata acggtttggt 7740
 tgatgcgagt gatthtgatg acgagcgtaa tggtggcct gttgaacaag tctggaaga 7800
 aatgcataaa cttttgcat tctcaccgga ttcagtcgtc actcatggtg atttctcact 7860
 tgataacctt atttttgacg aggggaaatt aataggttgt attgatgttg gacgagtcgg 7920
 aatcgagac cgataccagg atcttgccat cctatggaac tgcctcggtg agttttctcc 7980
 ttcattacag aaacggcttt ttcaaaaata tggattgat aatcctgata tgaataaatt 8040
 gcagtttcat ttgatgctcg atgagttttt ctaagaattc gcgcaattaa ccctcactaa 8100
 agggaacaaa agctgggtac cgggccc 8127