



등록특허 10-2162118



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년10월07일
(11) 등록번호 10-2162118
(24) 등록일자 2020년09월25일

- (51) 국제특허분류(Int. Cl.)
C07K 14/14 (2006.01) *A61K 39/15* (2006.01)
A61P 31/14 (2006.01) *A61P 37/04* (2006.01)
C12N 15/46 (2006.01) *C12N 15/82* (2006.01)
C12P 21/00 (2006.01)
- (21) 출원번호 10-2014-7034303
- (22) 출원일자(국제) 2013년05월10일
심사청구일자 2018년05월04일
- (85) 번역문제출일자 2014년12월05일
- (65) 공개번호 10-2015-0013713
- (43) 공개일자 2015년02월05일
- (86) 국제출원번호 PCT/CA2013/050364
- (87) 국제공개번호 WO 2013/166609
국제공개일자 2013년11월14일
- (30) 우선권주장
61/646,058 2012년05월11일 미국(US)
- (56) 선행기술조사문헌
KR1020010082143 A*
Sci China Life Sci. 54(1):82-89(2011)*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
메디카고 인코포레이티드
캐나다, 퀘벡 쥐1브이 3브이9, 세인트-포이, 슈트
600, 루트드 레그리즈, 1020
미쓰비시 타나베 파마 코퍼레이션
일본 오사카후 오사카시 쥐오쿠 도쇼마치 3쵸메
2방 10고
- (72) 발명자
다우스트 마르크-안드레
캐나다 퀘벡 쥐1엑스 4엔4 퀘벡 뿐 루이즈-피세
3925
란드리 나탈리에
캐나다 퀘벡 쥐6제트 1케이1 세인트-쟝-크리소스
톰 뿐 드 라 브와 웨스트 1060
(뒷면에 계속)

- (74) 대리인
특허법인와이에스장

전체 청구항 수 : 총 24 항

심사관 : 정지혜

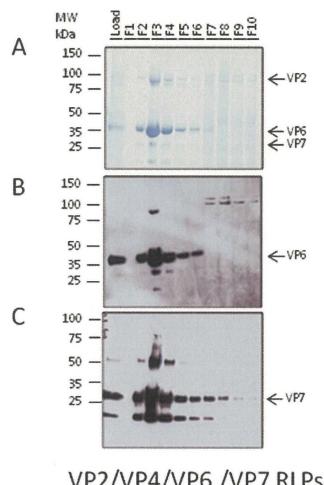
- (54) 발명의 명칭 식물에서 로타바이러스-유사 입자 생산

(57) 요 약

식물에서 바이러스-유사 입자(VLP)를 생산하는 방법이 제공된다. 방법은 제1 핵산을 식물, 또는 식물의 일부로 도입하는 단계를 포함한다. 하나 이상의 로타바이러스 구조 단백질, 제한은 아니지만, 예를 들어, 로타바이러스 단백질 VP2를 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제1 조절 영역을 포함하

(뒷면에 계속)

대 표 도



VP2/VP4/VP6 /VP7 RLPs

는 제1 핵산, 뉴클레오티드 서열은 하나 또는 그 이상 증폭 요소 및/또는 구획 표적화 서열을 더 포함할 수도 있다. 제2 핵산은 식물, 또는 식물의 일부로 도입될 수도 있다. 제2 핵산은 하나 이상의 로타바이러스 구조 단백질, 제한은 아니지만, 예를 들어, 로타바이러스 단백질 VP6을 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제2 조절 영역을 포함한다. 선택적으로, 제3 핵산 및/또는 제4 핵산은 식물, 또는 식물의 일부로 도입될 수도 있다. 제3 핵산은 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질, 제한은 아니지만, 예를 들어, 로타바이러스 단백질 VP4를 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된 제3 조절 영역을 포함한다. 제4 핵산은 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질, 제한은 아니지만, 예를 들어, 로타바이러스 단백질 VP7을 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된 제4 조절 영역을 포함한다. 식물 또는 식물의 일부는 핵산의 발현을 허용하는 조건 하에 배양되고, 이로 인해 VLP를 생산한다.

(72) 발명자

라브와 피에르-올리비에

캐나다 퀘벡 쥐1엔 2엘7 퀘벡 프랭클린 621

아라이 마사아키일본 오사카 541-8505 오사카시 추오쿠 기타하마
2-초메 6반 18고 미쓰비시 타나베 파마 코퍼레이션
내**아사하라 나오미**일본 오사카 541-8505 오사카시 추오쿠 기타하마
2-초메 6반 18고 미쓰비시 타나베 파마 코퍼레이션
내**무텝파 테이비드 레비 루텐도**

영국 셔우드 노팅햄 엔5취 3이와이 보닝턴 크레센트 26

히체로스 잉가 이사벨

남아프리카 공화국 케이프 타운 브레데호크 클리포드 애비뉴 11

리비키 에드워드 피터

남아프리카 공화국 7405 케이프 타운 파인랜즈 위트블루그트 로드 34

명세서

청구범위

청구항 1

식물, 식물의 일부 또는 식물 세포에서 로타바이러스 유사 입자(RLP)를 생산하는 방법으로서,

a) VP2, VP6 및 VP7로부터 선택되는 제1 로타바이러스 구조 단백질을 암호화하는 제1 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제1 조절 영역을 포함하는 제1 핵산,

VP2, VP6 및 VP7로부터 선택되는 제2 로타바이러스 구조 단백질을 암호화하는 제2 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제2 조절 영역을 포함하는 제2 핵산, 및

VP2, VP6 및 VP7로부터 선택되는 제3 로타바이러스 구조 단백질을 암호화하는 제3 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제3 조절 영역을 포함하는 제3 핵산

을 식물, 식물의 일부 또는 식물 세포로 도입하는 단계로서, 여기서 VP7을 암호화하는 제1, 제2 또는 제3 뉴클레오티드 서열은 식물 폴리펩티드로부터의 절단된(truncated) 고유 신호 펩티드 또는 비-고유 신호 펩티드를 포함하고, 여기서 RLP는 VP2, VP6 및 VP7을 포함하는, 단계; 및

b) 제1, 제2 및 제3 핵산의 일과성(一過性) 발현을 허용하는 조건 하에 식물, 식물의 일부 또는 식물 세포를 배양하고, 이로 인해 RLP를 생산하는 단계

를 포함하는 방법.

청구항 2

제1 항에 있어서, 단계 a)에서 식물에서 활성이고 제4 로타바이러스 구조 단백질을 암호화하는 제4 뉴클레오티드 서열에 작동 가능하게 결합된 제4 조절 영역을 포함하는 제4 핵산이 식물, 식물의 일부 또는 식물 세포로 도입되고, 단계 b)에서 제4 로타바이러스 구조 단백질은 식물, 식물의 일부 또는 식물 세포를 배양할 때 발현되며, 여기서 제4 로타바이러스 구조 단백질은 VP4인 것을 특징으로 하는 방법.

청구항 3

제1 항에 있어서, 제1 로타바이러스 구조 단백질은 VP2이고, 제2 로타바이러스 구조 단백질은 VP6이며, 제3 로타바이러스 구조 단백질은 VP7인 것을 특징으로 하는 방법.

청구항 4

제2 항에 있어서, 제1 로타바이러스 구조 단백질은 VP2이고, 제2 로타바이러스 구조 단백질은 VP6이며, 제3 로타바이러스 구조 단백질은 VP7이고, 제4 로타바이러스 구조 단백질은 VP4인 것을 특징으로 하는 방법.

청구항 5

제1 항에 있어서, 제1, 제2 또는 제3 뉴클레오티드 서열 또는 이것들의 조합은 동부콩 모자이크 바이러스(CPMV) 조절 영역에 작동 가능하게 결합되는 것을 특징으로 하는 방법.

청구항 6

제2 항에 있어서, 제1, 제2, 제3 또는 제4 뉴클레오티드 서열 또는 이것들의 조합은 동부콩 모자이크 바이러스(CPMV) 조절 영역에 작동 가능하게 결합되는 것을 특징으로 하는 방법.

청구항 7

제4 항의 방법에 의해 생산된 RLP로서, RLP는 70-100 nm의 크기이고 식물-특이적 N-글리칸, 또는 변형된 N-글리칸을 포함하는 하나 이상의 로타바이러스 구조 단백질을 포함하는 RLP.

청구항 8

제7 항의 RLP의 유효량 및 약학적으로 허용 가능한 담체를 포함하는, 대상에서의 면역 반응 유발용 조성물.

청구항 9

제3 항에 있어서, VP2를 암호화하는 뉴클레오티드 서열은 SEQ ID NO:13, SEQ ID NO:14, 또는 SEQ ID NO:45에 의해 정의된 뉴클레오티드 서열을 포함하고, VP6을 암호화하는 뉴클레오티드 서열은 SEQ ID NO:17, SEQ ID NO:18 또는 SEQ ID NO:46에 의해 정의된 뉴클레오티드 서열을 포함하고, VP7을 암호화하는 뉴클레오티드 서열은 SEQ ID NO:19, 20, 48, 49, 52, 53, 54 또는 57에 의해 정의된 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 방법.

청구항 10

제4 항에 있어서, VP2를 암호화하는 뉴클레오티드 서열은 SEQ ID NO:13, SEQ ID NO:14, 또는 SEQ ID NO:45에 의해 정의된 뉴클레오티드 서열을 포함하고, VP6을 암호화하는 뉴클레오티드 서열은 SEQ ID NO:17, SEQ ID NO:18 또는 SEQ ID NO:46에 의해 정의된 뉴클레오티드 서열을 포함하고, VP7을 암호화하는 뉴클레오티드 서열은 SEQ ID NO:19, 20, 48, 49, 52, 53, 54 또는 57에 의해 정의된 뉴클레오티드 서열을 포함하고, VP4는 SEQ ID NO:15, 16, 47, 50, 또는 51에 의해 정의된 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 방법.

청구항 11

제1 항에 있어서, 제1, 제2 및 제3 핵산의 비율이 1:1:1인 혼합물을 포함하는 아그로박테리아의 접종으로 아그로-침투(agro-infiltration)하여 제1, 제2 및 제3 핵산이 식물, 식물의 일부 또는 식물 세포로 도입되고, 제1, 제2 및 제3 핵산은 식물에서 일과성(一過性)으로 발현되며, 식물은 니코티아나 벤타미아나(*Nicotiana benthamiana*)인 것을 특징으로 하는 방법.

청구항 12

제2 항에 있어서, 제1, 제2, 제3 및 제4 핵산의 비율이 1:1:1:1인 혼합물을 포함하는 아그로박테리아의 접종으로 아그로-침투(agro-infiltration)하여 제1, 제2, 제3 및 제4 핵산이 식물, 식물의 일부 또는 식물 세포로 도입되고, 제1, 제2, 제3 및 제4 핵산은 식물에서 일과성(一過性)으로 발현되며, 식물은 니코티아나 벤타미아나(*Nicotiana benthamiana*)인 것을 특징으로 하는 방법.

청구항 13

제1 항 또는 제2 항에 있어서, 뉴클레오티드 서열의 코돈 사용은 인간 코돈 사용, 증가된 GC 함량 또는 이것들의 조합에 대해 최적화되어 있는 것을 특징으로 하는 방법.

청구항 14

제1 항에 있어서, 상기 제1, 제2 및 제3 뉴클레오티드 서열은 전사 인핸서 요소에 작동가능하게 연결되는 것을 특징으로 하는 방법.

청구항 15

제1 항에 있어서, 비-고유 신호 펩티드는 알팔파(alfalfa) 단백질 이황화 이소머라제(PDI) 신호 펩티드인 것을 특징으로 하는 방법.

청구항 16

제1 항에 있어서,

- c) 식물, 식물의 일부 또는 식물 세포를 수확하는 단계, 및
 - d) 식물, 식물의 일부 또는 식물 세포로부터 칼슘의 존재 하에 RLP를 추출하고 정제하는 단계
- 를 더 포함하는 것을 특징으로 하는 방법.

청구항 17

식물, 식물의 일부 또는 식물 세포에서 로타바이러스 유사 입자(RLP)를 생산하는 방법으로서,

a) VP2, VP6 및 VP7로부터 선택되는 제1 로타바이러스 구조 단백질을 암호화하는 제1 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제1 조절 영역을 포함하는 제1 핵산,

VP2, VP6 및 VP7로부터 선택되는 제2 로타바이러스 구조 단백질을 암호화하는 제2 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제2 조절 영역을 포함하는 제2 핵산, 및

VP2, VP6 및 VP7로부터 선택되는 제3 로타바이러스 구조 단백질을 암호화하는 제3 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제3 조절 영역을 포함하는 제3 핵산

을 포함하는 식물, 식물의 일부 또는 식물 세포를 제공하는 단계로서, 여기서 VP7을 암호화하는 제1, 제2 또는 제3 뉴클레오티드 서열은 식물 폴리펩티드로부터의 절단된(truncated) 고유 신호 펩티드 또는 비-고유 신호 펩티드를 포함하고, 여기서 RLP는 VP2, VP6 및 VP7을 포함하는, 단계; 및

b) 제1, 제2 및 제3 핵산의 일과성(一過性) 발현을 허용하는 조건 하에 식물, 식물의 일부 또는 식물 세포를 배양하고, 이로 인해 RLP를 생산하는 단계

를 포함하는 방법.

청구항 18

제17 항에 있어서, 단계 a)에서 식물에서 활성이고 제4 로타바이러스 구조 단백질을 암호화하는 제4 뉴클레오티드 서열에 작동 가능하게 결합된 제4 조절 영역을 포함하는 제4 핵산이 식물, 식물의 일부 또는 식물 세포로 도입되고, 단계 b)에서 제4 로타바이러스 구조 단백질은 식물, 식물의 일부 또는 식물 세포를 배양할 때 발현되며, 여기서 제4 로타바이러스 구조 단백질은 VP4인 것을 특징으로 하는 방법.

청구항 19

제17 항의 방법에 의해 생산된 RLP로서, RLP는 70-100 nm의 크기이고 식물-특이적 N-글리칸, 또는 변형된 N-글리칸을 포함하는 하나 이상의 로타바이러스 구조 단백질을 포함하는 RLP.

청구항 20

제17 항의 RLP의 유효량 및 약학적으로 허용 가능한 담체를 포함하는, 대상에서의 면역 반응 유발용 조성물.

청구항 21

제17 항에 있어서, 뉴클레오티드 서열의 코돈 사용은 인간 코돈 사용, 증가된 GC 함량 또는 이것들의 조합에 대해 최적화되어 있는 것을 특징으로 하는 방법.

청구항 22

제17 항에 있어서, 상기 제1, 제2 및 제3 뉴클레오티드 서열은 전사 인핸서 요소에 작동 가능하게 연결되는 것을 특징으로 하는 방법.

청구항 23

제17 항에 있어서, 비-고유 신호 펩티드는 알파파(alphafalfa) 단백질 이황화 이소머라제(PDI) 신호 펩티드인 것을 특징으로 하는 방법.

청구항 24

제17 항에 있어서,

c) 식물, 식물의 일부 또는 식물 세포를 수확하는 단계, 및

d) 식물, 식물의 일부 또는 식물 세포로부터 칼슘의 존재 하에 RLP를 추출하고 정제하는 단계

를 더 포함하는 것을 특징으로 하는 방법.

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

발명의 설명

기술 분야

[0001]

본 발명은 식물에서 로타바이러스(rotavirus) 구조 단백질의 생산에 관한 것이다. 더 구체적으로, 본 발명은 식물에서 로타바이러스 구조 단백질을 포함하는 바이러스-유사 입자의 생산에 관한 것이다.

배경 기술

[0002]

로타바이러스 감염은 5살 미만의 어린이들에게 주로 영향을 미치는 세계적인 문제이다. 그것은 심각한 위장염(gastroenteritis) 및 최악의 경우에는 사망을 초래한다.

[0003]

로타바이러스는 위장관계(gastrointestinal system) 및 기도에 영향을 미치는 바이러스의 레오비리대(Reoviridae) 과의 멤버 (로타바이러스 속)이다. 명칭은 역상 대비 전자 현미경 검사(negative contrast electron microscopy) (도 1a; 선행 기술)로 볼 때 비리온의 바퀴와 유사한 모양으로부터 유래된다. 로타바이러스는 보통 구의 형태이고 같은 것의 외부 및 내부 쉘(shell) 또는 이중-쉘 캡시드(capsid) 구조에 따라 명명된다. 각각 외부 캡시드는 지름이 약 70 nm이고, 내부 캡시드는 약 55 nm이다. 로타바이러스의 이중-쉘 캡시드는 내부 단백질 쉘 및 게놈을 포함하는 코어를 둘러싸고 있다. 로타바이러스의 게놈은 적어도 11개의 로타바이러스 단백질을 암호화하는 이중 가닥 RNA 세그먼트로 구성된다.

[0004]

dsRNA는 여섯 개의 구조 단백질 (VP) 및 여섯 개의 비-구조 단백질 (NSP)을 암호화한다 (도 1c; 선행 기술). 구조 단백질은 VP1, VP2, VP3, VP4, VP6 및 VP7을 포함한다 (도 1b; 선행 기술). 세 개의 동심원 층은 각각 VP2, VP6 및 VP7의 조립에 의해 형성되며, VP4는 바이러스 구조의 표면에 "스파이크(spike)"를 형성한다. NSP는 감염된 세포에서 합성되고 복제 주기의 다양한 부분에서 기능하거나 숙주 단백질의 일부와 상호작용하여 감염에 대한 발병 또는 면역 반응에 영향을 미친다 (Greenberg and Estes, 2009).

- [0005] VP2는 102 kDa 단백질이고 바이러스 코어의 가장 풍부한 단백질이다. 그것은 가장 안쪽의 구조 단백질 층을 형성하고 바이러스 코어의 구성요소 및 전사 효소의 올바른 조립을 위한 스캐폴드(scaffold)를 제공한다 (Lawton, 2000). VP1, 125 kDa로 가장 큰 바이러스 단백질은 로타바이러스에 대한 RNA-의존적 폴리머라제로서 작용하며, 코어 복제 중간물을 생성하고, 그것의 20면체의 꼭지점에서 VP2와 결합한다 (Varani and Allain, 2002; Vende et al., 2002). VP3, 98 kDa 단백질은 또한 바이러스 계놈에 직접적으로 결합되며, 바이러스 mRNA에 5' 캡(cap) 구조를 추가하는 mRNA 캡핑(capping) 효소로서 작용한다. 더불어, VP1 및 VP3은 VP2 캡시드 층의 외부 5배의 꼭지점에 부착되는 복합체를 형성한다 (Ange1, 2007). VP6은 바이러스 코어의 중간 쉘을 형성하는 42 kDa 단백질이고, 주요 캡시드 단백질이며, 비리온의 총 단백질 질량의 50% 이상을 차지한다 (Gonzalez et al., 2004; Estes, 1996). 그것은 유전자 전사를 필요로 하고 코어에서 VP1을 VP2에 고정함으로써 로타바이러스 RNA의 캡슐화에서 역할을 할 수도 있는데, 블루텅 바이러스(bluetongue virus), 레오비리대 과의 또 다른 멤버에서 보이는 바와 같다. 그것은 또한 로타바이러스의 다섯 개의 군 (A 내지 E)으로의 분류를 결정하며, 군 A는 가장 일반적으로 사람에게 영향을 미친다 (Palombo, 1999). 로타바이러스 군 A에서 VP6는 적어도 네 개의 부분군(SG)을 가지며, 이것들은 SG 특이적 예피토프, SG I, SG II, SG (I+II) 및 SG 비-(I+II)의 존재 또는 부재에 의존적이다. 군 B 및 C는 일반적인 군 A 공통 항원이 없지만 사람을 감염시키는 것으로도 알려져 있는 한편, 군 D는 동물, 예를 들어, 닭 및 소에만 영향을 미친다 (Thongprachum, 2010).
- [0006] 두 개의 외부 캡시드 단백질 VP7, 37 kDa 당단백질 (G) 및 87 kDa 프로테아제 민감성 VP4 (P)는 바이러스의 혈청형을 한정한다. 이 두 개의 단백질들은 중화 항체 반응을 유발하고 따라서 로타바이러스 혈청형을 이중 명명 시스템으로 분류하기 위해 사용되며, G-P 항원 조합에 따른다 (예를 들어, G1 P[8] 또는 G2 P[4]) (Sanchez-Padilla et al., 2009). VP4 단백질은 다이머화하여 바이러스의 외부 쉘에서 60개의 스파이크(spike)를 형성하며, 이것은 숙주 세포 진입의 초기 단계에 직접 수반된다. 스파이크 단백질은 아미노산(aa) 위치 248에서 분할 부위를 함유한다. 감염 시, 그것은 프로테아제 트립신에 의해 분할되어 VP5 (529 aa, 60 kDa) 및 VP8 (246 aa, 28 kDa)을 생산한다 (Denisova et al., 1999). 이 공정은 바이러스 감염성 (숙주 세포의 세포 부착 및 침입)을 항상시키고 스파이크 구조를 안정화한다 (Glass, 2006). VP7 당단백질은 바이러스의 세 번째 또는 바깥 층을 형성한다. 현재, 27 G 및 35 P 유전자형이 알려져 있다 (Greenberg and Estes, 2009). VP4 및 VP7은 바이러스 중화에 수반된 주요 항원이고 백신 개발에 대한 중요한 표적이다 (Dennehy, 2007).
- [0007] 감염된 포유동물 세포에서, 로타바이러스는 독특한 방식의 형태 형성(morphogenesis)을 경험하여 완전한 3층화된 VP2/6/4/7 바이러스 입자를 형성한다 (Lopez et al., 2005). 3층 캡시드는 소화기를 통한 (faecal-oral) 전파 및 바이러스의, 응모의 끝 근처에서 비-분할 분화된 장세포를 감염시키는 소장으로의 전달을 가능하게 하는 매우 안정한 복합체이다 (Greenberg and Estes, 2009). 처음에, 온전한 바이러스는 바이러스의 표면에서 60개의 VP4 다이머 스파이크를 통해 시알산-독립적 수용체에 부착된다 (Lundgren and Svensson, 2001). 바이러스의 표면에서 60개의 VP4 다이머 스파이크는 바이러스가 이 세포 수용체들에 부착되게 한다. VP4는 트립신에 의한 단백질 가수분해 분할에 민감하며 일련의 공동-수용체와의 상호작용을 위해 당단백질의 표면에서 추가적인 부착 부위를 노출시키는 형태적 변화를 일으킨다.
- [0008] 하지만, 다단계 부착 및 진입 공정이 분명하게 이해되어 있지는 않지만, 바이러스는 숙주의 원형질막을 거쳐 전달된다. 진입 공정에 또한 수반되는 VP7 외부 캡시드 쉘은 공정에서 제거되고 소낭에서 이중층 입자 (DLP)는 세포질로 전달된다 (도 2; 선행 기술). DLP는 소낭에서 빠져나와서 비-막 결합된 세포질 봉입체(cytoplasmic inclusion)로 들어간다. VP1에 의한 계놈의 초기 전사는 dsRNA가 절대 세포질에 노출되지 않도록 입자에서 시작한다. RNA 복제 및 코어 형성은 이 비-막-결합된 세포질 봉입체들에서 일어난다. 그 다음 발생기 (+) RNA는 세포질로 수송되고 바이러스 단백질 합성을 위한 주형의 역할을 한다. VP4는 시토졸에서 생산되고 조면 소포체(RER)로 수송되며, VP7은 RER로 분비된다. VP2 및 VP6이 생산되고 비로솜(virosome)의 시토졸에서 조립되고 그 후에 RER 구획으로 버깅되며(budding), 공정에서 일시적 막 외피를 받는다 (Lopez et al., 2005; Tian et al., 1996). RER에서, 바이러스 입자의 일시적 외피가 제거되고 VP4 및 VP7 단백질 모노머에 의해 대체되며, 로타바이러스 당단백질 NSP4의 수반이 중요하다 (Tian et al., 1996; Lopez et al., 2005; Gonzalez et al., 2000). NSP4는 ER 막에서 세포 내 수용체로서 기능하고 새롭게 만들어진 바이러스 일부 입자 및 아마 스파이크 단백질 VP4에도 결합한다 (Tian et al., 1996). NSP4는 또한 사람에게 유독성이고 설사(diarrhea)의 원인 인자이다. 그 후에 완벽한, 성숙한 입자는 분비를 위해 RER에서 골지체를 통하여 원형질막으로 전송된다 (Lopez et al., 2005).
- [0009] 다양한 다른 접근법이 다양한 혈청형의 로타바이러스로부터 인류를 보호하는데 적합한 로타바이러스 백신을 생성하기 위해 취해졌다. 이 접근법들은 다양한 제네리안 접근법 (Jennerian approach), 살아있는 약화된 바이러

스의 사용, 면역원으로서 바이러스-유사 입자, 핵산 백신 및 바이러스 서브유닛의 사용을 포함한다. 현재 시중에서 이용 가능한 두 개의 경구용 백신이 있지만, 이것들은 균주 변형 및 다른 병원체의 존재로 인해 일부 개발도상국에서는 낮은 효능을 갖는다.

[0010] 미국 특허 번호 제4,624,850호, 제4,636,385호, 제4,704,275호, 제4,751,080호, 제4,927,628호, 제5,474,773호, 및 제5,695,767호, 각각은 다양한 로타바이러스 백신 및/또는 같은 것의 제조 방법을 설명한다. 이 군의 멤버들에 의해 공유된 공통점은 이 백신들 각각이 궁극적인 로타바이러스 백신을 생성하기 위해 전체 바이러스 입자의 사용에 따른다는 것이다. 효과적인 다원자가 백신에 대한 오래된 필요성을 고려해 볼 때, 이 작품은 이러한 백신에 대한 필요성을 해결하는데 있어서 단지 부분적으로만 성공적이었다는 것은 분명하다.

[0011] 백신 생성의 전통적인 방법에서 벗어나, 분자 생물학 분야의 발전은 개개의 로타바이러스 단백질의 발현을 허용하였다. Crawford et al. (J Virol. 1994 September; 68(9): 5945-5952)은 주요 캡시드 단백질을 암호화하는 VP2, VP4, VP6, 및 VP7을 바클로바이러스(baculovirus) 발현 시스템으로 클로닝하였고 곤충 세포에서 각 단백질을 발현시켰다. 다른 조합의 로타바이러스 주요 구조 단백질의 동시 발현은 안정한 바이러스-유사 입자 (VLP)의 형성을 일으켰다. VP2 및 VP6의 동시 발현은 단독으로 또는 VP4와 함께 VP2/6 또는 VP2/4/6 VLP의 생산을 일으키며, 이것들은 이중층 로타바이러스 입자와 유사했다. VP4와 함께 또는 없이, VP2, VP6, 및 VP7의 동시 발현은 삼중층 VP2/6/7 또는 VP2/4/6/7 VLP를 생산하였으며, 이것들은 고유한 감염성 로타바이러스 입자와 유사하였다. VLP는 고유한 입자의 구조적 및 기능적 특성을 유지하였으며, 입자의 전자 현미경 검사, VP4 및 VP7에서 비-중화 및 중화 에피토프, 및 VP2/4/6/7 VLP의 해마글루티닌화 활성에 의해 결정된 바와 같다.

[0012] 다른 단백질 조성물의 바이러스-유사 입자로부터 생성된 백신 후보물질은 서브유닛 백신으로서의 가능성을 나타냈다. O'Neal et al. ("Rotavirus Virus-like Particles Administered Mucosally Induce Protective Immunity," J. Virology, 71(11):8707-8717 (1997))은 VP 2 및 6 또는 VP 2, 6, 및 7을 함유하는 VLP가 콜레라(cholera) 독소의 추가와 함께 및 없이 마우스에 투여될 때 면역화된 마우스에서 보호 면역을 유발했지만, 보호는 VLP가 콜레라 독소 (CT)와 함께 투여될 때 더 효과적이었다는 것을 나타냈다.

[0013] 코어-유사 입자 (CLP) 및 VLP는 또한 소를 면역화하기 위해 사용되었다. Fernandez, et al., ("Passive Immunity to Bovine Rotavirus in Newborn Calves Fed Colostrum Supplements From Cows Immunized with Recombinant SA11 Rotavirus core-like particle (CLP) 또는 virus-like particle (VLP) vaccines," Vaccine, 16(5):507-516 (1998)). 이 연구에서, 수동 면역을 생성하는 CLP 및 VLP의 능력이 연구되었다. 이 군은 VLP가 수동 면역을 유발하는데 있어서 CLP보다 더 효과적이었다고 결론 내렸다.

[0014] 식물은 재조합 단백질의 대규모 생산에 점점 더 많이 사용된다. 예를 들어, 제US 2003/0175303호는 안정한, 형질전환된 토마토 식물에서 재조합 로타바이러스 구조 단백질 VP6, VP2, VP4 또는 VP7의 발현을 개시한다.

[0015] Saldana et al.은 꽃양배추 모자이크 바이러스 (cauliflower mosaic virus; CaMV) 35S 프로모터 및 재조합 아그로박테리움 투메파시엔스(*A. tumefaciens*)를 사용하여 토마토 식물의 세포질에서 VP2 및 VP6을 발현시켰다 (Saldana et al., 2006). 전자 현미경 연구는 적은 비율의 입자가 2/6 VLP로 조립되었다는 것을 나타냈다. 보호 면역 반응이 마우스에서 검출되었고 이것은 어느 정도로는 비-조립된 VP에 의해 기여될 수도 있었다. 개개의 단백질은 VP8 및 VP6의 경우에서와 마찬가지로 마우스의 면역 반응을 유도하는 것으로 나타났다 (Zhou et al., 2010).

[0016] Matsumura et al., (2002)는 처음으로 유전자 이식 토마토 식물에서 소 로타바이러스 A VP6 발현 및 조립을 보고하였다. 그들의 연구에서, 그들은 꽃양배추 모자이크 바이러스 (CaMV) 35S 프로모터 및 VP6 유전자를 운반하는 재조합 아그로박테리움 투메파시엔스(*Agrobacterium tumefaciens*)에 의해 조절되는 유전자 이식 토마토 식물을 사용하였다. 단백질은 발현되었고, 정제되었으며 면역원 연구가 수행되었다. 성체 마우스의 면역-반응은 혈청에서 VP6 항체의 존재를 나타냈다. 하지만, 그것들은 조립된 VP6 단백질의 증거를 나타내지는 않았다. 그것은 마우스에서 면역 반응을 유도할 수 있는 단순한 모노머 또는 트라이머였을 수도 있다. 또 다른 군의 작업은 감자 바이러스 X (PVX) 백터를 사용하는 니코티아나 벤탄미아나(*Nicotiana benthamiana*)의 VP6 조립을 나타냈다 (O'Brien et al., 2000). VP6 단백질이 식물에서 발현될 때, 그것은 PVX 단백질 막대에 융합될 때만 조립된다는 것이 발견되었다. 분할이 발생하면, VP6은 Marusic et al., (2001)에 의한 HIV-PVX의 유사한 연구에서 보이는 바와 같이 20면체 VLP로 조립되었다. 이 결과는 아마 로타바이러스 단백질이 VLP를 형성하기 위해 추가적 요인 또는 향상이 필요할 수도 있다고 제안한다.

[0017] VLP의 생산은, 하나 이상의 재조합 단백질의 합성 및 조립 둘 다가 필요하기 때문에, 도전적인 과제이다. 이것

은 네 개의 다른 단백질의 1860개의 모노머에 의해 형성된 캡시드를 가진 RNA 바이러스인 로타바이러스의 VLP에 대한 경우이다. VLP 생산을 위해 두 개 내지 세 개의 재조합 단백질의 동시 발현 및 조립이 필요하다. 이것들은 120개 분자의 VP2 (내층), 780개 분자의 VP6 (중간층) 및 780개 분자의 당단백질 VP7 (외층)을 포함하며, 궁극적으로는 이중층 또는 삼중층 입자를 형성한다. 게다가, 대부분의 VLP의 생산은 여러 재조합 단백질의 동시 발현 및 조립을 필요로 하며, 이것은 - 로타바이러스 유사 입자 (RLP)의 경우에 - 단일 숙주 세포에서 발생하는 것을 필요로 한다.

[0018] 더 최근의 연구는 비트 블랙 스코치 바이러스 (Beet black scorch virus; BBSV) 매개된 발현 시스템을 사용하여 체노포듐 아마란티컬러(Chenopodium amaranticolor)에서 코돈(codon)-최적화된 인간 로타바이러스 VP6의 성공적인 발현을 보여주었다. 단백질은 BBSV의 코딩 단백질 오픈 리딩 프레임(open reading frame)에 대한 대체로서 설계되었다. 암컷 BALB/c 마우스의 식물 기반 VP6 단백질로의 경구 면역화는 항-VP6 점막 IgA 및 혈청 IgG의 높은 역ガ를 유발하였다 (Zhou et al., 2010). 하지만, 상기 군은 VP6 단백질이 VLP 또는 입자로 조립되었는지를 언급하지는 않았다.

[0019] 로타바이러스 VP7은 또한 담배에서 성공적으로 발현되었고 마우스에서 그것의 중화 면역 반응을 유지하는 것으로 나타났다 (Yu and Langridge, 2001). VP7을 발현하기 위해 유전자 이식 토마토 식물을 사용하는 또 다른 연구는 VP7 유전자가 형질전환된 식물에서 50 세대에 걸쳐 안정하다는 것을 나타냈다. 50번째 세대의 VP7 단백질은 성체 마우스에서 보호 및 중화 항체 둘 다를 유발하였다 (Li et al., 2006).

[0020] Yang et al. (Yang Y M, Li X, Yang H, et al. 2011)은 담배에서 군 A RV (P[8]G1)의 세 개의 로타바이러스 캡시드 단백질 VP2, VP6 및 VP7을 동시에 발현시켰고 이 단백질의 발현 수준, 뿐만 아니라 로타바이러스-유사 입자의 형성 및 면역성이 연구되었다. VLP는 유전자 이식 담배로부터 정제되었고 전자 현미경 검사 및 웨스턴 븍로트(Western blot)에 의해 분석되었다. Yang et al.의 결과는 식물 유래된 VP2, VP6 및 VP7 단백질이 지름이 60-80 nm인 2/6 또는 2/6/7 로타바이러스 유사 입자로 자가-조립(self-assemble) 했다는 것을 나타낸다.

발명의 내용

[0021] 본 발명은 식물에서 로타바이러스 구조 단백질의 생산에 관한 것이다. 더 구체적으로, 본 발명은 또한 식물에서 로타바이러스 구조 단백질을 포함하는 바이러스-유사 입자의 생산에 관한 것이다.

[0022] 본 발명에 따르면 다음을 포함하는 식물에서 로타바이러스-유사 입자 (LP)를 생산하는 방법 (A)가 제공된다:

[0023] a) 제1 로타바이러스 구조 단백질을 암호화하는 제1 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제1 조절 영역을 포함하는 제1 핵산, 제2 로타바이러스 구조 단백질을 암호화하는 제2 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제2 조절 영역을 포함하는 제2 핵산 및 제3 로타바이러스 구조 단백질을 암호화하는 제3 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제3 조절 영역을 포함하는 제3 핵산을 식물, 식물의 일부 또는 식물 세포로 도입하는 단계,

[0024] b) 제1, 제2 및 제3 핵산의 일시적 발현을 허용하는 조건 하에 식물, 식물의 일부 또는 식물 세포를 배양하고, 이로 인해 RLP를 생산하는 단계.

[0025] 게다가 식물에서 활성이고 제4 로타바이러스 구조 단백질을 암호화하는 제4 뉴클레오티드 서열에 작동 가능하게 결합된 제4 조절 영역을 포함하는 제4 핵산은 단계 a)에서 식물, 식물의 일부 또는 식물 세포로 도입될 수도 있고, 단계 b)에서 식물, 식물의 일부 또는 식물 세포를 배양할 때 발현된다.

[0026] 상기 설명된 방법 (A)에서 제1 로타바이러스 구조 단백질은 VP2일 수도 있고, 제2 로타바이러스 구조 단백질은 VP6일 수도 있으며 제3 로타바이러스 구조 단백질은 VP4 또는 VP7일 수도 있다. 게다가, 제4 로타바이러스 구조 단백질은 VP7 또는 VP4일 수도 있다. VP4는 VP5 및 VP8을 생산하도록 가공되거나 분할될 수도 있다. VP4의 분할은 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 또는 서브틸리신을 사용하여 수행될 수도 있다. 프로테아제는 식물 내에서 동시에 발현될 수도 있다.

[0027] 본 발명은 또한 다음을 포함하는 로타바이러스-유사 입자 (LP)를 생산하는 방법 (B)를 제공한다:

[0028] a) 하나 이상의 로타바이러스 구조 단백질을 암호화하는 제1 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 조절 영역을 포함하는 핵산을 식물, 식물의 일부 또는 식물 세포로 도입하는 단계,

[0029] b) 제1 핵산의 일과성 발현을 허용하는 조건 하에 식물, 식물의 일부 또는 식물 세포를 배양하고, 이로 인해 RLP를 생산하는 단계.

- [0030] 상기 설명된 방법 (B)는 단계 a)에서 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질을 암호화하는 제2 뉴클레오티드 서열에 작동 가능하게 결합된 제2 조절 영역을 포함하는 제2 핵산을 도입하는 단계 및 단계 b)에서 식물, 식물의 일부 또는 식물 세포를 배양할 때 제2 핵산을 발현하는 단계를 더 포함할 수도 있다.
- [0031] 상기 설명된 방법 (B)는 단계 a)에서 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질을 암호화하는 제3 뉴클레오티드 서열에 작동 가능하게 결합된 제3 조절 영역을 포함하는 제3 핵산을 식물, 식물의 일부 또는 식물 세포로 도입하는 단계 및 단계 b)에서 식물, 식물의 일부 또는 식물 세포를 배양할 때 제3 핵산을 발현하는 단계를 더 포함할 수도 있
- [0032] 게다가, 방법 (A) 또는 (B)에서 추가 핵산이 식물, 식물의 일부 또는 식물 세포에서 발현될 수도 있으며, 여기에서 추가 핵산은 침묵의 억제인자를 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 조절 영역을 포함한다.
- [0033] 뉴클레오티드 서열의 코돈 사용은 선호하는 인간 코돈 사용, 증가된 GC 함량 또는 이것들의 조합으로 조정될 수도 있다.
- [0034] 로타바이러스 구조 단백질은 절단된, 고유 또는 비-고유 신호 펩티드를 포함할 수도 있다. 비-고유 신호 펩티드는 단백질 이황화 이소머라제 신호 (PDI) 펩티드일 수도 있다.
- [0035] 제1, 제2, 제3 또는 제4 뉴클레오티드 서열 또는 이것들의 조합은 동부콩 모자이크 바이러스 (Cowpea Mosaic Virus; CPMV) 조절 영역에 작동 가능하게 결합될 수도 있다.
- [0036] 상기 설명된 방법 (A) 또는 (B)는 다음 단계를 더 포함할 수도 있다:
- c) 식물, 식물의 일부 또는 식물 세포를 수확하는 단계, 및
 - d) 식물, 식물의 일부 또는 식물 세포로부터 RLP를 정제하는 단계.
- [0037] 방법 (A) 또는 (B)에서 수확 단계 또는 정제 단계 중에, VP4는 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 사용하여 VP5 및 VP8을 생산하도록 가공되거나 분할될 수도 있다.
- [0038] RLP는 크기가 70-100 nm의 범위에 있을 수도 있고 칼슘의 존재시 정제될 수도 있다.
- [0039] 본 발명은, 상기 설명된 바와 같이 방법 (A) 또는 (B)에 의해 생산된 RLP를 제공한다. 생산된 RLP는 적어도 VP4 로타바이러스 구조 단백질을 포함할 수도 있다. VP4는 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 사용하여 VP5 및 VP8로 분할될 수도 있다. 프로테아제는 식물 내에서 동시-발현되거나 수확 단계, 정제 단계, 또는 둘 다의 동안에 추가될 수도 있다. 게다가 방법 (A) 또는 (B)에 의해 생산된 RLP는 이중층 RLP 및/또는 삼중층 RLP일 수도 있다.
- [0040] 게다가, 뉴클레오티드 서열이 제공된다. VP2를 암호화하는 뉴클레오티드 서열은 SEQ ID NO:13, SEQ ID NO:14, 또는 SEQ ID NO:45에 의해 정의된 뉴클레오티드 서열과 80% 내지 100% 동일성을 포함할 수도 있다. VP6을 암호화하는 뉴클레오티드 서열은 SEQ ID NO:17, SEQ ID NO:18 또는 SEQ ID NO:46에 의해 정의된 뉴클레오티드 서열과 80% 내지 100% 동일성을 포함할 수도 있다. VP7을 암호화하는 뉴클레오티드 서열은 SEQ ID NO:19, 20, 48, 49, 52, 53, 54 또는 57에 의해 정의된 뉴클레오티드 서열과 80% 내지 100% 동일성을 포함할 수도 있다. 그리고 VP4를 암호화하는 뉴클레오티드 서열은 SEQ ID NO:15, 16, 47, 50, 또는 51에 의해 정의된 뉴클레오티드 서열과 80% 내지 100% 동일성을 포함할 수도 있다. 게다가, VP2는 SEQ ID NO:1 또는 SEQ ID NO:25에 의해 정의된 아미노산 서열과 80% 내지 100% 동일성을 포함하는 아미노산 서열에 의해 암호화될 수도 있다. VP6은 SEQ ID NO:3 또는 SEQ ID NO:31에 의해 정의된 아미노산 서열과 80% 내지 100% 동일성을 포함하는 아미노산 서열에 의해 암호화될 수도 있다. VP7은 SEQ ID NO:4, 39, 43 또는 59에 의해 정의된 아미노산 서열과 80% 내지 100% 동일성을 포함하는 아미노산 서열에 의해 암호화될 수도 있다. VP4는 SEQ ID NO:2 또는 SEQ ID NO:36에 의해 정의된 아미노산 서열과 80% 내지 100% 동일성을 포함하는 아미노산 서열에 의해 암호화될 수도 있다. 하나 이상의 로타바이러스 구조 단백질은 VP2, VP4, VP6 및/또는 VP7일 수도 있다. VP4는 VP5 및 VP8로 가공될 수도 있다. 하나 이상의 로타바이러스 구조 단백질은 로타바이러스 균주 G9 P[6], 로타바이러스 A WA 균주, 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주, 및 로타바이러스 SA11 균주로부터 선택될 수도 있다.
- [0041] 상기 설명된 방법 (A)에서, 제1, 제2 또는 제3 핵산 서열 또는 이것들의 조합은 하나 또는 그 이상의 코모바이러스(comovirus) 인핸서(enhancer), 하나 또는 그 이상의 증폭 요소 및 로타바이러스 구조 단백질을 암호화하는

뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 조절 영역을 포함할 수도 있으며, 여기에서 레플리카제를 암호화하는 제4 핵산은 식물, 식물의 일부 또는 식물 세포로 도입될 수도 있다.

[0044] 상기 설명된 방법 (B)에서, 제1, 제2, 제3 또는 제4 핵산 서열 또는 이것들의 조합은 하나 또는 그 이상 코모바이러스 인핸서, 하나 또는 그 이상 증폭 요소 및 로타바이러스 구조 단백질을 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 조절 영역을 포함할 수도 있으며, 여기에서 레플리카제를 암호화하는 제5 핵산은 식물, 식물의 일부 또는 식물 세포로 도입될 수도 있다.

[0045] 게다가, 본 발명에 따르면 다음을 포함하는 식물에서 로타바이러스-유사 입자 (RLP)를 생산하는 방법 (C)가 제공된다:

a) 하나 이상의 로타바이러스 구조 단백질을 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 조절 영역을 포함하는 핵산을 식물, 또는 식물의 일부로 도입하는 단계,

b) 제1 핵산의 일과성 발현을 허용하는 조건 하에 식물, 식물의 일부를 배양하고, 이로 인해 RLP를 생산하는 단계.

[0048] 게다가, 제2 핵산은 식물 또는 식물의 일부로 도입될 수도 있고, 제2 핵산은 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질을 암호화하는 제2 뉴클레오티드 서열에 작동 가능하게 결합된 제2 조절 영역을 포함하며 여기에서 제2 핵산은 단계 b)에서 식물 또는 식물의 일부를 배양할 때 발현된다.

[0049] 게다가 제3 뉴클레오티드 서열은 식물 또는 식물의 일부로 도입될 수도 있고, 제3 핵산은 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질을 암호화하는 제3 뉴클레오티드 서열에 작동 가능하게 결합된 제3 조절 영역을 포함하며 여기에서 제3 핵산은 단계 b)에서 식물 또는 식물의 일부를 배양할 때 발현된다.

[0050] 상기 설명된 방법 (C)는 식물의 수학 단계 및 RLP의 추출 단계를 더 포함할 수도 있다.

[0051] 단계 (C)의 하나 이상의 로타바이러스 구조 단백질은 로타바이러스 단백질 VP2, VP4 또는 VP6일 수도 있다. 제1 또는 제2 뉴클레오티드 서열에 의해 암호화된 하나 이상의 로타바이러스 구조 단백질은 VP2 또는 VP6일 수도 있다. 제3 뉴클레오티드 서열에 의해 암호화된 하나 이상의 로타바이러스 구조 단백질은 VP4일 수도 있다. VP4는 VP5 및 VP8을 생산하도록 분할될 수도 있다.

[0052] 제1, 제2 또는 제3 뉴클레오티드 서열은 하나 또는 그 이상 구획 표적화 서열 및/또는 증폭 요소를 더 암호화하고, 포함하거나, 암호화하고 포함할 수도 있다. 하나 이상의 구획 표적화 서열은 하나 이상의 로타바이러스 구조 단백질을 식물 세포의 소포체 (ER), 엽록체, 색소체 또는 아포플라스트(Apoplast)로 보낸다.

[0053] 본 발명은 또한 다음을 포함하는 로타바이러스-유사 입자 (RLP)를 생산하는 방법 (D)를 제공한다,

[0054] a) 하나 이상의 로타바이러스 구조 단백질을 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 조절 영역을 포함하는 핵산을 포함하는 식물 또는 식물의 일부를 제공하는 단계,

[0055] b) 핵산의 일과성 발현을 허용하는 조건 하에 식물, 식물의 일부 또는 식물 세포를 배양하고, 이로 인해 RLP를 생산하는 단계.

[0056] 게다가 방법 (D)의 식물 또는 식물의 일부는

[0057] i) 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질을 암호화하는 제2 뉴클레오티드 서열에 작동 가능하게 결합된 제2 조절 영역을 포함하는 제2 핵산 또는,

[0058] ii) 제2 핵산은 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질을 암호화하는 제2 뉴클레오티드 서열에 작동 가능하게 결합된 제2 조절 영역을 포함하고 제3 핵산은 식물에서 활성이고 하나 이상의 로타바이러스 구조 단백질을 암호화하는 제3 뉴클레오티드 서열에 작동 가능하게 결합된 제3 조절 영역을 포함하는 제2 및 제3 핵산을 더 포함할 수도 있으며:

[0059] 여기에서 제2 또는 제2 및 제3 핵산은 단계 b)에서 식물 또는 식물의 일부를 배양할 때 발현된다.

[0060] 방법 (D)에서 하나 이상의 구조 단백질은 로타바이러스 단백질 VP2, VP4 또는 VP6일 수도 있다. 제1 또는 제2 뉴클레오티드 서열에 의해 암호화된 하나 이상의 로타바이러스 구조 단백질은 VP2 또는 VP6일 수도 있다. 제3 뉴클레오티드 서열에 의해 암호화된 하나 이상의 로타바이러스 구조 단백질은 VP4일 수도 있다. VP4는 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브릴리신을 사용하여 VP5 및 VP8로 분할될 수도 있다. 프로테아제는 식물 내에서 동시-발현되거나 수학, 정제, 또는 둘 다

의 동안에 추가될 수도 있다.

[0061] 본 발명은, 상기 설명된 바와 같이, 방법 (A), 방법 (B), 방법 (C), 방법 (D), 또는 이것들의 조합에 의해 생산된 RLP를 제공한다. RLP는 식물-특이적 N-글리칸, 또는 변형된 N-글리칸을 포함할 수도 있는 하나 이상의 로타바이러스 구조 단백질을 포함할 수도 있다.

[0062] 본 발명은, 상기 설명된 바와 같이, 면역 반응을 유발하기 위해, 방법 (A), 방법 (B), 방법 (C), 방법 (D), 또는 이것들의 조합에 의해 만들어진 RLP의 유효량, 및 약학적으로 허용 가능한 담체를 포함하는 조성물을 포함한다.

[0063] 본 발명은 또한 대상에서 로타바이러스 감염에 대한 면역력을 유발하는 방법을 포함하며, 상기 설명된 RLP를 대상에 투여하는 단계를 포함한다. RLP는 대상에 경구, 피부 내, 비강 내, 근육 내, 복강 내, 정맥 내, 또는 피하로 투여될 수도 있다.

[0064] 본 발명은 또한, 상기 설명된 바와 같이, 방법 (A), 방법 (B), 방법 (C), 방법 (D), 또는 이것들의 조합에 의해 생산된 RLP를 포함하는 식물 재료를 제공한다. 식물 재료는 대상에서 로타바이러스 감염에 대한 면역력을 유발하는데 사용될 수도 있다. 식물 재료는 또한 식품 보충제로서 혼합될 수도 있다.

[0065] 상기 설명된 방법 (방법 A, B, C 또는 D)에서, 식물 또는 식물의 일부는 침묵의 억제인자를 암호화하는 또 다른 핵산 서열과 함께 더 투여될 수도 있거나, 이것을 더 포함할 수도 있다.

[0066] 게다가, 본 발명은 또한 다음을 포함하는 식물에서 로타바이러스 구조 단백질을 생산하는 방법 (E)를 제공한다

[0067] a) 하나 이상의 로타바이러스 구조 단백질을 암호화하는 뉴클레오티드 서열에 작동 간으하게 결합된, 식물에서 활성인 조절 영역을 포함하는 핵산을 식물, 또는 식물의 일부로 도입하는 단계;

[0068] b) 핵산의 일과성 발현을 허용하는 조건 하에 식물 또는 식물의 일부를 배양하고, 이로 인해 하나 이상의 로타바이러스 구조 단백질을 생산하는 단계.

[0069] 본 발명의 요약은 반드시 본 발명의 모든 특징을 설명하는 것은 아니다.

도면의 간단한 설명

[0070] 본 발명의 여러 특징은 첨부된 도면에 대한 참고가 이루어지는 다음 설명으로부터 더 분명해질 것이다:

도 1은 로타바이러스 구조 및 유전자-단백질 배치를 나타낸다. (A) 로타바이러스 입자의 투과 전자 현미경 검사 (transmission electron microscopy) (막대는 100nm를 나타낸다). (B) 내부, 중간 및 외부를 포함하는 바이러스 캡시드 단백질의 조직. (C) 크기 및 기능에 따라 배열된 로타바이러스 dsRNA 세그먼트. dsRNA는 폴리아크릴 아미드 겔 전기영동으로 분리될 수 있다 (D). (C)의 단백질은 (D)에서 dsRNA 세그먼트에 의해 지시된다. Crawford et al., 1997의 이미지 (A), Swiss Institute of Bioinformatics, 2008 (B) 및 Greenberg and Estes, 2009 (D).

도 2는 로타바이러스 세포 진입 및 복제를 나타낸다. 로타바이러스가 세포에 진입할 때, VP4 및 VP7이 손실되며, 이중층 입자 (DLP)를 형성한다. dsRNA의 전사는 VP2, VP4, VP6 및 VP7의 변역을 일으키기 시작한다. 레플리카제 활성을 가진 자손 코어는 바이러스 제작소 (비로플라즘(viroplasm)으로도 불림)에서 생산된다. 늦은 전사가 이 자손 코어들에서 일어난다. 바이러스 제작소의 주변부에서, 이 코어들은 VP6으로 코팅되며, 소포체의 막을 거쳐 베딩하는 미성숙한 DLP를 형성하고, ER 고유 바이러스 당단백질 NSP4 및 VP7로 변형된 일시적 지질막을 얻는다; 이 외피 보유 입자들은 또한 VP4를 함유한다. 입자들이 ER 시스터네(cisternae)의 내부로 이동하기 때문에, 일시적 지질막 및 비구조 단백질 NSP4가 손실되는 한편, 바이러스 표면 단백질 VP4 및 VP7은 가장 외부의 바이러스 단백질 층을 형성하도록 재배열되며, 성숙한 감염성 삼중층 입자를 수득한다 (Swiss Institute of Bioinformatics(ViralZone):viralzone.expasy.org/viralzone/all_by_species/107.html 참조).

도 3은 아그로박테리움 벡터 pTRAc, pTRAkc-rbcS1-cTP 및 pTRAkc-ERH, P35SS, 이중 전사 인핸서를 가진 CaMV 35S 프로모터; CHS, 칼콘(chalcone) 신타제 5' 번역되지 않은 영역; pA35S, CaMV 35S 폴리아데닐화 신호; SAR, 담배 Rb7 유전자의 스캐폴드 부착 영역; LB 및 RB, T-DNA 통합을 위한 왼쪽 및 오른쪽 보더(border); ColElori, 대장균(E. coli)에 대한 복제 기원; RK2ori, 아그로박테리움에 대한 복제 기원; bla, 앰피실린/카르베니실린-내성 bla 유전자; LPH, 쥐 mAb24 중쇄의 신호-펩티드 서열; his6, 6 x His 태그(tag) 서열; SEKDEL, ER-잔류 신호 서열; rbcS1-cTP, 감자(Solanum tuberosum)의 Rubisco 작은-서브유닛 유전자 (rbcS1)의 엽록체-통과 펩티드 서열; npt II, 카나마이신 내성 npt II 유전자; Pnos 및 pAnos, 노팔린 신타제 유전자의 프로모터

및 폴리아데닐화 신호를 나타낸다 (Maclean et al., 2007).

도 4는 로타바이러스 클로닝(cloning) 및 침투 과정의 개요를 나타낸다.

도 5는 아포플라스드 단백질 추출 과정을 나타낸다. (A) 식물 세포 및 아포플라스트의 장소의 도시, VP 단백질은 시토솔에서 발현되고 아포플라스트로 표적화된다 (빨간 화살표). (B) 타임 트라이얼 후, 식물 잎은 PBS와 함께 진공 침투되고 (1) 친공된 스핀 컬럼에 배치되고 (2) 그 다음에 2ml 에펜도르프튜브(Eppendorf tube)에서 원심분리하여 수액(sap)을 수거한다.

도 6은 식물 잎 세포 구획에서 7일 동안 로타바이러스 VP6 단백질의 발현의 웨스턴 블롯을 나타낸다. 마우스 항-로타바이러스 VP6 항체 (1:5000)는 막을 탐침하는데 사용되었다. (+) 및 (-)는 각각 침묵 억제인자가 있거나 없는 발현을 나타낸다. 빨간 선은 분석된 다양한 샘플에서 VP6 단백질의 위치를 나타낸다 (~ 40 kDa). VP6의 발현 및 추출 효능은 세포질에서 최고였다.

도 7은 제3 일에 니코티아나 벤타미아나(*N. benthamiana*) 식물 잎의 세포질에서 his-태그된 로타바이러스 단백질의 개개의 발현을 나타내는 웨스턴 블롯을 나타낸다. +ve-박테리아 발현된 로타바이러스 VP2; M-분자량 마커; VP-로타바이러스 캡시드 단백질. VP7 침투는 잎의 황화(yellowing)를 일으켰다 (b).

도 8은 제3 일에 다양한 니코티아나 벤타미아나 식물 잎 세포 구획으로 표적화된 로타바이러스 VP2 (a) 및 VP4 (b)의 발현을 나타낸다. VP2 및 VP4에 대한 각각의 닦 항-로타바이러스 혈청 (1:2000)은 단백질의 탐침에 사용되었다. cTp-엽록체; ER-소포체; pTRAc-세포질; A-아포플라스트; 음성 대조군 (-ve)-침묵 억제인자만 침투된 식물; (a)-박테리아 발현된 VP2, (b)-박테리아 발현된 VP4에서 양성 대조군 (+ve); (- 및 +) 침묵 억제인자 유무; M-분자량 마커. 화살표는 문제의 단백질 밴드의 위치를 나타낸다.

도 9는 니코티아나 벤타미아나 잎의 세포질에서 동시-발현된 VP2/6/4의 제3 일 추출물의 웨스턴 블롯 분석을 나타낸다. 단백질은 닦 항-로타바이러스 혈청 (항-VP2 (1/5000) 및 항-VP4 (1/5000)) 및 마우스 항-VP6 항체 (1:5000)의 혼합물로 탐침되었다. 재조합 아그로박테리움의 침투는 침묵 억제인자. 음성 대조군 (-ve)-전체 침묵 억제인자만 침투된 식물; M-분자량 마커로 실행되었다.

도 10은 유라닐 아세테이트로 염색된 제3 일 세포질 추출된 로타바이러스 단백질의 전자 현미경 사진을 나타낸다. (a) 침묵 억제인자가 있는 음성 단백질 샘플 추출물; (b) VP6 단백질 추출물; (c) VP2/6 단백질 추출물 및 (d) VP2/6/4 단백질 추출물. 막대 = 200 nm. 검출된 모든 RLP는 지름이 70-100 nm였다. (b)에서 화살표는 VP6 씨스(sheath)/매트(mat)를 나타낸다. (c)에서 화살표는 aRLP의 예를 나타낸다. 모든 단백질은 침묵 억제인자의 존재시 발현되었다. 모든 것은 마우스-항 VP6 항체 (1:2000)로 캡쳐되었다 (captured).

도 11은 동시-발현된 VP2/6 및 VP2/6/4의 수크로스 구배 정제를 나타낸다 (a). 수크로스 구배 정제된 VP2/6 (b) 및 VP2/6/4 (c)의 도트 블롯 (dot blot). 단백질 추출물은 수크로스 구배 (10-60%)에 로딩되었고 초원심분리되었다. 분획은 (b) 마우스 항-VP6 항체 (1:5000) 및 (c) 닦 항-VP2 및 VP4 혈청 (1:5000)으로 탐침함으로써 분석되었다.

도 12는 VP2/6 분획 (a)의 웨스턴 블롯 분석, 분획 VP2/6 분획 16 및 17 (b)의 SDS-PAGE 쿠마시 염색된 겔 사진, 및 분획 16 및 17 (c)의 웨스턴 블롯 분석을 나타낸다. 마우스 항-VP6 (1:5000) 및 닦 항-VP2 혈청 (1:5000)이 웨스턴 블롯에서 사용되었고 (a), (c)에서는 마우스 항-VP6 (1:5000)만이 사용되었다. 음성 대조군 (-ve) (a) 및 (c)에서 박테리아 발현된 VP4, 및 (b)에서는 침묵 억제인자가 침투되고 수크로스 구배 정제된 식물; 원래의-비-정제된 VP2/6 추출물; 양성 대조군 (a)에서 박테리아 발현된 VP2, (b) 및 (c)에서는 식물 발현된 VP6; SF9 곤충세포에서 발현된 알려진 농도의 VP6-SF9-VP6 단백질. 화살표는 원하는 단백질 밴드를 나타낸다.

도 13은 수크로스 밀도 구배 정제된 VP2/6의 분획에서 총 수용성 단백질 검정을 나타낸다. (a)-IgG 표준 곡선, (b)-750 nm에서 얻은 분획의 흡광도 판독값. 원하는 지점: 분획 16 내지 19.

도 14는 세포질 동시-발현된 VP2/6/4 분획의 수크로스 밀도 구배 분석을 나타낸다. 750 nm에서 미가공 흡광도 판독값은 VP2/6/4의 도트 블롯에서 이전에 검출된 단백질 피크를 확인하기 위해 취해졌다.

도 15는 수크로스 밀도 구배 정제된 VP2/6 입자의 투과 전자 현미경 사진을 나타낸다. (a) 및 (b) 둘 다는 구리 그리드(grid)에서 보이는 두 개의 다른 섹션을 나타낸다. 검출된 모든 RLP는 지름이 70-100 nm였다. 샘플은 마우스-항 VP6 항체 (1:2000)로 캡쳐되었다. 막대는 200 nm를 나타낸다.

도 16A는 로타바이러스 VP2의 아미노산 서열 (SEQ ID NO:1) 및 뉴클레오티드 서열 (SEQ ID NO:13 및 14)을 나

타낸다. 도 16B는 로타바이러스 VP4의 아미노산 서열 (SEQ ID NO:2) 및 뉴클레오티드 서열 (SEQ ID NO:15 및 16)을 나타낸다. 도 16C는 로타바이러스 VP6의 아미노산 서열 (SEQ ID NO:3) 및 뉴클레오티드 서열 (SEQ ID NO:17 및 18)을 나타낸다. 도 16D는 로타바이러스 VP7의 아미노산 서열 (SEQ ID NO:4) 및 뉴클레오티드 서열 (SEQ ID NO:19 및 20)을 나타낸다.

도 17A는 프라이머 IF-WA_VP2(opt).s1+3c의 뉴클레오티드 서열 (SEQ ID NO:21)을 나타낸다. 도 17B는 프라이머 IF-WA_VP2(opt).s1-4r의 뉴클레오티드 서열 (SEQ ID NO:22)을 나타낸다. 도 17C는 구조 1191의 개략도를 나타낸다. 플라스미드 선형화에 사용된 SacII 및 StuI 제한 효소 부위는 도면에 주석이 달려 있다.

도 18은 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로 구조 1191의 뉴클레오티드 서열 (SEQ ID NO:23)을 나타낸다 (밀줄 그어져 있음). 플라스토시아닌-P19-플라스토시아닌 침묵 억제자 발현 카세트가 있는 2X35S/CPMV-HT/NOS.

도 19는 로타바이러스 A WA 균주의 VP2(opt)를 암호화하는 뉴클레오티드 서열 (SEQ ID NO:45)을 나타낸다.

도 20은 로타바이러스 A WA 균주의 VP2의 아미노산 서열 (SEQ ID NO:25)을 나타낸다.

도 21은 구조 번호 1710의 개략도를 나타낸다.

도 22A는 구조 193의 개략도를 나타낸다. 플라스미드 선형화에 사용된 SacII 및 StuI 제한 효소 부위는 도면에 주석이 달려 있다. 도 22B는 구조 193의 뉴클레오티드 서열 (SEQ ID NO:26)을 나타낸다. 구조 193은 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로 나타난다 (밀줄 그어져 있음). 플라스토시아닌-P19-플라스토시아닌 침묵 억제자 발현 카세트가 있는 BeYDV(m)+레플리카제 증폭 시스템 내 2X35S/CPMV-HT/NOS.

도 23은 발현 카세트 1710의 뉴클레오티드 서열 (SEQ ID NO:27)을 나타낸다. 발현 카세트 번호 1710은 2X35S 프로모터에서 NOS 종결자까지 나타난다. 로타바이러스 A WA 균주의 VP2(opt)는 밀줄 그어져 있다.

도 24는 구조 번호 1711의 개략도를 나타낸다.

도 25A는 프라이머 IF-WA_VP6(opt).s1+3c의 뉴클레오티드 서열 (SEQ ID NO:28)을 나타낸다. 도 25B는 프라이머 IF-WA_VP6(opt).s1-4r의 뉴클레오티드 서열 (SEQ ID NO:29)을 나타낸다. 도 25C는 2X35S 프로모터에서 NOS 종결자까지의 발현 카세트 번호 1713 (SEQ ID NO:30)을 나타낸다. 로타바이러스 A WA 균주의 VP6(opt)는 밀줄 그어져 있다. 도 25D는 로타바이러스 A WA 균주의 VP6(opt)를 암호화하는 뉴클레오티드 서열 (SEQ ID NO:46)을 나타낸다.

도 26은 로타바이러스 A WA 균주의 VP6의 아미노산 서열 (SEQ ID NO:31)을 나타낸다.

도 27은 구조 번호 1713의 개략도를 나타낸다.

도 28은 2X35S 프로모터에서 NOS 종결자까지의 발현 카세트 번호 1714의 뉴클레오티드 서열 (SEQ ID NO:32)을 나타낸다. 로타바이러스 A WA 균주의 VP6(opt)은 밀줄 그어져 있다.

도 29는 구조 번호 1714의 개략도를 나타낸다.

도 30A는 프라이머 IF-Rtx_VP4(opt).s1+3c의 뉴클레오티드 서열 (SEQ ID NO:33)을 나타낸다. 도 30B는 프라이머 IF-Rtx_VP4(opt).s1-4r의 뉴클레오티드 서열 (SEQ ID NO:34)을 나타낸다.

도 31A는 2X35S 프로모터에서 NOS 종결자까지의 발현 카세트 번호 1731의 뉴클레오티드 서열 (SEQ ID NO:35)을 나타낸다. 로타바이러스 A Rotarix 균주의 VP4(opt)는 밀줄 그어져 있다. 도 31B는 균주 RVA/백신 /USA/Rotarix-A41CB052A/1988/G1P1A[8]의 로타바이러스 A VP4의 최적화된 암호화 서열 (SEQ ID NO:47)을 나타낸다. 도 31C는 2X35S 프로모터에서 NOS 종결자까지 발현 카세트 번호 1730의 뉴클레오티드 서열 (SEQ ID NO:44)을 나타낸다. 로타바이러스 A Rotarix 균주의 VP4(opt)는 밀줄 그어져 있다.

도 32는 로타바이러스 A Rotarix 균주의 VP4의 아미노산 서열 (SEQ ID NO:36)을 나타낸다.

도 33A는 구조 번호 1730의 개략도를 나타낸다. 도 33B는 구조 번호 1731의 개략도를 나타낸다.

도 34A는 프라이머 IF-Rtx_VP7(opt).s1+3c의 뉴클레오티드 서열 (SEQ ID NO:37)을 나타낸다. 도 34B는 프라이머 IF-Rtx_VP7(opt).s1-4r의 뉴클레오티드 서열 (SEQ ID NO:38)을 나타낸다. 도 34C는 2X35S 프로모터에서 NOS 종결자까지의 발현 카세트 번호 1733의 뉴클레오티드 서열을 나타낸다. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8]의 VP7 (SEQ ID NO:24)은 밀줄 그어져 있다. 도 34D는 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP7을 암호화하는 뉴클레오티드 서열 (SEQ ID NO:48)을

나타낸다. 도 34E는 균주 RVA/백신/USA/Rotarix-A41CB052A/1988/G1P1A[8]의 로타바이러스 A VP7의 최적화된 암호화 서열 (SEQ ID NO 54)을 나타낸다.

도 35는 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP7의 아미노산 서열 (SEQ ID NO:39)을 나타낸다.

도 36은 구조 번호 1733의 개략도를 나타낸다.

도 37은 프라이머 IF-Rtx_VP7(opt).s2+4c의 뉴클레오티드 서열 (SEQ ID NO:40)을 나타낸다.

도 38은 구조 1192의 개략도를 나타낸다. 플라스미드 선형화에 사용된 SacII 및 StuI 제한 효소 부위는 도면에 주석이 달려 있다.

도 39는 원쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로 구조 1192의 뉴클레오티드 서열 (SEQ ID NO:41)을 나타낸다 (밀줄 그어져 있음). 플라스토시아닌-P19-플라스토시아닌 침묵 억제자 발현 카세트가 있는 2X35S/CPMV-HT/PDISP/NOS가 나타난다.

도 40A는 2X35S 프로모터에서 NOS 종결자까지의 발현 카세트 번호 1735의 뉴클레오티드 서열 (SEQ ID NO:42)을 나타낸다. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 PDISP/VP7(opt)는 밀줄 그어져 있다. 도 40B는 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 PDISP/VP7(opt)을 암호화하는 뉴클레오티드 서열 (SEQ ID NO:49)을 나타낸다.

도 41은 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 PDISP/VP7의 아미노산 서열 (SEQ ID NO:43)을 나타낸다.

도 42는 구조 번호 1735의 개략도를 나타낸다.

도 43A는 균주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의 로타바이러스 A VP4의 암호화 서열 (SEQ ID NO:50)을 나타낸다. 도 43B는 균주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의 로타바이러스 A VP4의 최적화된 암호화 서열 (SEQ ID NO:51)을 나타낸다. 도 43C는 균주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의 로타바이러스 A VP7의 암호화 서열 (SEQ ID NO:52)을 나타낸다. 도 43D는 균주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의 로타바이러스 A VP7의 최적화된 암호화 서열 (SEQ ID NO:53)을 나타낸다.

도 44A는 프라이머 IF-TrSP+Rtx_VP7(opt).s1+3c의 뉴클레오티드 서열 (SEQ ID NP: 55)을 나타낸다. 도 44B는 프라이머 IF-Rtx_VP7(opt).s1-4r의 뉴클레오티드 서열 (SEQ ID NO:56)을 나타낸다. 도 44C는 균주 RVA/백신/USA/Rotarix-A41CB052A/1988/G1P1A[8]의 로타바이러스 A VP7의 최적화된 암호화 서열의 뉴클레오티드 서열 (SEQ ID NO:57)을 나타낸다. 도 44D는 2X35S 프로모터에서 NOS 종결자까지 발현 카세트 번호 1734의 뉴클레오티드 서열 (SEQ ID NO 58)을 나타낸다. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP7은 밀줄 그어져 있다. 도 44E는 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 TrSp-VP7의 아미노산 서열 (SEQ ID NO:59)을 나타낸다. 도 44F는 구조 번호 1734의 개략도를 나타낸다.

도 45는 이오디사놀 밀도 구배 원심분리에 의한 VP2 및 VP6을 포함하는 로타바이러스-유사 입자의 정체를 나타낸다. 도 45A 원심분리 전 로딩된 것들 및 분획 1 내지 10 (분획 1은 튜브의 바닥에 있음)의 쿠마시-염색된 SDS-PAGE 분석. 로타바이러스 항원의 위치는 화살표로 나타난다. 도 45B 토끼 다클론성 항-로타바이러스 항체를 사용하여 (A)와 같은 분획의 웨스턴 블롯 분석. 도 45C 토끼 다클론성 항-VP2 항체를 사용하여 (A)와 같은 분획의 웨스턴 블롯 분석.

도 46은 이오디사놀 밀도 구배 원심분리에 의한 VP2, VP6 및 VP7을 포함하는 로타바이러스-유사 입자의 정체를 나타낸다. 도 46A 원심분리 전 로딩된 것들 및 분획 1 내지 10 (분획 1은 튜브의 바닥에 있음)의 쿠마시-염색된 SDS-PAGE 분석. 로타바이러스 항원의 위치는 화살표로 나타난다. 도 46B 토끼 다클론성 항-로타바이러스 항체를 사용하여 (A)와 같은 분획의 웨스턴 블롯 분석. 도 46C 토끼 다클론성 항-VP2 항체를 사용하여 (A)와 같은 분획의 웨스턴 블롯 분석.

도 47은 이오디사놀 밀도 구배 원심분리에 의한 VP2, VP4, VP6 및 VP7을 포함하는 로타바이러스-유사 입자의 정체를 나타낸다. 도 47A 원심분리 전 로딩된 것들 및 분획 1 내지 10 (분획 1은 튜브의 바닥에 있음)의 쿠마시-염색된 SDS-PAGE 분석. 로타바이러스 항원의 위치는 화살표로 나타난다. 도 47B 토끼 다클론성 항-로타바이러스 항체를 사용하여 (A)와 같은 분획의 웨스턴 블롯 분석. 도 47C 토끼 다클론성 항-VP7 항체를 사용하여 (A)와 같은 분획의 웨스턴 블롯 분석.

도 48은 항-VP4 특이적 ELISA에 의한 VP2, VP4, VP6 및 VP7을 포함하는 정제된 로타바이러스-유사 입자에서 VP4 함량의 평가를 나타낸다.

도 49는 VP2 및 VP6 (왼쪽 패널)과 VP2, VP4, VP6 및 VP7 (오른쪽 패널)을 포함하는 정제된 로타바이러스-유사 입자의 냉동-전자 현미경 이미지를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0071]

다음 설명은 바람직한 구체예이다.

[0072]

본 발명은 하나 이상의 로타바이러스 구조 단백질 (즉, 로타바이러스 유사 입자, 로타바이러스 VLP 또는 RLP)을 포함하는 바이러스-유사 입자 (VLP), 및 식물에서 로타바이러스-유사 입자 (RLP)를 생산하는 방법에 관한 것이다. 그러므로 로타바이러스 유사 입자 (RLP)는 하나 이상의 로타바이러스 구조 단백질을 포함할 수도 있다. RLP는 이중층 또는 삼중층일 수도 있다.

[0073]

본 발명은 식물에서 로타바이러스-유사 입자 (RLP)를 생산하는 방법을 부분적으로 제공한다. 방법은 하나 이상의 로타바이러스 구조 단백질을 암호화하는 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 조절 영역을 포함하는 하나 이상의 핵산을 식물, 또는 식물의 일부로 도입하는 단계를 포함할 수도 있다. 핵산의 일과성 발현을 허용하는 조건 하에 식물 또는 식물의 일부를 배양한 후, RLP를 생산하는 단계로 이어진다.

[0074]

게다가, 본 발명은 식물에서 로타바이러스-유사 입자 (RLP) 백신 후보물질을 생산하는 방법을 부분적으로 제공한다. 방법은 제1 로타바이러스 구조 단백질을 암호화하는 제1 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제1 조절 영역을 포함하는 제1 핵산, 제2 로타바이러스 구조 단백질을 암호화하는 제2 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제2 조절 영역을 포함하는 제2 핵산 및 제3 로타바이러스 구조 단백질을 암호화하는 제3 뉴클레오티드 서열에 작동 가능하게 결합된, 식물에서 활성인 제3 조절 영역을 포함하는 제3 핵산을 식물, 식물의 일부 또는 식물 세포로 도입하는 단계를 포함할 수도 있다. 제1, 제2 및 제3 핵산의 일과성 발현을 허용하는 조건 하에 식물, 식물의 일부 또는 식물 세포를 배양한 후, 이어서 RLP를 생산하는 단계로 이어진다. RLP는 단층, 이중층 또는 삼중층일 수도 있다.

[0075]

"로타바이러스 구조 단백질"은 로타바이러스로부터 분리된 로타바이러스 구조 단백질 서열 모두 또는 일부를 나타낼 수도 있으며, 어떤 자연 발생한 또는 변이체 로타바이러스 균주 또는 분리체에도 존재한다. 따라서, 용어로타바이러스 구조 단백질 등은 바이러스 생활 주기 중에 돌연변이에 의해 생산되거나 선택압 (예를 들어, 약물치료, 숙주 세포 향성(tropism)의 확장 또는 감염성, 등)에 반응하여 생산되는 자연 발생한 로타바이러스 구조 단백질 서열 변이체를 포함한다. 당업자가 인정하는 바와 같이, 이러한 로타바이러스 구조 단백질 서열 및 이것들의 변이체는 또한 재조합 기술을 사용하여 생산될 수도 있다.

[0076]

게다가, 구조 단백질은 예를 들어, VP2 및 VP6과 같은 캡시드 단백질 및/또는 예를 들어, VP4와 같은 표면 단백질을 포함할 수도 있다. 구조 단백질은 예를 들어, VP7을 더 포함할 수도 있다.

[0077]

로타바이러스 구조 단백질의 비-제한 예는 로타바이러스 단백질 VP2, VP4, VP6 및 VP7, 및 VP2, VP4, VP6 및 VP7의 단편이다. 본 발명에 따라 사용될 수도 있는 VP2, VP4, VP6 및 VP7, 또는 VP2, VP4, VP6 및 VP7의 단편 단백질의 비-제한 예는 로타바이러스 균주 G9 P[6], 로타바이러스 A WA 균주, 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주 및 로타바이러스 SA11 균주의 상기 VP2, VP4 VP6 및 VP7 단백질을 포함한다.

[0078]

VP2 구조 단백질의 예는, 제한하는 것으로 간주되는 것은 아니지만, SEQ ID NO:1 및 SEQ ID NO:25의 아미노산 서열에서 제시된다. 게다가, VP2 구조 단백질은 SEQ ID NO:1, SEQ ID NO:25, 또는 이것들에 대하여 적어도 약 90-100% 서열 유사성, 예를 들어, 이것들에 대하여, 91, 92, 93, 94, 95, 96, 97, 98, 99% 서열 유사성과 같이, 이 범위 내에서 어떤 퍼센트 유사성을 갖는 서열에서 제시된 서열을 포함할 수도 있다. 게다가, VP2 구조 단백질은 SEQ ID NO:13, 14, 25, 또는 45 또는 이것들에 대하여 적어도 약 80-100% 서열 유사성, 예를 들어, 이것들에 대하여, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% 서열 유사성과 같이, 이 범위 내에서 어떤 퍼센트 유사성을 갖는 서열에서 제시된 뉴클레오티드 서열에 의해 암호화될 수도 있다.

[0079]

VP4 구조 단백질의 예는, 제한하는 것으로 간주되는 것은 아니지만, SEQ ID NO:2 및 SEQ ID NO:36의 아미노산 서열에서 제시된다. 게다가, VP4 구조 단백질은 SEQ ID NO:2, SEQ ID NO:36 또는 이것들에 대하여 적어도 약 90-100% 서열 유사성, 예를 들어, 이것들에 대하여, 91, 92, 93, 94, 95, 96, 97, 98, 99% 서열 유사성과

같이, 이 범위 내에서 어떤 퍼센트 유사성을 갖는 서열에서 제시된 서열을 포함할 수도 있다. 게다가, VP4 구조 단백질은 SEQ ID NO:15, 16, 47, 50 또는 51 또는 이것들에 대하여 적어도 약 80-100% 서열 유사성, 예를 들어, 이것들에 대하여, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% 서열 유사성과 같이, 이 범위 내에서 어떤 퍼센트 유사성을 갖는 서열에서 제시된 뉴클레오티드 서열에 의해 암호화될 수도 있다.

[0080] VP6 구조 단백질의 예는, 제한하는 것으로 간주되는 것은 아니지만, SEQ ID NO:3 및 SEQ ID NO:31의 아미노산 서열에서 제시된다. 게다가, VP6 구조 단백질은 SEQ ID NO:3, SEQ ID NO:31 또는 이것들에 대하여 적어도 약 90-100% 서열 유사성, 예를 들어, 이것들에 대하여, 91, 92, 93, 94, 95, 96, 97, 98, 99% 서열 유사성과 같이, 이 범위 내에서 어떤 퍼센트 유사성을 갖는 서열에서 제시된 서열을 포함할 수도 있다. 게다가, VP6 구조 단백질은 SEQ ID NO:17, 18, 또는 46 또는 이것들에 대하여 적어도 약 80-100% 서열 유사성, 예를 들어, 이것들에 대하여, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% 서열 유사성과 같이, 이 범위 내에서 어떤 퍼센트 유사성을 갖는 서열에서 제시된 뉴클레오티드 서열에 의해 암호화될 수도 있다.

[0081] VP7 구조 단백질의 예는, 제한하는 것으로 간주되는 것은 아니지만, SEQ ID NO:4, SEQ ID NO:39, SEQ ID NO:43, 및 SEQ ID NO:47의 아미노산 서열에서 제시된다. 게다가, VP7 구조 단백질은 SEQ ID NO:4, SEQ ID NO:39 및 SEQ ID NO:43, 또는 이것들에 대하여 적어도 약 90-100% 서열 유사성, 예를 들어, 이것들에 대하여, 91, 92, 93, 94, 95, 96, 97, 98, 99% 서열 유사성과 같이, 이 범위 내에서 어떤 퍼센트 유사성을 갖는 서열에서 제시된 서열을 포함할 수도 있다. 게다가, VP7 구조 단백질은 SEQ ID NO:19, 20, 48, 49, 52, 53 또는 54 또는 이것들에 대하여 적어도 약 80-100% 서열 유사성, 예를 들어, 이것들에 대하여, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% 서열 유사성과 같이, 이 범위 내에서 어떤 퍼센트 유사성을 갖는 서열에서 제시된 뉴클레오티드 서열에 의해 암호화될 수도 있다.

[0082] 아미노산 서열 유사성 또는 동일성은 BLAST (기초적 부분 정렬 검색 도구) 2.0 알고리즘을 이용하는 BLASTP 및 TBLASTN 프로그램을 사용함으로써 계산될 수도 있다. 아미노산 서열 유사성 또는 동일성을 계산하는 기술은 당업자에게 잘 알려져 있고, BLAST 알고리즘의 사용은 ALTSCHUL et al. (1990, J Mol. Biol. 215: 403-410) 및 ALTSCHUL et al. (1997, Nucleic Acids Res. 25: 3389-3402)에서 설명된다.

[0083] 용어 "바이러스-유사 입자" (VLP), 또는 "바이러스-유사 입자" 또는 "VLP"는 자가-조립하여 하나 이상의 구조 단백질, 예를 들어, 로타바이러스 구조 단백질을 포함하지만, VP2, VP4, VP6 및/또는 VP7 구조 단백질에 제한되지 않는 구조를 나타낸다. 로타바이러스 구조 단백질을 포함하는 VLP는 또한 "로타바이러스 VLP", "로타바이러스-유사 입자 (RVLP)", "로타바이러스-유사 입자 (RLP)", "로타바이러스-유사 입자", "RVLP" 또는 "RLP"로도 불릴 수 있다. VLP 또는 RLP는 보통 감염 시 생산된 비리온과 형태적으로 및 항원적으로 유사하지만, 복제하는데 충분한 유전적 정보가 부족하고 따라서 비-감염성이다. VLP는 식물 숙주 세포를 포함하는 적합한 숙주 세포에서 생산될 수도 있다. 숙주 세포로부터 추출 및 분리 및 적합한 조건 하에서 추가의 정제 후, VLP는 온전한 구조로 정제될 수도 있다. RLP는 단층, 이중층 또는 삼중층 RLP일 수도 있다. 단층 RLP는 VP2 또는 VP6과 같은 로타바이러스 구조 단백질을 발현함으로써 얻어질 수도 있다. 이중층 RLP는 두 개의 로타바이러스 구조 단백질을 발현함으로써, 예를 들어, VP4가 있거나 없이, VP2 및 VP6 둘 다를 동시 발현함으로써 얻어질 수도 있다. 삼중층 RLP는 적어도 세 개의 로타바이러스 구조 단백질의 동시에 발현, 예를 들어, VP4가 있거나 없이, VP2, VP6 및 VP7의 동시-발현에 의해 얻어질 수도 있다. VP의 동시-발현은 고유 로타바이러스와 유사한 스파이크를 가진 입자를 발생시킨다. VP4는 VP5 및 VP8을 생산하도록 가공되거나 분할될 수도 있다. 이 가공은 내인성 프로테아제를 사용하여, 또는 적합한 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 세르인-유사 프로테아제, 서브틸리신을 동시에 발현함으로써 숙주 내에서 발생할 수도 있다. 대안으로, VP4는 RLP 추출 과정의 어떤 단계 중에, 또는 RLP 정제 후에 적합한 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 추가함으로써 VP5 및 VP8을 생산하도록 가공될 수도 있다.

[0084] 각각의 로타바이러스 구조 단백질은 다른 특성 및 크기를 가지며, RLP로의 조립에 다른 양이 필요하다. 용어 "로타바이러스 VLP", "로타바이러스 바이러스-유사 입자 (RVLP)", "로타바이러스 바이러스-유사 입자 (RLP)", "로타바이러스 바이러스-유사 입자", "RVLP" 또는 "RLP"는 하나 이상의 로타바이러스 구조 단백질을 포함하는 바이러스-유사 입자 (VLP)를 나타낸다. 로타바이러스 구조 단백질의 예는 VP2, VP4 (또는 VP5 및 VP8), VP6 및 VP7 구조 단백질을 포함하지만, 이에 제한되지 않는다.

[0085]

본 발명은 또한 식물에서 RLP를 생산하는 방법을 제공하며, 제1 로타바이러스 구조 단백질, 예를 들어, VP2 또는 VP6 단백질을 암호화하는 제1 핵산 (제1 핵산)은 제2 로타바이러스 구조 단백질, 예를 들어, VP6 또는 VP2 단백질을 암호화하는 제2 핵산과 동시-발현된다. 게다가, 제3 로타바이러스 구조 단백질, 예를 들어, VP4 또는 VP7을 암호화하는 제3 핵산은 식물에서 제1 핵산, 제2 핵산 및 제3 핵산이 동시-발현되도록 제1 및 제2 핵산과 동시에 발현될 수도 있다. 제1 핵산, 제2 핵산 및 제3 핵산은 같은 단계에서 식물에 도입될 수도 있거나, 순차적으로 식물에 도입될 수도 있다. VP4는 적합한 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 암호화하는 핵산을 동시에-발현함으로써 숙주 내에서 VP5 및 VP8을 생산하도록 가공되거나 분할될 수도 있다. 대안으로, VP4는 만족스러운 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 추가함으로써 RLP 추출의 어떤 단계 중에, 또는 RLP 정제 후에 가공될 수도 있다.

[0086]

게다가, 제1 로타바이러스 구조 단백질을 암호화하는 제1 핵산, 제2 로타바이러스 구조 단백질을 암호화하는 제2 핵산 및 제3 로타바이러스 구조 단백질을 암호화하는 제3 핵산을 발현하는 식물은 제1 핵산, 제2 핵산, 제3 및 제4 핵산이 식물에서 동시에-발현되도록 제4 로타바이러스 구조 단백질, 예를 들어, VP7 또는 VP4 단백질을 암호화하는 제4 핵산으로 더 형질전환될 수도 있다. VP4는 적합한 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 암호화하는 핵산을 동시에-발현함으로써 숙주 내에서 VP5 및 VP8을 생산하도록 가공되거나 분할될 수도 있다. 대안으로, VP4는 만족스러운 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 추가함으로써 RLP 추출의 어떤 단계 중에, 또는 RLP 정제 후에 가공될 수도 있다.

[0087]

게다가, 하나 이상의 로타바이러스 구조 단백질, 예를 들어, VP2 또는 VP6 단백질을 암호화하는 제1 핵산을 발현하는 제1 식물은 각각 VP2 및 VP6 또는 VP6 및 VP2를 암호화하는 제1 및 제2 핵산을 동시에-발현하는 자손 식물 (제3 식물)을 생산하기 위해, 하나 이상의 로타바이러스 구조 단백질, 예를 들어, 제한은 아니지만, VP6 또는 VP2 단백질을 암호화하는 제2 핵산을 발현하는 제2 식물과 교배될 수도 있다. 게다가, 각각 VP2 및 VP6 또는 VP6 및 VP2를 암호화하는 제1 및 제2 핵산을 발현하는 제3 식물은 VP2, VP6, 및 VP4 또는 VP7을 암호화하는 제1, 제2 및 제3 핵산을 동시에-발현하는 추가의 자손 식물 (제5 식물)을 생산하기 위해, 하나 이상의 로타바이러스 구조 단백질 예를 들어, 제한은 아니지만, VP4 또는 VP7을 암호화하는 제3 핵산을 발현하는 제4 식물과 교배될 수도 있다. VP4는 숙주 프로테아제를 사용하여, 또는 제1, 제2, 제3 또는 제4 식물 중 하나의 내에서 적합한 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 암호화하는 핵산을 동시에-발현함으로써 식물 내에서 VP5 및 VP8을 생산하도록 가공되거나 분할될 수도 있다. 대안으로, VP4는 만족스러운 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 추가함으로써 RLP 추출의 어떤 단계 중에, 또는 RLP 정제 후에 가공될 수도 있다.

[0088]

하기 더 상세한 설명에서 설명된 바와 같이, RLP는 하나 이상의 로타바이러스 구조 단백질, 예를 들어, 제한은 아니지만, VP2, VP6 또는 VP7을 암호화하는 핵산 (제1 핵산)을 발현함으로써 식물에서 생산될 수도 있다. 제2 로타바이러스 구조 단백질, 예를 들어, 제한은 아니지만, VP7, VP6 또는 VP2를 암호화하는 제2 핵산은 식물에서 동시에-발현될 수도 있다. 게다가, 제3 로타바이러스 구조 단백질, 예를 들어, 제한은 아니지만, VP6, VP7 또는 VP2를 암호화하는 제3 핵산은 식물에서 동시에-발현될 수도 있다. 핵산, 제2 핵산 및 제3 핵산은 같은 단계에서 식물에 도입될 수도 있거나, 그것들은 순차적으로 식물에 도입될 수도 있다. 핵산, 제2 핵산 및 제3 핵산은 일시적 방식으로, 또는 안정한 방식으로 식물에서 도입될 수도 있다.

[0089]

게다가, 제1 로타바이러스 구조 단백질, 예를 들어, VP2 단백질을 암호화하는 제1 핵산을 발현하는 식물은 제1 및 제2 핵산 둘 다가 식물에서 동시에-발현되도록 제2 로타바이러스 구조 단백질, 예를 들어, 제한은 아니지만, VP6 또는 VP7을 암호화하는 제2 핵산으로 형질전환될 수도 있다. 식물은 제3 로타바이러스 구조 단백질, 예를 들어, 제한은 아니지만, VP7 또는 VP6을 암호화하는 제3 핵산으로 더 형질전환될 수도 있다.

[0090]

대안으로, VP6 또는 VP7 단백질을 발현하는 식물, (제2 핵산)은 제1 및 제2 핵산 둘 다가 식물에서 동시에-발현되도록, VP2 단백질을 암호화하는 제1 핵산으로 형질전환 될 수도 있다. 식물은 제3 로타바이러스 구조 단백질, 예를 들어, 제한은 아니지만, VP7 또는 VP6을 암호화하는 제3 핵산으로 더 형질전환될 수도 있다.

[0091]

게다가, 제1 및 제2 로타바이러스 구조 단백질, 예를 들어, VP2 및 VP6 단백질을 암호화하는 제1 및 제2 핵산을 발현하는 식물은 제3 로타바이러스 구조 단백질, 예를 들어, VP4 또는 VP7을 암호화하는 제3 핵산으로 형질전환될 수도 있다. VP4는 적합한 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모

트립신-유사 프로테아제, 서브틸리신을 암호화하는 핵산을 동시-발현함으로써 VP5 및 VP8을 생산하도록 가공되거나 분할될 수도 있다. 대안으로, VP4는 만족스러운 프로테아제, 예를 들어, 트립신, 트립신-유사 프로테아제, 세린 프로테아제, 키모트립신-유사 프로테아제, 서브틸리신을 추가함으로써 RLP 추출의 어떤 단계 중에, 또는 RLP 정제 후에 가공될 수도 있다.

[0092] 본 발명은 또한 식물에서 활성인 조절 영역에 작동 가능하게 결합된 하나 이상의 로타바이러스 구조 단백질을 암호화하는 하나 이상의 핵산, 및 하나 또는 그 이상의 구획 표적화 서열 및/또는 중복 요소를 식물, 식물의 일부 또는 식물 세포로 도입하는 단계를 포함하는 식물에서 RLP를 생산하는 방법을 제공한다. 그 다음에 식물, 식물의 일부 또는 식물 세포는 하나 이상의 핵산의 발현을 허용하는 조건 하에서 배양되고, 이로 인해 RLP를 생산한다. 하나 이상의 로타바이러스 구조 단백질은 VP2, VP4 (또는 VP5 및 VP8), VP6, VP7, VP2, VP4 (또는 VP5 및 VP8), VP6, VP7의 단편 또는 이것들의 조합일 수도 있다.

[0093] 본 발명은 하나 이상의 로타바이러스 구조 단백질, 예를 들어, 제한은 아니지만, VP2, VP4 (또는 VP5 및 VP8), VP6, VP7 또는 이것들의 조합을 포함하는 RLP를 더 제공한다. RLP는 본 발명에 의해 제공된 방법 중 하나 이상에 의해 생산될 수도 있다.

[0094] RLP의 발생은 어떤 적합한 방법, 예를 들어, 밀도 구배 원심분리 또는 크기 배제 크로마토그래피(size exclusion chromatography)를 사용해서 검출될 수 있다. RLP는, 예를 들어, 전자 현미경 검사, 또는 크기 배제 크로마토그래피에 의해 구조 및 크기에 대하여 평가될 수도 있다.

[0095] 크기 배제 크로마토그래피를 위해, 전체 수용성 단백질은 냉동-분쇄된 식물 재료, 및 원심분리에 의해 제거된 불용성 재료의 샘플을 균질화함으로써 (Polytron) 식물 조직으로부터 추출될 수도 있다. 매우 차가운 아세톤 또는 PEG로의 침전은 또한 유익할 수도 있다. 수용성 단백질은 정량화되고, 추출물은 Sephadryl™ 컬럼, 예를 들어, Sephadryl™ S500 컬럼을 통과하였다. Blue Dextran 2000은 보정 기준으로서 사용될 수도 있다. 크로마토그래피 후, 분획은 분획의 단백질 보충을 결정하기 위해 면역블롯에 의해 더 분석될 수도 있다.

[0096] 분리된 분획은, 예를 들어, 예를 들어, 상층액 (원심분리, 침강, 또는 침전되면), 또는 여과물 (여과되면)일 수도 있고, 단백질, 또는 상기 구조 단백질, 예를 들어, 나노튜브, 나노구체 또는 단층 (s1), 이중층 (d1) 또는 삼중층 (t1) RLP와 같은 고차원, 고분자량 입자가 풍부하다.

[0097] 분리된 분획은, 예를 들어, 추가 원심분리 단계, 침전, 크로마토그래피 단계 (예를 들어, 크기 배제, 이온 교환, 친화도 크로마토그래피), 접선 유동 여과(tangential flow filtration), 또는 이것들의 조합에 의해 단백질, 상기 구조 단백질 또는 고차원 입자를 분리하거나, 정제하거나, 농축하거나, 또는 분리하고, 정제하고 농축하도록 더 가공될 수도 있다. 정제된 단백질, 상기 구조 단백질 또는 RLP와 같은 고차원 입자의 존재는, 예를 들어, 고유한 분석 또는 SDS-PAGE 분석, 적절한 검출 항체를 사용하는 웨스턴 분석, 모세관 전기영동(capillary electrophoresis), 전자 현미경 검사, 또는 당업자에게 분명한 어떤 다른 방법에 의해서도 확인될 수 있다.

[0098] 본 발명에 따라 생산된 RLP는 식물, 식물의 일부 또는 식물 재료로부터 정제되거나, 부분적으로 정제될 수도 있거나, 또는 당업자에게 알려져 있는 방법을 사용하여 경구 백신으로서 투여될 수도 있다.

[0099] RLP 정제는 구배 원심분리, 예를 들어, 수크로스, 이오덕사놀, OptiPrep™ 또는 세슘 클로라이드 (CsCl) 밀도 구배를 수반할 수도 있으며 형질전환된 식물 바이오매스(biomass)로부터 RLP를 정제하거나 부분적으로 정제하는데 사용될 수도 있다. 예를 들어, 도 45에서 나타난 바와 같이, 이오덕사놀 단계 구배 또는 이오덕사놀 연속 구배는 RLP 및/또는 발현된 로타바이러스 구조 단백질을 정제하는데 사용될 수도 있다.

[0100] 칼슘 (Ca^{2+}) 농도는 삼중층 입자 (TLP) 내지 이중층 입자 (DLP) 형질전환에 중요한 것으로 나타났고 균주 의존적이다 (예를 들어, Martin et al. Journal of Virology, Jan 2002를 참고하면 되는데, 이것은 본원에 참고로 포함된다). TLP에서 외부-캡시드 단백질의 완전한 상실 (TLP 탈캡시드화(decapsidation))은 $[\text{Ca}^{2+}]$ 의 나노몰 범위에서 일어난다. 그러므로 RLP의 정제 및/또는 추출은 칼슘의 존재시 수행될 수도 있고, 구배 원심분리의 단계는 칼슘의 존재시, 예를 들어, CaCl_2 의 존재시 수행될 수도 있다. CaCl_2 의 농도는, 예를 들어, 1 mM와 1000 mM 사이, 또는 그 사이의 어떤 양, 예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 50, 600, 650, 700, 750, 800, 850, 900, 950 mM 또는 그 사이의 어떤 양도 될 수 있다.

[0101] 식물, 또는 식물 단편은 최소한으로 가공될 수도 있다. 용어 "최소 가공"은 식물 재료, 예를 들어, 식물 추출물, 균질액, 식물 균질액의 분획 등을 수득하기 위해 부분적으로 정제된 (즉, 최소한으로 가공된) 원하는

단백질 및/또는 RLP를 포함하는 식물 또는 이것의 일부를 의미한다. 부분적 정제는 식물 세포 구조를 분열시켜서, 예를 들어, 제한은 아니지만, 원심분리, 여과 또는 이것들의 조합에 의해 분리될 수도 있는 수용성 식물 구성요소, 및 불용성 식물 구성요소를 포함하는 조성물을 생성하는 단계를 포함할 수도 있지만, 이에 제한되지 않는다. 이 점에서, 일 또는 다른 조직의 세포 외 공간 내에 분비된 단백질은 진공 또는 원심분리 추출을 사용하여 쉽게 얻어질 수 있거나, 조직은 세포 외 공간 내에서 단백질을 짜내거나 해방시키기 위한 롤러(roller) 통과 또는 그라인딩(grinding)에 의한 압력 하에 추출될 수 있다. 최소 가공은 또한 수용성 단백질의 조추출물(crude extract)의 제조를 수반할 수 있는데, 이 조제물은 2차 식물 생성물에서 미량의 오염을 가질 것이기 때문이다. 게다가, 최소 가공은 일의 수용성 단백질의 수성 추출에 이어서, 어떤 적합한 염으로의 침전을 수반할 수도 있다. 다른 방법은 추출물의 직접적인 사용을 허용하기 위해 대규모 매서레이션(maceration) 및 착즙(juice extraction)을 포함할 수도 있다. RLP는 어떤 적합한 방법, 예를 들어, 기계적 또는 생화학적 추출을 사용하여 정제되거나 추출될 수도 있다.

[0102] 하나 이상의 로타바이러스 구조 단백질은 식물 생중량(fresh weight)의 킬로그램 당 최대 2 g의 양으로 합성될 수도 있다. 예를 들어, 합성된 구조 단백질의 양은 생중량의 킬로그램 당 1 내지 2 g, 또는 그 사이의 어떤 값, 예를 들어, 생중량의 킬로그램 당 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2 g 또는 그 사이의 어떤 값도 될 수 있다. 예를 들어, 구조 단백질은 식물 생중량의 킬로그램 당 최대 1.54 g의 양으로 합성될 수도 있다.

[0103] 게다가, RLP는 식물 생중량의 킬로그램 당 최대 1.5 g의 양으로 합성될 수도 있는데, 예를 들어, 합성된 RLP의 양은 생중량의 킬로그램 당 0.5 내지 1.5 g, 또는 그 사이의 어떤 값, 예를 들어, 생중량의 킬로그램 당 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5 g도 될 수 있다. 예를 들어, RLP는 식물 생중량의 킬로그램 당 최대 1.1 g의 양으로 합성될 수도 있다.

[0104] 상기-정의된 RLP의 크기 (즉, 지름)는, 예를 들어, 동적 광 산란 (dynamic light scattering; DLS) 또는 전자 현미경 (EM) 기술에 의해 측정될 수도 있고, 보통 50 내지 110 nm 사이, 또는 그 사이의 어떤 크기도 된다. 예를 들어, 온전한 RLP 구조의 크기는 약 70 nm 내지 약 110 nm, 또는 그 사이의 어떤 크기, 예를 들어, 75 nm, 80 nm, 85 nm, 90 nm, 95 nm, 100 nm, 105 nm 또는 그 사이의 어떤 크기의 범위에도 있을 수 있다.

[0105] 본 발명은 식물에서 활성성인 조절 영역에 작동 가능하게 결합된 하나 이상의 로타바이러스 구조 단백질을 암호화하는 뉴클레오티드 서열을 포함하는 핵산을 더 제공한다. 뉴클레오티드 서열은, 예를 들어, 인간 코돈 사용 또는 식물 코돈 사용에 최적화될 수도 있다. 게다가 하나 이상의 로타바이러스 구조 단백질은 하나 또는 그 이상의 증폭 요소에 작동 가능하게 결합될 수도 있다. 게다가, 하나 이상의 로타바이러스 구조 단백질은 하나 또는 그 이상의 구획 표적화 서열에 작동 가능하게 결합될 수도 있다. 뉴클레오티드 서열에 의해 암호화된 하나 이상의 로타바이러스 구조 단백질은, 예를 들어, VP2, VP4, VP6 또는 VP7일 수도 있다. 게다가 뉴클레오티드 서열에 의해 암호화된 하나 이상의 로타바이러스 구조 단백질은, 예를 들어, 어떤 로타바이러스 군 A 내지 G의 것일 수도 있지만, 더 바람직하게는 로타바이러스 군 A의 것이다. 게다가, 뉴클레오티드 서열에 의해 암호화된 하나 이상의 로타바이러스 구조 단백질은 G1 내지 G27 및 P1 내지 P34, 및 더 바람직하게 G1 내지 G19 및 P1 내지 P27의 G-및 P-타입의 어떤 조합의 유전자형을 갖는 어떤 로타바이러스 균주의 것도 될 수 있으며, G1P[8], G2P[4], G2P[8], G3P[8], G4P[8], G9P[6], G9P[8], 로타바이러스 A WA 균주, 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주 또는 로타바이러스 SA11 균주를 포함하지만, 이에 제한되지 않는다.

[0106] 본 발명에서 언급된 핵산 서열은 핵산 서열이 본원에서 정의된 바와 같이 염증 합종화(hybridisation) 조건 하에서 하나 또는 그 이상의 뉴클레오티드 서열 또는 핵산 서열의 보체에 합종화하면 서열, 또는 서열의 보체에 "실질적으로 상동성"이거나, "실질적으로 유사"하거나 또는 "실질적으로 동일" 할 수도 있다. 서열은 뉴클레오티드 중 적어도 약 70%, 또는 70 내지 100%, 또는 그 사이의 어떤 양, 예를 들어, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100%, 또는 그 사이의 어떤 양이 정의된 길이의 뉴클레오티드 서열에 걸쳐 일치할 때 "실질적으로 상동성"이거나, "실질적으로 유사"하거나, "실질적으로 동일"하며, 이러한 상동성 서열은 서열의 속성, 또는 본원에서 설명된 암호화된 생성물 중 하나 또는 그 이상을 나타낸다는 것을 제공한다.

[0107] 예를 들어, 본 발명은 예를 들어, SEQ ID NO:13, 14, 15, 16, 17, 18, 19, 20, 45, 46, 47, 49, 50, 51, 52, 53, 54에서 정의된 서열과 적어도 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 100% 또는 그 사이의 어떤 양이 동일한 하나 이상의 로타바이러스 구조 단백질을 암호화하는 핵산을 포함하는 분리된 폴리뉴클레오티드를 제공한다. 폴리뉴클레오티드는 업계에 알려져 있는 방

법 중 어떤 것에 의해서도 최적화되는 인간 코돈일 수 있다.

[0108] 게다가, 본 발명은, 예를 들어, SEQ ID NO:13, 14, 15, 16, 17, 18, 19, 20, 45, 46, 47, 49, 50, 51, 52, 53, 54에서 정의된 서열과 적어도 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 100% 또는 그 사이의 어떤 양이 동일한 핵산에 의해 암호화된 로타바이러스 구조 단백질을 포함하는 RLP를 제공한다.

[0109] 이러한 서열 유사성 또는 동일성은 DNASIS 내에서 제공된 것과 같은, 뉴클레오티드 서열 비교 프로그램을 사용하여 결정될 수도 있다 (예를 들어, 다음 파라미터를 사용하지만, 이에 제한되지 않는다: GAP 패널티(GAP penalty) 5, 대각선 상단의 # 5, 고정된 GAP 패널티 10, k 투플(tuple) 2, 플로팅 갭(floating gap) 10, 및 창크기 5). 하지만, 비교를 위한 서열 정렬의 다른 방법은 업계에 잘 알려져 있으며, 예를 들어, Smith & Waterman (1981, Adv. Appl. Math. 2:482), Needleman & Wunsch (J. Mol. Biol. 48:443, 1970), Pearson & Lipman (1988, Proc. Nat'l. Acad. Sci. USA 85:2444)의 알고리즘, 및 이 알고리즘의 컴퓨터된 구현 (GAP, BESTFIT, FASTA, 및 BLAST, NIH를 통해 이용 가능함), 또는 수동 정렬 및 외관 검사 (예를 들어, Current Protocols in Molecular Biology, Ausubel et al., eds. 1995 supplement를 참고하면 된다), 또는 염증 조건 하에 서던(Southern) 또는 노던(Northern) 잡종화의 사용 (Maniatis et al., in Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory, 1982를 참고하면 된다)가 있다. 바람직하게, 실질적으로 상동성인 서열은 분자의 정의된 길이에 걸쳐 적어도 약 80% 및 가장 바람직하게는 적어도 약 90% 서열 유사성을 나타낸다.

[0110] 이러한 하나의 염증 잡종화 조건의 예는 65°C에서 4X SSC에서 하룻밤 동안 (약 16-20시간) 잡종화에 이은, 65°C에서 한 시간 동안 0.1 X SSC로 세척, 또는 65°C에서 각각 20 또는 30분 동안 0.1 X SSC로 2회 세척일 수도 있다. 대안으로, 예시적 염증 잡종화 조건은 42°C에서 50% 포름아미드, 4 X SSC에서 하룻밤 동안 (16-20시간), 이어서 65°C에서 한 시간 동안 0.1 X SSC로 세척, 또는 65°C에서 각각 20 또는 30분 동안 0.1 X SSC로 2회 세척, 또는 50°C에서 각각 20 또는 30분 동안 0.1 X SSC, 0.1% SDS로 2회 세척되거나, 또는 독특한 서열 영역에 대하여 65°C에서 각각 20 또는 30분 동안 2 X SSC, 0.1% SDS로 2회 세척되는, 하룻밤 동안 (16-20시간), 또는 처치 (Church) 수성 포스페이트 베퍼 (7% SDS; 0.5M NaPO4 베퍼 pH 7.2; 10 mM EDTA) 65°C에서 잡종화일 수 있다.

[0111] 로타바이러스 구조 폴리펩티드를 암호화하는 핵산은 "로타바이러스 핵산", "로타바이러스 뉴클레오티드 서열", "로타바이러스 핵산", 또는 "로타바이러스 뉴클레오티드 서열"로서 설명될 수도 있다. 제한하는 것으로 간주되는 것은 아니지만, 예를 들어, 하나 이상의 로타바이러스 구조 단백질 또는 로타바이러스 구조 폴리펩티드를 포함하는 바이러스-유사 입자는 "로타바이러스 VLP", "RVLP" 또는 "RLP"로서 설명될 수도 있다.

[0112] 많은 유기체들은 증가하는 웹티드 사슬에서 특정 아미노산의 삽입물을 암호화하기 위해 특정 코돈에 사용되는 바이어스(bias)를 나타낸다. 코돈 선호도 또는 코돈 바이어스, 유기체 간에 코돈 사용의 차이는 유전 암호의 퇴화에 의해 제공되고, 많은 유기체 사이에서 문서화된다. 코돈 바이어스는 종종 메신저 RNA (mRNA)의 번역의 효율과 연관성이 있으며, 이것은 그 중에서도 번역되고 있는 코돈의 속성 및 특정 전이 RNA (tRNA) 분자의 이용 가능성에 의존하는 것으로 차례로 생각된다. 세포에서 선택된 tRNA의 우세는 일반적으로 웹티드 합성에서 가장 빈번하게 사용되는 코돈의 반영이다. 따라서, 유전자는 코돈 최적화에 기초하여 특정 유기체에서 최적의 유전자 발현에 맞춰질 수도 있다. 이종 기원의 방식으로 발현된 단백질을 암호화하는 뉴클레오티드 서열을 최적화하는 공정은 발현 수율을 개선하는데 중요한 단계일 수 있다. 최적화 요건은 외부 단백질을 생산하는 숙주의 능력을 개선하는 단계를 포함할 수도 있다.

[0113] "코돈 최적화"는 고유 서열의 코돈 중 적어도 하나, 하나 이상, 또는 많은 수를 또 다른 유기체 또는 종의 유전자에서 더 빈번하게 또는 가장 빈번하게 사용될 수도 있는 코돈으로 대체함으로써 원하는 세포에서 향상된 발현을 위해 핵산 서열을 변형시키는 것으로 정의된다. 다양한 종은 특정 아미노산의 특정 코돈에 대한 특정 바이어스를 나타낸다.

[0114] 본 발명은 코돈 최적화된 합성 폴리뉴클레오티드 서열을 포함하는데, 예를 들어, 서열은 인간 코돈 사용 또는 식물 코돈 사용에 최적화되었다. 그 다음에 코돈 최적화된 폴리뉴클레오티드 서열은 식물에서 발현될 수도 있다. 더 구체적으로, 인간 코돈 사용 또는 식물 코돈 사용에 최적화된 서열은 식물에서 발현될 수도 있다. 어떤 이론에도 결부되지 않으면서, 인간 코돈에 최적화된 서열은 서열의 구아닌-시토신 함량 (GC 함량)을 증가시키고 식물의 발현 수율을 개선하는 것으로 생각된다.

[0115] 번역으로 비효율적인 단백질 암호화 영역의 번역 동역학을 개선하기 위한 다른 코돈-최적화 기술이 업계에 알려

져 있다. 이 기술들은 주로 특정 숙주 유기체에 대한 코돈 사용을 확인하는 것에 의존한다. 특정 유전자 또는 서열이 이 유기체에서 발현되어야 하는 경우에, 이러한 유전자 및 서열의 암호화 서열은 그 다음에 변형될 것이고 따라서 원하는 서열의 코돈을 숙주 유기체의 더 빈번하게 사용되는 코돈에 의해 대체할 것이다.

[0116] 로타바이러스 구조 단백질 또는 폴리펩티드는 바이러스 기반 DNA 또는 RNA 발현 시스템, 제한은 아니지만, 예를 들어, 코모바이러스-기반 발현 카세트 및 제미니바이러스-기반 증폭 요소를 포함하는 발현 시스템에서 발현될 수도 있다.

[0117] 본원에서 설명된 발현 시스템은 이분화(bipartite) 바이러스, 또는 이분화 게놈을 가진 바이러스를 기반으로 하는 발현 카세트를 포함할 수도 있다. 예를 들어, 이분화 바이러스는 코모비리대(Comoviridae) 과일 수도 있다. 코모비리대 과의 속은 코모바이러스, 네포바이러스(Nepovirus), 파바바이러스(Fabavirus), 체라바이러스(Cheravirus) 및 사드와바이러스(Sadwavirus)를 포함한다. 코모바이러스는 동부콩 모자이크 바이러스(CPMV), 동부콩 중증 모자이크 바이러스 (Cowpea severe mosaic virus; CPSMV), 호박 모자이크 바이러스 (Squash mosaic virus; SqMV), 붉은 클로버 반점 바이러스 (Red clover mottle virus; RCMV), 콩껍질 반점 바이러스 (Bean pod mottle virus; BPMV), 순무 원형반점 바이러스 (Turnip ringspot virus; TuRSV), 잠두 진성 모자이크 바이러스 (Broad bean true mosaic virus; BBtMV), 잠두 염색 바이러스 (Broad bean stain virus; BBSV), 무 모자이크 바이러스 (Radish mosaic virus; RaMV)를 포함한다.

[0118] 본 발명의 다양한 양태에 유용할 수도 있는 인핸서 요소를 포함하는 코모바이러스 RNA-2 서열의 예는 CPMV RNA-2 (GenBank 수납 번호 NC_003550), RCMV RNA-2 (GenBank 수납 번호 NC_003738), BPMV RNA-2 (GenBank 수납 번호 NC_003495), CPSMV RNA-2 (GenBank 수납 번호 NC_003544), SqMV RNA-2 (GenBank 수납 번호 NC_003800), TuRSV RNA-2 (GenBank 수납 번호 NC_013219.1), BBtMV RNA-2 (GenBank 수납 번호 GU810904), BBSV RNA2 (GenBank 수납 번호 FJ028650), RaMV (GenBank 수납 번호 NC_003800)를 포함하지만, 이에 제한되지 않는다.

[0119] 이분화 코모바이러스 RNA 게놈의 세그먼트는 RNA-1 및 RNA-2로 불린다. RNA-1은 복제에 수반된 단백질을 암호화하는 한편 RNA-2는 세포-대-세포 이동에 필요한 단백질 및 두 개의 캡시드 단백질을 암호화한다. CPMV, CPSMV, SqMV, RCMV, 또는 BPMV를 포함하는 어떤 적합한 코모바이러스-기반 카세트도 사용될 수 있는데, 예를 들어, 발현 카세트는 CPMV에 기초할 수도 있다.

[0120] "발현 카세트"는 숙주 세포에서 원하는 핵산의 전사를 위해 적절한 프로모터 또는 다른 조절 요소의 조절 하에, 및 이것들에 작동 가능하게 결합된, 원하는 핵산을 포함하는 뉴클레오티드 서열을 나타낸다.

[0121] 발현 시스템은 또한 제미니바이러스의 증폭 요소, 예를 들어, 콩 황반 난장이 바이러스(BeYDV)의 증폭 요소를 포함할 수도 있다. BeYDV는 쌍자엽 식물에 적용된 매스트레바이러스(Mastrevirus) 속에 속한다. BeYDV는 한 가닥 원형 DNA 게놈을 갖는 단립형(monopartite)이고 회전환 메커니즘에 의해 매우 많은 카피 수로 복제할 수 있다. BeYDV-유래된 DNA 래플리콘 벡터 시스템은 식물에서 신속한 고수율 단백질 생산에 사용되어 왔다.

[0122] 본원에서 사용된 바와 같이, 구절 "증폭 요소"는 제미니바이러스 게놈의 하나 이상의 긴 유전자간 영역 또는 긴 유전자간 반복 (LIR) 중 적어도 일부를 포함하는 핵산 세그먼트를 나타낸다. 본원에서 사용된 바와 같이, "긴 유전자간 영역" 또는 "긴 유전자간 반복"은 제미니바이러스 Rep 단백질에 의한 절제 및 복제를 막개할 수 있는 rep 결합 부위를 함유하는 긴 유전자간 영역의 영역을 나타낸다. 일부 양태에서, 하나 이상의 LIR을 포함하는 핵산 세그먼트는 제미니바이러스 게놈의 짧은 유전자간 영역 또는 작은 유전자간 영역 (SIR)을 더 포함할 수도 있다. 본원에서 사용된 바와 같이, "짧은 유전자간 영역" 또는 "작은 유전자간 영역"은 상보성 가닥 (매스트레바이러스의 짧은 IR (SIR))을 나타낸다. 어떤 적합한 제미니바이러스-유래된 증폭 요소도 본원에서 사용될 수 있다. 예를 들어, 제WO2000/20557호; 제WO2010/025285호; Zhang X. et al. (2005, Biotechnology and Bioengineering, Vol. 93, 271-279), Huang Z. et al. (2009, Biotechnology and Bioengineering, Vol. 103, 706-714), Huang Z. et al. (2009, Biotechnology and Bioengineering, Vol. 106, 9-17)을 참고하면 되고; 이것들은 본원에 참고로 포함된다. 하나 이상의 LIR, 예를 들어, 두 개의 LIR이 구조에서 사용되면, 그 다음에 프로모터, CMPV-HT 영역 및 원하는 핵산 서열 및 종결자가 두 개의 LIR 각각에 의해 나누어진다. 게다가, 증폭 요소는, 예를 들어, Halley-Stott et al. (2007) Archives of Virology 152: 1237-1240에서 개시된 서열로부터 유래할 수도 있고, GenBank 수납 번호 DQ458791 하에 기탁되며, 이것은 본원에 참고로 포함된다. LIR을 포함하는 핵산 세그먼트는 뉴클레오티드 2401 내지 2566 및 1 내지 128에 연결된다. SIR을 포함하는 핵산 세그먼트는 뉴클레오티드 1154 내지 1212이다.

[0123] 본원에서 설명된 바와 같이, 니코티아나 벤타미아나 잎의 아그로인필트레이션에 의한 콩 황반 난장이 바이러스

(bean yellow dwarf virus; BeYDV)-유래된 벡터 및 Rep/RepA-공급 벡터의 동시-전달은 효과적인 레플리콘 (replicon) 증폭 및 활발한 단백질 생산을 일으킨다.

[0124] 코모바이러스-기반 발현 카세트 및 제미니바이러스(geminivirus)-유래된 증폭 요소는 별개의 벡터에서 포함될 수도 있거나, 구성요소 일부는 한 벡터에 포함될 수고 있다. 두 개의 벡터가 사용되면, 제1 및 제2 벡터는 식물 세포로 동시에, 또는 별도로 도입될 수도 있다.

[0125] 바이러스 레플리카제는 또한 원하는 핵산의 발현을 증가시키기 위해 본원에서 설명된 발현 시스템에 포함될 수도 있다. 레플리카제의 비-제한 예는 BeYDV Rep 및 RepA를 암호화하는 BeYDV 레플리카제 (pREPHO)이다 (C2/C1; Huang et al., 2009, Biotechnol. Bioeng. 103, 706-714; 이것은 본원에 참고로 포함된다). 레플리카제의 또 다른 비-제한 예는 Halley-Stott et al. (2007) Archives of Virology 152: 1237-1240에서 개시되고, GenBank 수납 번호 DQ458791 하에 기탁되며, 이것은 본원에서 참고로 포함된다. C1:C2 유전자를 포함하는 핵산 세그먼트는 뉴클레오티드 1310 내지 2400이다.

[0126] "동시-발현된"은 둘 또는 그 이상의 뉴클레오티드 서열이 식물 내에서, 및 식물의 같은 조직 내에서 거의 동시에 발현된다는 것을 의미한다. 하지만, 뉴클레오티드 서열은 정확하게 동시에 발현될 필요는 없다. 오히려, 둘 이상의 뉴클레오티드 서열은 암호화된 생성물이 상호작용하는 기회를 갖는 방식으로 발현된다. 둘 또는 그 이상의 뉴클레오티드 서열은 일과성 발현 시스템을 사용하여 동시-발현될 수 있으며, 둘 이상의 서열이 식물 내에서 거의 동시에 두 개의 서열이 발현되는 조건 하에서 도입된다. 대안으로, 뉴클레오티드 서열 중 하나를 포함하는 플랫폼(platform) 식물은 안정한 방식으로 형질전환될 수도 있으며, 추가의 서열은 원하는 단백질, 예를 들어, 하나 이상의 로타바이러스 구조 단백질을 암호화하고, 일시적 방식으로 플랫폼 식물에 도입된다.

[0127] 단백질의 올바른 폴딩(folding)은 단백질의 안정성, 멀티머의 형성, RLP의 형성 및 기능에 중요할 수도 있다. 단백질의 폴딩은 단백질의 서열, 단백질의 상대적 빈도, 세포 내 과밀(crowding)의 정도, 폴딩된, 부분적으로 폴딩된 또는 폴딩되지 않은 단백질과 결합하거나 이것들과 일시적으로 결합될 수도 있는 공통인자의 이용 가능성을 포함하지만, 이에 제한되지 않는, 하나 이상의 인자들에 의해 영향을 받을 수도 있다. 게다가 단백질이 발현되는 식물 내 구획 또는 하위-구획은 단백질의 발현 수준 및 폴딩에 영향을 미칠 수도 있다.

[0128] 하나 이상의 로타바이러스 구조 단백질의 발현은 유전자 이식 식물에서 아그로침투(agroinfiltration)에 의해 특이적 식물 세포 구획 및/또는 하위-구획으로 표적화될 수도 있다. 구획 또는 하위-구획은, 예를 들어, 색소체, 소포체 (ER), 엽록체 또는 아포플라스트일 수도 있다. 어떤 이론에도 결부되지 않으면서, 구획 또는 하위-구획 표적화는 세포질 축적을 통해 표적화된 구획 또는 하위-구획으로의 단백질 축적을 증가시킬 수도 있다. 구획 또는 하위-구획 축적은 세포질에 존재하는 프로테아제에 의한 분해로부터 단백질을 보호하고 및/또는 식물 세포의 기능에 영향을 미치지 않으면서 그것을 더 높은 농도로 축적되게 할 수도 있다.

[0129] 그러므로, 발현 카세트 또는 벡터는 식물에서 벡터 또는 로타바이러스 구조 단백질 또는 벡터로부터 발현된 폴리펩티드를 원하는 구획 또는 하위-구획으로 보내도록 적용될 수도 있다.

[0130] 예를 들어, 발현 카세트 또는 벡터는 발현된 로타바이러스 구조 단백질 또는 폴리펩티드가 색소체의 틸라코이드 막(thylakoid membrane)과 상호작용할 수 있는 부분, 특히 틸라코이드 막의 전송 메커니즘을 포함하게 함으로써 표적 색소체에 적용될 수도 있다. 이 상호작용은 로타바이러스 구조 단백질 또는 폴리펩티드가 그것이 발현되는 세포질의 색소체로 유입되게 할 수도 있다. 어떤 이론에도 결부되지 않으면서, 세포질로부터의 유입 메커니즘은 단백질의 적절한 폴딩에 중요할 수도 있다. 발현 카세트 또는 벡터가 형질전환되는 색소체 자체를 표적화하는데 적용될 수도 있고 로타바이러스 구조 단백질 또는 폴리펩티드의 발현이 전적으로 색소체 내에서 발생할 수도 있다는 것이 인정될 것이다.

[0131] 용어 "표적화 서열"은 표적화 서열이 벡터 또는 발현 카세트에 포함될 수도 있다는 것을 의미한다. 이러한 표적화 서열은 벡터 또는 이것의 생성물을, 색소체와 같이, 식물에서 원하는 구획 또는 하위-구획으로 보내는 웨პ티드로 번역될 수도 있다. 예를 들어, 단백질을 색소체로 표적화하기 위한 색소체 신호 웨პ티드 (업계에서는 "색소체 통과 웨პ티드"로도 불림)는 업계에 알려져 있다. 사용될 수도 있는 색소체 통과 웨პ티드의 비 제한 예는 rbcS1-cTP의 것이다. 엽록체-통과 웨პ티드 서열의 적합한 예는, 예를 들어, 감자의 Rubisco 작은-서브유닛 유전자 (rbcS1)이다.

[0132] 그러므로, 로타바이러스 구조 단백질 또는 폴리펩티드는 폴리펩티드 또는 단백질의 나머지와 같거나, 또는 이것과 이종 기원인 신호 웨პ티드를 포함할 수도 있다. 용어 "신호 웨პ티드"는 업계에 잘 알려져 있고 일반적으로 아미노산의 짧은 (약 5-30개의 아미노산) 서열을 나타내며, 일반적으로 새로 번역된 폴리펩티드의 특정 세포 기관

으로의 위치 이동을 지시할 수도 있는 폴리펩티드의 N-말단에서 발견되거나, 다른 것을 기준으로 폴리펩티드 사슬의 특이적 도메인의 위치 결정을 돋는다. 비-제한 예로서, 신호 웨პ티드는 성숙한 단백질, 예를 들어, 제한하는 것으로 간주되는 것은 아니지만, 로타바이러스 구조 단백질의 분할 및 풀딩을 돋기 위해 단백질의 소포체로의 위치 이동을 표적화하고 및/또는 발생기 폴리펩티드의 막-고정 도메인을 기준으로 N-말단 근위 도메인의 위치 결정을 돋는다.

[0133] 신호 웨პ티드 (SP)는 단백질 또는 바이러스 단백질에 고유할 수도 있거나, 신호 웨პ티드는 발현되는 단백질 또는 바이러스 단백질의 1차 서열에 관하여 이종 기원일 수도 있다. 예를 들어, 로타바이러스 구조 단백질의 고유 신호 웨პ티드는 식물 시스템에서 로타바이러스 구조 단백질을 발현하는데 사용될 수도 있다.

[0134] 신호 웨პ티드는 또한, 예를 들어, 단백질, 바이러스 단백질 또는 로타바이러스 이외의 바이러스의 고유 구조 단백질, 또는 식물, 동물 또는 박테리아 폴리펩티드로부터 고유하지 않을 수도 있다. 사용될 수도 있는 신호 웨პ티드의 비 제한 예는 알팔파(alfalfa) 단백질 이황화 이소머라제 (PDISP) (수납 번호 Z11499의 뉴클레오티드 32-103)의 것이다. 게다가, 신호 웨პ티드는 완전히 결실되거나 절단될 수도 있다. 절단 또는 절단된 아미노산 잔기의 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100% 또는 그 사이의 어떤 양이 신포 웨პ티드로부터 결실된다는 것을 의미한다. 바람직하게, 절단된 아미노산 잔기는 연속적이고, 절단은 제2 메티오닌으로부터 전방으로 발생한다.

[0135] 그러므로 본 발명은, 예를 들어, VP2, VP4, VP6 및/또는 VP7과 같은 로타바이러스 구조 단백질을 제공하며, 고유, 비-고유 신호 웨პ티드 또는 절단된 신호 웨პ티드, 및 이러한 로타바이러스 구조 단백질을 암호화하는 핵산을 포함한다.

[0136] 본 발명의 하나 또는 그 이상의 유전적 구조는 본 발명의 뉴클레오티드 서열 또는 구조, 또는 벡터에 의해 형질 전환되는 어떤 적합한 식물 숙주에서도 발현될 수 있다. 적합한 숙주의 예는 알팔파, 캐놀라(canola), 브라시카 속(*Brassica* spp.), 옥수수, 니코티아나 속(*Nicotiana* spp.), 감자, 인삼, 완두콩, 귀리, 쌀, 대두, 밀, 보리, 해바라기, 목화 등을 포함하는 농작물을 포함하지만, 이에 제한되지 않는다.

[0137] 로타바이러스 구조 단백질을 암호화하는 뉴클레오티드 서열은 1, 2, 3, 4 또는 5개의 바이너리(binary) 플라스미드 벡터를 사용하여 식물 숙주로 전송될 수도 있다. 그러므로 각각의 바이너리 플라스미드 벡터는 로타바이러스 구조 단백질을 암호화하는 1, 2, 3, 4 또는 5개의 뉴클레오티드 서열을 함유할 수도 있다.

[0138] 본 발명의 하나 이상의 유전적 구조는 3' 번역되지 않은 영역을 더 포함할 수도 있다. 3' 번역되지 않은 영역은 폴리아데닐화 신호 및 mRNA 가공 또는 유전자 발현에 영향을 미칠 수 있는 어떤 다른 조절 신호를 함유하는 DNA 세그먼트를 포함하는 유전자의 일부를 나타낸다. 폴리아데닐화 신호는 보통 mRNA 전구체의 3' 끝으로의 폴리아데닐산 트랙의 추가에 영향을 미치는 것을 특징으로 한다. 폴리아데닐화 신호는 보통 기본형 5'-AATAAA-3'에 대한 상동성의 존재에 의해 인식되지만 변화가 드문 일은 아니다. 적합한 3' 영역의 비-제한 예는 노팔린 신타제 (NOS) 유전자와 같은 아그로박테리움 종양 유발 (Ti) 플라스미드 유전자, 대두 저장 단백질 유전자와 같은 식물 유전자, 리볼로스-I의 작은 서브유닛, 5-비스포스페이트 카르복실라제 유전자 (ssPvUBISCO; 제US 4,962,028호; 이것은 본원에 참고로 포함된다), 제US 7,125,978호에서 설명된, 플라스토시아닌 발현을 조절하는데 사용된 프로모터 (이것은 본원에 참고로 포함된다)의 폴리아데닐화 신호를 함유하는 3' 전사된 번역되지 않은 영역이다.

[0139] 본 발명의 유전적 구조 중 하나 이상은 또한, 필요할 수도 있기 때문에, 추가의 인핸서, 번역 또는 전사 인핸서를 포함할 수도 있다. 인핸서는 전사도는 서열의 5' 또는 3'에 자리할 수도 있다. 인핸서 영역은 당업자에게 잘 알려져 있고, ATG 개시 코돈, 인접 서열 등을 포함할 수도 있다. 개시 코돈은, 존재하면, 전사된 서열의 올바른 번역을 제공하기 위해 암호화 서열의 리딩 프레임(reading frame) 상에 있을 수도 있다 ("인 프레임(in frame)").

[0140] 본 발명의 구조는 Ti 플라스미드, Ri 플라스미드, 식물 바이러스 벡터, 직접적인 DNA 형질전환, 미량 주사 (micro-injection), 전기천공(electroporation), 등을 사용하여 식물 세포 내로 도입될 수 있다. 이러한 기술의 재검토를 위해서는, 예를 들어, Weissbach and Weissbach, Methods for Plant Molecular Biology, Academy Press, New York VIII, pp. 421-463 (1988); Geierson and Corey, Plant Molecular Biology, 2d Ed. (1988); 및 Miki and Iyer, Fundamentals of Gene Transfer in Plant를 참고하면 된다. 식물 대사에서, 2d Ed. DT. Dennis, DH Turpin, DD Lefebvre, DB Layzell (eds), Addison Wesley, Langmans Ltd. London, pp. 561-579 (1997). 다른 방법들은 직접적인 DNA 흡수, 리포솜의 사용, 예를 들어, 원생동물을 사용하는 전기천공, 미량 주사, 미세투사물(microparticle) 또는 위스커(wisker), 및 진공 침투를 포함한다. 예를 들어, Bilang, et

al. (Gene 100: 247-250 (1991), Scheid et al. (Mol. Gen. Genet. 228: 104-112, 1991), Guerche et al. (Plant Science 52: 111-116, 1987), Neuhause et al. (Theor. Appl Genet. 75: 30-36, 1987), Klein et al., Nature 327: 70-73 (1987); Howell et al. (Science 208: 1265, 1980), Horsch et al. (Science 227: 1229-1231, 1985), DeBlock et al., Plant Physiology 91: 694-701, 1989), Methods for Plant Molecular Biology (Weissbach and Weissbach, eds., Academic Press Inc., 1988), Methods in Plant Molecular Biology (Schuler and Zielinski, eds., Academic Press Inc., 1989), Liu and Lomonossoff (J Virol Meth, 105:343-348, 2002,), 미국 특허 번호 제4,945,050호; 제5,036,006호; 및 제5,100,792호, 1995년 5월 10일에 출원된 미국 특허 출원 일련 번호 제08/438,666호, 및 1992년 9월 25일에 출원된 제07/951,715호 (이것들 모두는 본원에서 참고로 포함된다)를 참고하면 된다.

[0141] 일과성 발현

[0142] 어떤 이론에도 결부되지 않으면서, 다른 로타바이러스 구조 단백질의 단백질 농도 및 비율은 RLP의 조립 효율에 중요할 수도 있다. 그러므로 감염의 다양성 및 시기는 식물에서 RLP의 단백질 농도 및 전체적인 조립 효율을 조작하는데 중요할 수도 있다.

[0143] 본 발명의 구조는 식물 또는 식물의 일부에서 일과성으로 발현될 수도 있다. 식물, 식물의 일부 또는 식물 세포에서 재조합 아그로박테리움 투메파시엔스의 에피염색체(epichromosomal) 발현에 의존적인 일과성 발현은 로타바이러스 구조 단백질을 발현하는데 사용될 수도 있으며, 다양한 세포 구획 또는 하위-구획에 표적화된다. 일과성 발현 시스템은 높은 생산 속도를 허용한다. 게다가, 대량의 단백질이 식물에서 재조합 아그로박테리움의 침투 후 몇 일 내에 얻어질 수 있다 (Rybicki, 2010; Fischer et al., 1999). 또한 긴 유전자 서열을 발현하고 같은 세포에서 동시에 발현된 하나 이상의 유전자를 갖는 것이 가능하며, 멀티머 단백질의 효율적인 조립을 허용한다 (Lombardi et al., 2009).

[0144] 로타바이러스 구조 단백질을 암호화하는 뉴클레오티드 서열은 1, 2, 3, 4 또는 5개의 형질전환된 아그로박테리움 투메파시엔스 균주의 식물 숙주로 전송될 수도 있다.

[0145] 하지만, 일과성 발현 중에 전사 후 유전자 침묵은 식물에서 이종 기원 단백질의 발현을 제한할 수도 있다. 침묵의 억제인자, 예를 들어, 제한은 아니지만, 토마토 반점 위조 바이러스(Tomato spotted wilt virus)의 NSS의 동시 발현은 이식유전자 mRNA의 특이적 분해에 대응하기 위해 사용될 수도 있다 (Brigneti et al., 1998). 침묵의 대체 억제인자, 예를 들어, 제한은 아니지만, HcPro, TEV-p1/HC-Pro (담배 식각 바이러스(Tobacco etch virus)-p1/HC-Pro), BYV-p21, 토마토 덤불 위축 바이러스(Tomato bushy stunt virus)의 p19 (TBSV p19), 토마토 주름 바이러스(Tomato crinkle virus)의 캡시드 단백질 (TCV-CP), 오이 모자이크 바이러스(Cucumber mosaic virus)의 2b; CMV-2b), 감자 바이러스 X의 p25 (PVX-p25), 감자 바이러스 M의 p11 (PVM-p11), 감자 바이러스 S의 p11 (PVS-p11), 블루베리 스코치 바이러스(Blueberry scorch virus)의 p16, (BScV-p16), 시트러스 트리스택사 바이러스(Citrus tristeza virus)의 p23 (CTV-p23), 포도덩쿨 잎말이병-관련 바이러스(Grapevine leafroll-associated virus)-2의 p24 (GLRaV-2 p24), 포도덩쿨 바이러스 A의 p10 (GVA-p10), 포도덩쿨 바이러스 B의 p14 (GVB-p14), 헤라클리움 잡복 바이러스(Heracleum latent virus)의 p10 (HLV-p10), 또는 마늘 일반 잡복 바이러스(Garlic common latent virus)의 p16 (GCLV-p16)가 업계에 잘 알려져 있고 본원에서 설명된 바와 같이 사용될 수도 있다 (Chiba et al., 2006, Virology 346:7-14; 이것은 본원에 참고로 포함된다). 그러므로, 침묵의 억제인자, 예를 들어, HcPro, TEV-p1/HC-Pro, BYV-p21, TBSV p19, TCV-CP, CMV-2b, PVX-p25, PVM-p11, PVS-p11, BScV-p16, CTV-p23, GLRaV-2 p24, GBV-p14, HLV-p10, GCLV-p16 또는 GVA-p10은 식물 또는 식물의 일부 내에서 높은 수준의 단백질 생산을 더 보장하기 위해 하나 이상의 로타바이러스 구조 단백질, 예를 들어, VP2, VP4, VP6, 또는 이것의 조합과 함께 동시-발현될 수도 있다.

[0146] 본 발명은 또한 상기 설명된 방법을 제공하며, 추가의 (제2, 제3, 제4, 또는 제5) 뉴클레오티드 서열은 식물 내에서 발현되고, 침묵의 억제인자를 암호화하는 추가의 (제2, 제3, 제4, 또는 제5) 뉴클레오티드 서열은 식물에서 활성인 추가의 (제2, 제3, 제4, 또는 제5) 조절 영역과 작동 가능하게 결합된다. 침묵의 억제인자를 암호화하는 뉴클레오티드 서열은, 예를 들어, Nss, HcPro, TEV-p1/HC-Pro, BYV-p21, TBSV p19, TCV-CP, CMV-2b, PVX-p25, PVM-p11, PVS-p11, BScV-p16, CTV-p23, GLRaV-2 p24, GBV-p14, HLV-p10, GCLV-p16 또는 GVA-p10일 수도 있다.

[0147] 하기 설명된 바와 같이, 일과성 발현 방법은 본 발명의 구조를 발현하는데 사용될 수도 있다 (Liu and Lomonossoff, 2002, Journal of Virological Methods, 105:343-348을 참고하면 되고; 이것은 본원에 참고로 포함된다). 대안으로, Kapila et al., 1997 (이것은 본원에 참고로 포함된다)에 의해 설명된, 진공-기반 일과성

발현 방법이 사용될 수도 있다. 이 방법들은, 예를 들어, 제한은 아니지만, 아그로-접종(Agro-inoculation) 또는 아그로-침투, 주사기 침투의 방법을 포함하지만, 다른 일시적 방법들도 상기 언급된 바와 같이 사용될 수 있다.

[0148] 아그로-접종, 아그로-침투, 또는 주사기 침투로, 원하는 핵산을 포함하는 아그로박테리아의 혼합물이 조직의 세포 내 공간, 예를 들어, 잎, 식물의 공기 중의 부분 (줄기, 잎 및 꽃 등), 식물의 다른 일부 (줄기, 뿌리, 꽃), 또는 전체 식물에 전입한다. 상피를 통과한 후 아그로박테리아는 t-DNA 카페를 감염시키고 이것을 세포로 전송한다. t-DNA는 에피솜에 의해 전사되고 mRNA가 번역되며, 감염된 세포에서 원하는 단백질의 생산으로 이어지지만, 핵 내부에서 t-DNA의 통과는 일시적이다.

[0149] 형질전환된 식물 세포의 확인을 돋기 위해서, 본 발명의 구조는 식물 선택 가능한 마커를 포함하도록 더 조작될 수도 있다. 유용한 선택 가능한 마커는 항생제, 예를 들어, 젠타마이신, 히그로마이신, 카나마이신, 또는 포스피노트리신, 글리포세이트, 클로로솔푸론, 등과 같은 제초제와 같은 화학물질에 대한 내성을 제공하는 효소를 포함한다. 유사하게, GUS (베타-글루쿠로니다제)와 같이 색깔 변화에 의해, 또는 루시퍼라제 또는 GFP와 같이 발광(luminescence)에 의해 확인 가능한 화합물의 생산을 제공하는 효소가 사용될 수도 있다.

[0150] 또한 본 발명의 고려된 부분은 본 발명의 유전자 구조를 함유하는 유전자 이식 식물, 식물 세포 또는 씨앗이다. 식물 세포로부터 전체 식물을 재생하는 방법이 또한 업계에 알려져 있다. 일반적으로, 형질전환된 식물 세포는 적절한 배지에서 배양되는데, 이것은 항생제와 같은 선택적 약제를 함유할 수도 있으며, 선택 가능한 마커는 형질전환된 식물 세포의 확인을 용이하게 하는데 사용된다. 캘러스(callus)가 형성되면, 순(shoot) 형성은 알려진 방법에 따라 적절한 식물 호르몬을 이용함으로써 촉진될 수도 있고 순은 식물의 재생을 위해 발근 배지로 옮겨진다. 그 다음에 식물은, 종자로부터 또는 식물 번식 기술을 사용하여, 반복적 세대를 확립하는데 사용될 수도 있다. 유전자 이식 식물은 조직 배양을 사용하지 않고도 생성될 수 있다.

[0151] 본 출원에서 용어 "조절 영역", "조절 요소" 또는 "프로모터"의 사용은 항상 그렇지는 않지만, 전형적으로, 유전자의 단백질 암호화 영역의 업스트림에 있는 핵산의 일부를 반영한다는 것을 의미하며, 이것은 DNA 또는 RNA, 또는 DNA 및 RNA 둘 다로 구성될 수도 있다. 조절 영역이 원하는 유전자에서 활성이고, 이것과 작동 가능하게 연결되거나, 작동 가능하게 결합될 때, 이것은 원하는 유전자의 발현을 일으킬 수도 있다. 조절 요소는 기관 특이성을 매개하거나, 또는 발달적 또는 일시적 유전자 활성화를 제어할 수도 있다. "조절 영역"은 프로모터 요소, 기초 프로모터 활성을 나타내는 코어 프로모터 요소, 외부 자극에 반응하여 유발성 요소, 음성 조절 요소 또는 전사 인핸서와 같이 프로모터 활성을 매개하는 요소를 포함할 수도 있다. "조절 영역"은, 본원에서 사용된 바와 같이, 전사 후 활성인 요소, 예를 들어, 번역 및 전사 인핸서, 번역 및 전사 억제물질, 업스트림 활성화 서열, 및 mRNA 불안정 결정요인과 같이 유전자 발현을 조절하는 조절 요소도 포함할 수 있다. 이 후자의 요소들 중 다수는 암호화 영역의 근위에 자리할 수도 있다.

[0152] 본 개시의 맥락에서, 용어 "조절 요소" 또는 "조절 영역"은 전형적으로, 항상은 아니지만, 보통은, 구조 유전자의 암호화 서열에 대한 업스트림 (5')에 있는 DNA의 서열을 나타내며, 이것은 RNA 폴리머라제 및/또는 특정 부위에서 시작하는 전사에 필요한 다른 인자에 대한 인식을 제공함으로써 암호화 영역의 발현을 제어한다. 하지만, 인트론 내, 또는 서열의 3'에 자리한 다른 뉴클레오티드 서열 또한 원하는 암호화 영역의 발현의 조절에 기여할 수도 있다고 생각되어야 한다. RNA 폴리머라제 또는 특정 부위에서 개시를 보장하는 다른 전사 인자에 대한 인식을 제공하는 조절 요소의 예는 프로모터 요소이다. 모두는 아니지만, 대부분의 진핵생물 프로모터 요소는 TATA 상자, 보통 전사 시작 부위의 업스트림에서 대략 25개의 염기쌍에 있는 아데노신 및 티미딘 뉴클레오티드 염기쌍으로 구성된 보존된 핵산 서열을 함유한다. 프로모터 요소는 전사의 개시의 원인이 되는 기본 프로모터 요소, 뿐만 아니라 유전자 발현을 변형시키는 다른 조절 요소 (상기 나열된 바와 같음)를 포함한다.

[0153] 여러 타입의 조절 영역이 있으며, 발달적으로 조절되거나, 유발성 또는 구성적인 것들을 포함한다. 발달적으로 조절되거나, 그것의 제어 하에서 유전자의 차등적 발현을 제어하는 조절 영역은 상기 기관 또는 조직의 발달 중의 특정 시간에 특정 기관 또는 기관의 조직 내에서 활성화된다. 하지만, 발달적으로 조절되는 일부 조절 영역은 우선적으로 특정 발달 단계에서 특정 기관 또는 조직 내에서 활성일 수도 있고, 그것들은 또한 발달적으로 조절된 방식으로, 또는 식물 내 다른 장기 또는 조직의 기저 수준에서 활성화될 수도 있다. 조직-특이적 조절 영역의 예, 예를 들어, 씨앗-특이적 조절 영역은 나핀(napin) 프로모터, 및 크루시페린(cruciferin) 프로모터를 포함한다 (Rask et al., 1998, J. Plant Physiol. 152: 595-599; Bilodeau et al., 1994, Plant Cell 14: 125-130). 잎-특이적 프로모터의 예는 플라토시아닌 프로모터를 포함한다 (제US 7,125,978호를 참고하면 되며, 이것은 본원에 참고로 포함된다).

[0154]

유발성 조절 영역은 인듀서(inducer)에 반응하여 하나 이상의 DNA 서열 또는 유전자의 전사를 직접적으로 또는 간접적으로 활성화시킬 수 있는 것이다. 인듀서의 부재시 DNA 서열 또는 유전자는 전사되지 않을 것이다. 전형적으로 유발성 조절 영역에 특이적으로 결합하여 전사를 활성화시키는 단백질 인자는 불활성 형태로 존재할 수도 있으며, 그 다음에 이것은 인듀서에 의해 활성형으로 직접적으로 또는 간접적으로 전환된다. 하지만, 단백질 인자가 또한 부재일 수도 있다. 인듀서는 단백질, 대사산물, 성장 조절기, 제초제 또는 폐놀 화합물 또는 열, 추위, 염, 또는 독성 요소에 의해 직접적으로 또는 병원체 또는 바이러스와 같은 병원균(disease agent)의 작용을 통해 간접적으로 도입된 생리학적 스트레스와 같은 화학적 약제일 수도 있다. 유발성 조절 영역을 함유하는 식물 세포는 인듀서를 세포 또는 식물에 외부에서 적용함으로써, 예를 들어, 분무(spraying), 살수(watering), 가열 또는 유사한 방법에 의해 인듀서에 노출될 수도 있다. 유발성 조절 요소는 식물 또는 비-식물 유전자로부터 유래될 수도 있다 (예를 들어 Gatz, C. and Lenk, L.R.P., 1998, Trends Plant Sci. 3, 352-358을 참고하면 되고; 이것은 참고로 포함된다). 잠재적 유발성 프로모터의 예는 테트라시클린-유발성 프로모터 (Gatz, C., 1997, Ann. Rev. Plant Physiol. Plant Mol. Biol. 48, 89-108; 이것은 참고로 포함된다), 스테로이드 유발성 프로모터 (Aoyama, T. and Chua, N.H., 1997, 식물 1, 2, 397-404; 이것은 참고로 포함된다) 및 에탄올-유발성 프로모터 (Salter, M.G., et al, 1998, Plant Journal 16, 127-132; Caddick, M.X., et al, 1998, Nature Biotech. 16, 177-180, 이것은 참고로 포함된다), 시토카닌 유발성 IB6 및 CKI1 유전자 (Brandstatter, I. and Kieber, 1.1., 1998, Plant Cell 10, 1009-1019; Kakimoto, T., 1996, Science 274, 982-985; 이것은 참고로 포함된다) 및 옥신 유발성 요소, DR5 (Ulmasov, T., et al., 1997, 식물 세포 9, 1963-1971; 이것은 참고로 포함된다)를 포함하지만, 이에 제한되지 않는다.

[0155]

구성적 조절 영역은 식물의 다양한 부분을 통해 및 지속적으로 식물 발달을 통해 유전자의 발현을 나타낸다. 알려진 구성적 조절 요소의 예는 CaMV 35S 전사물 (Odell et al., 1985, Nature, 313: 810-812), 쌀 액틴 1 (Zhang et al, 1991, Plant Cell, 3: 1155-1165), 액틴 2 (An et al., 1996, Plant J., 10: 107-121), 또는 tms 2 (미국 제5,428,147호, 이것은 본원에 참고로 포함된다), 및 트리오스포스페이트 이소머라제 1 (Xu et. al., 1994, Plant Physiol. 106: 459-467) 유전자, 옥수수 유비퀴틴 1 유전자 (Cornejo et al, 1993, Plant Mol. Biol. 29: 637-646), 아라비돕시스(Arabidopsis) 유비퀴틴 1 및 6 유전자 (Holtorf et al, 1995, Plant Mol. Biol. 29: 637-646), 및 담배 번역 개시 인자 4A 유전자 (Mandel et al, 1995, Plant Mol. Biol. 29: 995-1004)와 관련된 프로모터를 포함한다.

[0156]

용어 "구성적"은 본원에서 사용된 바와 같이 반드시 유전자가 구성적 조절 영역의 제어 하에 모든 세포 타입에서 같은 수준으로 발현된다는 것을 나타내지는 않지만, 풍부함의 변화를 통해 광범위한 세포 타입에서 발현된다는 것이 종종 관찰된다. 구성적 조절 요소는 다른 서열과 커플링되어(coupled) 작동 가능하게 결합되는 뉴클레오티드 서열의 전사 및/또는 번역을 더 향상시킬 수도 있는데, 예를 들어, CPMV-HT 시스템은 동부콩 모자이크 바이러스 (CPMV)의 번역되지 않은 영역으로부터 유래되고 관련된 암호화 서열의 향상된 번역을 입증한다. "고유한"은 핵산 또는 아미노산 서열이 자연 발생하거나, 또는 "야생형"인 것을 의미한다. "작동 가능하게 결합된"은 특정 서열, 예를 들어, 원하는 조절 요소 및 암호화 영역이 유전자 발현의 매개 또는 조절과 같이 의도된 기능을 직접적으로 또는 간접적으로 수행하기 위해 상호작용한다는 것을 의미한다. 작동 가능하게 결합된 서열의 상호 작용은, 예를 들어, 작동 가능하게 결합된 서열과 상호작용하는 단백질에 의해 매개될 수도 있다는 것을 의미한다.

[0157]

식물 내에서 생산된 RLP는 식물-특이적 N-글리칸을 포함하는 로타바이러스 VP7 구조 단백질을 유도할 수도 있다. 그러므로, 본 발명은 또한 식물 특이적 N-글리칸을 갖는 VP7을 포함하는 RLP를 제공한다.

[0158]

게다가, 식물에서 N-글리칸의 변형이 알려져 있고 (예를 들어, 미국 제60/944,344호; 이것은 본원에 참고로 포함된다) 변형된 N-글리칸을 갖는 VP7이 생산될 수도 있다. 예를 들어, 푸코실화되거나, 자일로실화되거나, 또는 푸코실화되고 자일로실화된 N-글리칸이 감소된, 변형된 글리코실화 패턴을 포함하는 VP7이 얻어질 수도 있거나, 또는 단백질이 푸코실화, 자일로실화, 또는 둘 다가 부족하고, 증가된 갈락토실화를 포함하는, 변형된 글리코실화 패턴을 갖는 VP7이 얻어질 수도 있다. 게다가, 번역 후 변형, 예를 들어, 말단 갈락토스의 추가의 조절은 VP7을 발현하는 야생형 식물과 비교하여 발현된 VP7의 푸코실화 및 자일로실화의 감소를 일으킬 수도 있다.

[0159]

제한하는 것으로 간주되는 것이 아니라, 예를 들어, 변형된 글리코실화 패턴을 갖는 VP7의 합성은 베타-1.4 갈락토실트랜스퍼라제 (GalT), 제한은 아니지만, 예를 들어, 포유동물 GalT, 또는 인간 GalT를 암호화하는 뉴클레오티드 서열과 함께 VP7을 동시-발현함으로써 달성될 수도 있지만, 또 다른 공급원의 GalT 또한 사용될 수도 있다. GalT의 촉매 작용 도메인이 또한, GNT1-GalT 히브리드(hybrid) 효소를 생산하기 위해, N-아세틸글루코사미닐 트랜스퍼라제 (GNT1)의 CTS 도메인 (즉, 세포질 꼬리, 막관통 도메인, 줄기 영역)에 융합될 수도 있고, 히브

리드 효소는 VP7과 동시-발현될 수도 있다. VP7은 또한 N-아세틸글루코사미닐 트랜스퍼라제 III (GnT-III)를 암호화하는 뉴클레오티드 서열과 함께 동시-발현될 수도 있고, 제한은 아니지만, 예를 들어, 포유동물 GnT-III 또는 인간 GnT-III, 다른 공급원의 GnT-III가 또한 사용될 수도 있다. 추가적으로, GnT-III에 융합된 GNT1의 CTS를 포함하는, GNT1-GnT-III 히브리드 효소가 또한 사용될 수 있다.

[0160] 그러므로 본 발명은 또한 변형된 N-글리칸을 갖는 VP7을 포함하는 RLP를 제공한다.

[0161] 어떤 이론에도 결부되지 않으면서, VP7에서 식물 N-글리칸의 존재는 항원 제공 세포(antigen presenting cell)에 의해 VP7의 결합을 촉진함으로써 면역 반응을 자극할 수도 있다. 식물 N 글리칸을 사용하는 면역 반응의 자극은 Saint-Jore-Dupas et al. (2007)에 의해 제안되었다.

[0162] 본 발명은 다음 실시예에서 더 예시될 것이다.

[0163] 실시예

실시예 1

니코티아나 벤타미아나 식물 잎에서 로타바이러스 단백질의 발현 및 VLP의 생산

[0166] 다음 분석은 G9 P[6] 로타바이러스 균주의 로타바이러스 캡시드 단백질을 이용하였고, 로타바이러스-유사 입자가 담배 니코티아나 벤타미아나 잎 세포의 다양한 구획에서 형성되는지를 평가하였다. 담배 잎에서 VP2 및 VP6, 뿐만 아니라 VP2, VP6, VP7 및 VP4의 다양한 조합의 동시-발현을 조사하였다.

재료 및 방법

플라스미드 구조

[0169] VP2, VP4, VP6 및 VP7에 대하여 식물 코돈 최적화된 로타바이러스 cDNA를 Geneart, Germany에 의해 공급하였다. 플라스미드 DNA를 제조사의 지시에 따라 DH5-α 화학적 컴피턴트(competent) 대장균 세포 (E. cloni™, Lucigen)로 형질전환하였다. Rainer Fischer (Fraunhofer Institute for Molecular Biology and Applied Ecology, IME, Germany)에 의해 공급된 새로운 바이너리 아그로박테리움 벡터 pTRAc (세포질), pTRAc-rbcS1-cTP (엽록체 표적화) 및 pTRAc-ERH (소포체 표적화)를 이 연구에서 사용하였다. 추가적인 벡터 벡터, pTRAc-A (아포플라스트)는 다수의 클로닝 부위의 부위 NcoI 및 XhoI에서 제한 효소 (RE) 분해에 의한 pTRAc-ERH의 변형으로부터 유래된다 (도 3). 이것은 히스티딘 태그 및 ER에서 단백질을 보유하는 KDEL 서열을 제거한다. 단백질은 대신에 아포플라스트에 표적화된다.

[0170] VP2, VP4 및 VP6 cDNA는 NcoI/XhoI로 분해된 제한 효소 (RE)인 한편 VP7은 Af1III/XhoI로 잘린다. 제한 효소 Af1III, NcoI 및 MluI는 호환성 접착성 끝(sticky end)을 갖는다. DNA의 pTRAc, pTRAc-rbcS-cTP 및 pTRAc-A로의 직접적인 클로닝을 위해, 벡터를 각각 부위 Af1III/XhoI, MluI/XhoI 및 NcoI/XhoI에서 각각 RE 분해하였다. 벡터에서 DNA의 클로닝에 이어서 화학적 컴피턴트 대장균 DH5-α 세포 (E. cloni™, Lucigen)로의 형질 전환을 표준 프로토콜에 따라 수행하였다. 선택된 재조합 콜로니를 콜로니 PCR에 의해 확인하였다. pTRAc-ERH에서 클로닝에 대하여, PCR 증폭에 의해 각각 네 개의 로타바이러스 cDNA의 종결 코돈을 대체하기 위해 NotI 제한 효소 부위를 추가하였다. cDNA를 표 1에서 상세히 설명된 프라이머로 증폭하였다. PCR 반응 조건은 95°C에서 5분 동안 변성 후 이어서, 95°C에서 30초 동안 5주기의 변성, 52°C에서 1분 동안 어닐링(annealing), 및 72°C에서 1.5분 동안 신장을 포함한다. 추가의 20주기를 다음과 같이 실행하였다: 95°C에서 30초 동안, 57°C에서 1분 동안, 72°C에서 1.5분 동안 및 72°C에서 5분 동안. 그 다음에 증폭된 단편을 제조사의 지시에 따라 pGEM-T-Easy (Promega)에 클로닝하였다. 형질 전환을 화학적 컴피턴트 대장균 DH5-α (E. cloni™, Lucigen)에서 수행하였다. 그 다음에 콜로니 PCR을 다른 세 개의 구조에 대하여 실행된 바와 같이 선택된 콜로니에서 수행하였다.

표 1

표 1: ER 벡터 클로팅을 위한 로타바이러스 cDNA 프라이머

프라이머	서열	추가된 제한 효소 부위	배향
VP2F	5'-TTCCATGGCTTACCGTAAAAGG-3'	-	SEQ ID NO: 5 정방향
VP2R	5'-ATGC GG CG CA AG CT CG TT CATA AT CCT CAT G-3'	NotI	SEQ ID NO: 6 역방향
VP4F	5'-TTCCATGGCTTCCCTCATCTAC-3'	-	SEQ ID NO: 7 정방향
VP4R	5'-ATGC GG CG CA AG AC GG ACT TG GAG AAT GAG-3'	NotI	SEQ ID NO: 8 역방향
VP6F	5'-TTCCATGGATGTGCTCTACTC-3'	-	SEQ ID NO: 9 정방향
VP6R	5'-ATGC GG CG CC TT CAC GAG C AT GG AAC G-3'	NotI	SEQ ID NO: 10 역방향
VP7F	5'-GTACATGTACGGAATCGAGTAC-3'	-	SEQ ID NO: 11 정방향
VP7R	5'-ATGC GG CG CC AC AC GG TAG TAG AAA AGC AGC-3'	NotI	SEQ ID NO: 12 역방향

[0171]

[0172] 양성 콜로니의 pGEM-VP DNA를 PCR의 정확도를 확인하기 위해 시퀀싱하였다(sequenced). DNA를 NcoI/NotI로 분해하였고 pTRAKc-ERH-VP를 형성하기 위해 적절한 DNA 단편을 부위 NcoI 및 NotI에서 pTRAKc-ERH로 클로닝하였다. 그 다음에 형질전환을 이전에 실행된 바와 같이 대장균 DH5-α 세포로 수행하였다. 선택된 콜로니에서 로타바이러스 DNA를 검사하기 위해 콜로니 PCR을 또한 수행하였다.

[0173]

아그로박테리움 형질전환

[0174] 아그로박테리움 투메파시엔스 GV3101 균주는 Rainer Fischer 교수 (Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Aachen, Germany)에 의해 제공받았고 이전에 설명된 바와 같이 일렉트로컴피던트(electrocompetent)로 만들었다 (Shen and Forde, 1989). 300 나노그램의 분리된 로타바이러스 pTRA-VP 구조를 0.1 cm 일렉트로갭(electrogap) 큐벳 (BioRadTM)에서 100 μl의 일렉트로컴피던트 GV3101 세포와 혼합하였고 그 다음에 GenePulser (BioRad)에서 다음 설정 하에 전기천공하였다: 1.8 kV, 25 μF 및 200 Ω. 50 μg/ml 카르베니실린 (carb), 30 μg/ml 카나마이신 (kan) 및 50 μg/ml 리팜피신 (rif)을 함유하는 LA 플레이트에 플레이팅(plating)하기 전 배양은 900 μl LB에서 27°C에서 1시간 동안 허용되었다. 플레이트를 27°C에서 3일 동안 배양하였다. 양성 형질전환주를 검사하기 위해, 플라스미드 DNA를 재조합 아그로박테리움 콜로니로부터 분리하였고 대장균 컴피던트 DH5-α로 역-형질전환하였다. 이것들을 그 다음에 100 μg/ml 앰피실린 (amp) LA에서 선택하였다. cDNA 상에서 콜로니 PCR 및 제한 효소 분해를 실행하여 성공적인 형질전환주를 확인하였다. 적절한 재조합 아그로박테리움의 글리세롤 스톡(stock)을 만들었고-70°C에서 저장하였다.

[0175]

재조합 아그로박테리움 침투

[0176] 이 연구에 사용된 아그로박테리움 투메파시엔스 LBA 4404 (pBIN-NSs)를 Marcel Prins (Laboratory of Virology, Wageningen University, Binnenhaven, Netherlands)로부터 얻었다. 그것은 토마토 반점 위조 바이러스 (TSWV)에서 발견된 NSs 침묵 억제인자를 함유한다. 글리세롤 스톡의 재조합 아그로박테리움 (pTRA-VPs)을 27 °C에서 하룻밤 동안 50 μg/ml carb, 30 μg/ml kan 및 50 μg/ml rif가 들어있는 LB에서 키웠다. 그 다음에 재조합 아그로박테리움 및 LBA4404 (pBIN-NSs)를 각각 유도 배지 (LB, 10 mM 2-(N-모폴리노) 에탄술폰산 MES, 2 mM MgSO4, 20 μM 아세토시링콘, 50 μg/ml carb, 30 μg/ml kan 및 50 μg/ml rif, 및 pH 5.6)에서 접종하였다.

[0177]

배양물을 27°C 하룻밤 동안 배양하였다. 아그로박테리움 세포를 4°C에서 5분 동안 4000 rpm으로 원심분리에 의해 수거하였고 그 다음에 2 ml 침투 배지 (10 mM MES, 10 mM MgCl2, 3% 수크로스, pH 5.6, 200 μM 아세토시링콘 및 증류수)에서 재현탁하였다. 세포의 흡광도 (OD600)를 확인하였고 0.25의 OD600을 얻기 위해 침투 배지로 희석하였다. 각각의 pTRA-VP 구조에 대하여, LBA4404를 0.5의 최종 OD600으로 재조합 아그로박테리움과 혼합하였다. 동시-발현 연구를 위해, 혼합물에 대한 OD600이 0.5와 같아질 때까지, 각 구조를 총 0.5의 OD600, 예를 들어, VP2-0.25 및 VP6-0.25의 OD600까지 추가하였다. 유도 및 침투 배지에 사용된 아세토시링콘은 아그로박테리움에서 vir 유전자의 활성화를 돋는다.

[0178]

상처난 식물 세포는 아그로박테리움에서 Vir 및 Vir G 유전자에 의해 검출되는 폐놀 화합물을 방출하고 그 후에 속주 세포에서 단백질 유전자의 유도로 이어진다 (Zupan, J. et al., 2000). 그 다음에 세포를 실온에서 1시간 동안 배양하여 아세토시링콘이 vir 유전자를 유발하게 하였다. 3주령 야생형 니코티아나 벤타미아나 식물에 VP 단백질을 발현하는 재조합 아그로박테리움이 침투되었다. 이것은 식물 잎의 복부의 배축 공간(abaxial air

space)으로 전체 식물의 진공 침투 또는 재조합 아그로박테리움 (pTRA-VP)의 주사를 수반한다. 재조합 아그로박테리움은 침묵 억제인자 LBA 4404 (pBIN-NSs)와 함께 또는 없이 침투되었다.

[0179] 초기에는, 아그로박테리움 침투 배지 혼탁액 2 ml을 구조 당 주사를 사용하여 각 식물에 주사하였다. 7일의 타임 트라이얼(time trial) 동안 구조 당 하나의 식물을 사용하였다. 로타바이러스 단백질의 동시-발현을 또한 수행하였으며 여기에서 VP2, VP6 및 VP4가 니코티아나 벤타미아나 식물 잎의 세포질에서 동시에 발현된다. 조사된 조합은 VP2/6 및 VP2/6/4였다. VP4 "스파이크" 단백질은 VP6에 결합할 수도 있고, 따라서 그것들이 RLP 구조에 추가될 가능성이 있다. VP7 클로닝을 시도하였지만, 숙주 세포에 대한 독성 문제는 문제인 것으로 증명되었다. 재조합 VP7 아그로박테리움은 침투 후 하루 내에 잎 세포를 살해하였다. 이것을 피하기 위해 17°C의 저온에서 식물 침투 및 타임 트라이얼의 제3 일 및/또는 제5 일 후 침투와 같이 여러 방법들을 시도하였다. 이와 같이, VP7을 담배에서 그것의 독성 성질로 인해 동시-발현 연구에서 제외하였다.

단백질 추출

[0181] 전체 잎 또는 구조 당 두 개의 잎 디스크(disc)를 수확하였고 액체 질소에서 갈았다. 갈아진 잎 물질을 Complete Protease Inhibitor (EDTA-없음; Roche) 함유 멸균 PBS에서 재현탁하였다. 그 다음에 이것을 13000 rpm으로 5분 동안 원심분리하였고 펠릿(식물 잎 물질)을 제거하였다. 그 다음에 각 구조의 100 μl를 5X SDS-PAGE 로딩 버퍼와 혼합하였고 95°C에서 2분 동안 끓여서, SDS-PAGE 겔 및 웨스턴 블롯에서 추가의 분석을 준비하였다. 샘플의 나머지는 향후의 사용을 위해 -20°C에 저장하였다. 도 4는 로타바이러스 cDNA에 대한 클로닝 및 침투 과정의 개요를 나타낸다.

아포플라스트 단백질 추출

[0183] 추가적인 추출 과정을 pTRAkc-A, 아포플라스트 구조에서 수행하였다. 아포플라스트는 식물 세포의 원형질막과 세포벽 사이의 자유 확산 공간이다 (도 5a). 세포질에서 발현된 단백질은 그것들을 아포플라스트로 표적화하는 수출 서열을 갖고 이런 이유로 그것들은 여기에 축적된다. 이어지는 추출 과정에서, 각 추출일의 전체 잎은 Complete Protease Inhibitor를 함유하는 멸균 PBS와 함께 진공 또는 주사 침투되었다. 진공 침투를 위해, 개개의 식물은 잎을 PBS에 혼탁하였고 진공 탱크에서 10분 동안 100 mbar의 진공 하에 두었다. 그 다음에 잎을 말아서 부드럽게 바닥에 구멍이 있는 스핀 컬럼 (Qiagen 스핀 컬럼과 유사함)에 두었다 (도 5b2). 구멍은 고체 잎 물질이 통과하게 하지 않으면서 잎의 유동체의 쉬운 통과를 허용한다. 스핀 컬럼을 2 ml 에펜도르프 튜브에 배치하였고 원심분리를 15분 동안 4000 rpm으로 수행하였다 (도 5 b3). 여과물을 수거하였고 SDS-PAGE 겔 및 웨스턴 블롯 분석용 단백질 로딩 염료를 각 여과물 샘플의 100 μl에 추가하였다.

웨스턴 블롯 및 쿠마시 염색

[0185] 웨스턴 블롯 및 쿠마시 블루 염색된 SDS-PAGE 겔을 이전에 설명된 바와 같이 사용하였다. 웨스턴 블롯에서 마우스 항-로타바이러스 VP6 항체 (US Biologicals) (1:5000), 항-마우스 히스티딘 태그 항체 (Sigma®) (1:2000), 닦 항-VP2 및 닦 항-VP4 혈청 (1:2000)을 사용하여 각각의 단백질을 검출하였다. Syngene Gel Documentation System을 사용하여 밴드를 밀도 스캐닝함으로써 단백질을 정량하기 위해 쿠마시 블루 염색된 SDS-PAGE 겔을 사용하였다.

전자 현미경 검사

[0187] 발현된 단백질이 RLP로 조립되는지 결정하기 위해서, 침묵 억제인자 Nss의 존재시 세포질 발현된 VP6, VP2/6 및 VP2/6/4, 모두의 발현 제3 일에 면역-포획된(immuno-trapped) 입자의 투과 전자 현미경 검사 (TEM)를 수행하였다. 글로(glow) 방전된 탄소/구리 그리드를 5분 동안 20 μl의 마우스 항-로타바이러스 VP6 항체 (1:5000)에 배치하였고 그 다음에 중류수로 3번 세척하였다. 그 다음에 그리드를 10 μl의 단백질 추출물에 배치하였고 중류수로 3번 세척하기 전 2분 동안 두었다. 최종적으로, 그리드를 TEM (Zeiss 912 OMEGA Energy Filter Transmission Electron Microscope, University of Cape Town)으로 보기 전에 1분 동안 2% 유라닐 아세테이트 20 μl에서 플로팅하였다 (floated).

[0188] 수크로스 구배로부터 분리된 샘플에 대하여, 수크로스를 먼저 구리 그리드에서 면역-포획 전 투석에 의해 제거해야 한다. 제거되지 안으면, 수크로스 결정이 TEM 하에서 샘플의 명확한 시야를 억제하는데 그것이 그리드에서 결정을 형성하여 결합된 탄소 및 재료의 구조를 파괴하기 때문이다. 수크로스 분획을 10 000 MW 투석 카세트에 배치하였고 버퍼를 교환하기 전 4시간 동안 0.4 M NaCl을 함유하는 멸균 PBS에서 투석하였고 교반하면서 4°C에서 하룻밤 동안 두었다. 투석으로 부피가 증가하기 때문에 단백질 샘플은 농축이 필요했다. 샘플을 3시간 동안

진공 동결 건조하였고 멸균 PBS 2 ml에서 재현탁하여 추가의 분석을 준비하였다.

[0189] RLP의 수크로스 구배 정제

[0190] 식물 단백질 추출물을 고체 식물 재료를 제거하기 위해 미라클로스(miracloth)를 통해 처음에 여과하였다. 10 내지 60% 수크로스의 수크로스 구배를 멸균 PBS (pH 7.4)에 용해된 수크로스 5 ml의 여섯 개의 층을 생성함으로써 각각 40ml 퓨브에서 설정하였다. 그 다음에 5 내지 10 ml 부피의 정화된 단백질 샘플을 각 구배 컬럼의 상단에 로딩하였다.

[0191] 150000 g에서 초원심분리 (SWTi28 매달린 버켓 로터(swinging bucket rotor), Beckman Coulter)를 4°C에서 1시간 30분 동안 수행하였다. 원심분리의 끝에, 2 ml 분획을 퓨브 천자에 의해 각 컬럼의 바닥으로부터 수거하였다. 그 다음에 도트 블롯을 수행하여 원하는 단백질을 갖는 분획을 결정하였다. 각 분획에 대하여, 1 μ l의 샘플을 니트로셀룰로스 막 위에 그리드에 로딩하였으며, 그 다음에 이것은 BSA 차단 버퍼로 차단하였다. 그 다음에 웨스턴 블롯 분석을 평소와 같이 수행하였다. 단백질을 VP6에 대하여 마우스 항-VP6 항체 (1:5000) 또는 다른 두 개의 단백질에 대하여 담 항-VP2 및 VP4 혈청 (1:5000)으로 검출하였다.

[0192] 전체 수용성 단백질 검정

[0193] 전체 수용성 단백질 (TSP)을 브래드포드 검정(Bradford assay)으로 결정하였다. 세포질 동시-발현된 VP2/6에서 축적된 단백질의 수준을 비교하기 위해 이것을 수행하였다. 단백질 IgG (1.43mg/ml 스톡)을 기준으로서 일련의 희석에 사용하였다. 5 μ l의 기준 및 샘플을 각각 깨끗한 건조 미세적정 플레이트에 추가하였다. 전체 수용성 단백질 시약 A 및 B를 제조사의 지시에 따라 추가하였다 (Bio-Rad Dc Protein Assay). 모든 실험을 3배수로 실행하였다. 흡광도 판독값을 마이크로플레이트 판독기 (Bio-tek PowerWave XS)를 사용하여 750 nm에서 기록하였다.

[0194] 결과

[0195] 식물 잎 세포 구획에서 VP6의 발현

[0196] VP6을 발현시켰고 모든 세포 구획에 대하여 표적화하였으며 (도 6; (~42 kDa에서 선은 VP6을 표시한다)), 침묵 억제 인자가 있거나 없다. 세포질에서, 단백질을 타임 트라이얼의 제1 일부터 발현시켰고, 1주일에 걸친 시험 중에 세포질에서 단백질 축적이 증가한다 (도 6a). ER, 단백질 축적은 제3 일에만 보였다 (도 6b). 단백질은 다른 단백질보다 더 높은 밴드 크기 (대략 11 kDa 이상)로 이동하였다. 이것은 단백질의 C-말단 끝, 뿐만 아니라 분할 부위 (pProEx 벡터 서열로 불림)에 추가된 6개의 히스티딘-태그의 결과일 수도 있다.

[0197] 염록체에서 단백질 축적은 제1 일 및 제3 일에 발생하였다 (도 6 "염록체"). 침묵 억제인자는 단백질에 대한 효과가 있었는데 부재시 단백질이 검출되지 않기 때문이다. 제5 일 및 제7 일에 단백질이 발현되지 않았다. ER에서와 같이, 아포플라스트는 타임 트라이얼의 제3 일 내지 제5 일에 단백질이 최고로 축적되었고 (도 6 "아포플라스트") 제1 일 및 제7 일에는 전혀 축적되지 않았다. 침묵 억제인자는 특히 제3 일에 긍정적인 효과가 있었으며, 그것이 없을 때와 비교하여 더 높은 단백질 검출 수준을 발생시킨다. 두 개의 밴드는 아마 VP6 단백질에서 신호 전달 태그의 분할의 결과로서 ~40 kDa 마크(mark)에서 가시적인 것으로도 볼 수 있다.

[0198] ER, 염록체 및 아포플라스트는 모두 제3 일에 가장 높은 단백질 발현을 나타냈으며, 침묵 억제인자의 존재시 가장 많은 단백질이 축적되었다. 세포질은 단백질 축적에 관하여 최고인데 그것이 타임 트라이얼 동안 높고 증가하는 단백질 발현을 나타냈기 때문이다.

[0199] 세포질에서 히스티딘-태그된 로타바이러스 단백질의 발현

[0200] 네 개의 로타바이러스 VP를 추가적인 벡터, pTRAc-HT에 클로닝하였다. 이 벡터는 세포질에 표적화된 단백질에 대한 6-히스티딘 태그를 포함하고 원하는 단백질에 대한 항체가 이용 불가능한 경우 항-히스티딘 태그 항체의 사용에 의해 검출을 쉽게 만든다. 본 발명의 경우에서, VP6만이 상업적으로 이용 가능한 항체를 갖고 이런 이유로 발명자들은 혈청을 기다리는 동안 모든 단백질의 초기 검출을 위해 이 과정을 시도하였다. 세포질은 또한 VP6 발현을 위해 잘 작동하였고 본 발명자에게 다른 단백질에 시도할 동기를 부여했다.

[0201] 제7 일 타임 트라이얼로부터 제3 일 추출물의 웨스턴 블롯 결과는 VP2, VP4 및 VP6의 성공적인 발현을 나타냈다 (도 7a). 식물에서 VP7의 발현을 얻기 위해서, 다양한 기술들을 시도하였다. 하지만, VP7 침투된 식물은 제1 일부터 황화 잎을 나타냈고 타임 트라이얼의 과정 중에 시드는 것이 진행된다 (도 7b). 이 조건들 하에서는, 식물이 여전히 상당히 양호해보일 때 침투 제1 일 후에도, 단백질의 발현이 검출되지 않았다.

[0202] **식물에서 VP2 및 VP4의 발현**

VP2 및 VP4를 니코티아나 벤타미아나 식물에 침투시켰고 ER, 엽록체, 세포질 및 아포플라스트에 표적화하였다. 발명자들은 대장균에서 어떤 양성 클론도 얻을 수 없었기 때문에 아포플라스트 백터에 표적화된 VP2를 발현시킬 수 없었다. 하지만, 단백질을 성공적으로 발현시켰고 모두 다른 3개의 구획에 표적화하였다 (도 8a). 닭 항-VP2 및 항-VP4 혈청 (1:2000)을 추출물의 웨스턴 블롯 분석에 사용하였다. VP2 및 VP4 밴드는 각각 도 8a 및 8b에서 보이는 바와 같이 100 kDa의 마크 미만에서 가시적이었다 (화살표로 지시된 단백질 밴드). 발현은 VP2에 대하여 세포질 및 ER에서, 한편, VP4에 대해서는 세포질 및 아포플라스트에서 최고인 것으로 나타났다. 침목 억제인자는 단백질의 발현에 대하여 유의한 효과를 갖지 않았다. 웨스턴 블롯에서 볼수 있는 바와 같이, 그것은 VP2 ER 구조에서 발현을 약간만 증가시켰고 나머지에서도 너무 많이 증가되지는 않았다. VP4 구조는 침목 억제인자의 존재시 모두 발현되었다.

[0204] **세포질에서 VP2/6 및 VP2/6/4의 동시-발현**

세포질은 로타바이러스 캡시드 단백질 발현에 대하여 최고인 것으로 나타났고 가장 높은 추출 효율을 나타냈다. 그러므로 모든 추가의 발현을 세포질에 표적화된 단백질로 실행하였다.

VP2 및 VP6은 마우스에서 보호 면역원성 반응으로 RLP를 형성하는 것으로 나타났고 그러므로 세포질에서 VP2/6 및 VP2/6/4의 동시-발현을 조사하였다. 동시-발현된 VP2/6/4의 제3 일 추출물을 항-VP2 및 VP4 혈청 (1/5000) 및 마우스 항-VP6 항체 (1:5000)로의 웨스턴 블롯에 의해 검출하였다 (도 9). VP6 발현은 이전에 결정된 바와 같이 매우 높았지만, VP2 및 VP4의 발현은 100 kDa 마크의 매우 희미한 밴드로부터 볼 수 있는 바와 같이 매우 낮았다. 이것은 VP2 및/또는 VP4에 대해서는 더 낮게 놔두면서, VP6의 과발현에서 활용되고 있는 더 많은 숙주 세포 자원에서 발생하는 동시-발현에 의한 것이었을 수도 있었다. 검출된 밴드가 VP2 및 VP4 또는 2개의 단백질 중 하나인지를 결정하는 것 또한 쉽지는 않았다. 130 kDa 이상으로 이동하는 매우 가시적인 밴드는 다이머화된 VP6일 수도 있다. 55 kDa 마크에서 가시적인 밴드는 아마도 풍부한 식물 효소 Rubisco일 것이다.

세포질-발현된 VP6, 뿐만 아니라 VP2/6 및 VP2/6/4에 대한 투과 전자 현미경 분석을 수행하여 단백질 입자 및 조립된 RLP에 대하여 검사하였다 (도 10). 이것은 또한 VP2 및/또는 VP4가 확실히 성공적으로 동시-발현되는지를 결정하였다. 도 10b에서 화살표로 나타난 바와 같이 VP6은 단독으로 발현될 때 조립되어 단백질의 씨스를 형성하였다 (도 10c). VP2는 다른 단백질이 조립되어 결국 완전한 로타바이러스 구조를 형성할 수 있게 하는 스캐폴딩(scaffolding) 단백질의 역할을 한다. 이와 같이 VP6은 VP2에 결합했지만 동시-발현된 VP2/6/4에서 VP4 구조를 결정하는 것은 여전히 쉽지 않았다. 도 10d의 전자 현미경 사진은 순수하게 조립된 VP2/6 입자일 수도 있다. 하지만, VP4가 단백질 조립 중에 VP6 단백질에 결합하는 것으로 나타났고, 이것은 VP7 결합 전에 발생한다. 이 VP4 구조들은 안정하지 않고 RLP 구조는 전자 현미경 검사를 위한 준비 과정 중에 줄어들 가능성이 크다.

[0208] **VP2/6 및 VP2/6/4의 수크로스 구배 정제**

VP2/6 및 VP2/6/4를 10 내지 60% 수크로스 범위의 수크로스 구배에서 정제하였다 (도 11a). 2 ml 분획을 각 튜브의 바닥에서 수거하였고 분획이 단백질을 함유하는지 결정하기 위해서 마우스 항-VP6 항체 및/또는 닭 항-VP2 및 VP4 혈청로 검출하였다. VP2/6에 대하여, 단백질을 분획 16 및 17에서 발견하였는데 이것들이 블롯에서 VP6 단백질에 대하여 양성이기 때문이다 (도 11b). 닭 항-VP2 및 VP4로 VP2/6/4 블롯 분석은 모든 분획에 걸쳐 양성의 결과를 나타냈다. 이것은 닭 혈청에 의한 높은 수준의 배경 단백질 검출의 결과일 수도 있다. 하지만, 도트의 강도는 도 11c에서 나타난 바와 같이 분획 17 및 18에서 가장 높았는데, 아마도 이 분획들에서 원하는 단백질의 더 높은 농도로 인한 것이다.

[0210] 이 결과들을 종합해보면 (도 11b 및 11c), 로타바이러스 단백질은 16 내지 20의 범위에 있는 분획에 있었다.

[0211] **분획의 웨스턴 블롯 및 쿠마시 염색**

[0212] 동시-발현된 V2/6의 웨스턴 블롯 및 SDS-PAGE를 실행하여 분획 13 내지 20에서 VP2 및 VP6 단백질의 존재를 확인하였다. 마우스 항-VP6 항체로 검출된 VP6 단백질에 대한 웨스턴 블롯 분석은 분획 16 내지 20에서 양성이었다 (도 12a, 아래 화살표 도 12c). 닭 항-VP2 혈청으로 탐침된 VP2 단백질을 분획 17 내지 20에서 검출하였다 (도 12a, 상단 화살표). VP2는 과거 동시-발현 연구에서 VP6보다 더 낮게 발현되는 것으로 나타났고 이것은 VP2 단백질 밴드가 VP6과 비교하여 더 낮은 강도를 갖는 것으로 도 12a에서도 나타났다.

[0213] 도트 블롯에 의해 VP6 단백질을 함유하는 것으로 이전에 결정된 동시-발현된 분획 16 및 17 (도 11b)을 SDS-PAGE 겔에서 전기영동하였다. 알려진 농도의 단백질, SF9 콘충 세포 발현된 VP6 ($0.91 \mu\text{g}/\mu\text{l}$)는 VP2/6 조단백

질(crude protein)의 농도를 결정하기 위해 포함되었다 (도 11b 및 c). 이것을 Syngene Gel Documentation System을 사용하여 조단백질 밴드 (crude로 표기된 레인)의 밀도 스캐닝으로 실행하였고 그 결과 발명자들이 잎 물질의 키로그램 당 VP2/6의 양을 결정할 수 있게 한다. 단백질 수율은 대략 1.54 g/kg 생중량 (FW)인 것으로 발견되었다. 정제된 RLP 1.1 mg을 식물 재료의 1 그램에서 얻었다 (1.1g/kg).

[0214] VP2/6의 전체 수용성 단백질 검정

[0215] 전체 수용성 단백질 (TSP)을 VP2/6 단백질의 상대적인 양을 결정하기 위해 동시-발현된 VP2/6 분획에서 결정하였다 (도 13). 단백질 농도를 IgG 기준의 사용으로 분획 17 및 18에 대하여 각각 0.538 mg/ml 및 1.012 mg/ml로 계산하였다 (도 13a). 이 분획들에서 VP2/6에 해당하는 단백질 밴드를 Syngene Gel Documentation System을 사용하여 밀도 스캐닝으로 계산하였고 각각 대략 0.108 mg/ml 및 0.202 mg/ml인 것으로 발견되었다.

[0216] 따라서, 분획 17 및 18의 VP2/6에 대한 TSP는 둘 다 대략 20% TSP였다. 수크로스 컬럼에서 대부분의 RLP는 15내지 25% 수크로스인 것으로 발견되었으며, 약 20으로 분획 15와 일치하고, 여기에서 그래프는 갑자기 정점에 도달하고 그 다음에 진정되는 것으로 주목해야 한다. 추출물에서 다양한 재료의 밀도 차이는 발명자들이 원하는 단백질을 분리하고 이로 인해 이것들을 정제하는 것을 허용하였다. 쿠마시 블루로 염색된 SDS-PAGE 젤은 단백질이 비교적 순수하다는 것을 나타내는 하나의 뚜렷한 밴드만을 나타냈다 (도 12b).

[0217] 정제된 VP2/6의 TEM

[0218] 정제된 VP2/6 분획을 모아서 전파 전자 현미경 검사에서 보기 전 수크로스를 제거하기 위해 고염도 PBS에서 투석하였다. TEM을 실행하여 순도를 결정하고 RLP가 정제 과정 후에 온전하게 남아있는지 검사하였다. 도 15에서 보이는 바와 같이, 주로 숙수 세포 생성물로 구성된 대부분의 배경 재료 (도 10b, c 및 d)를 제거하였으며, RLP를 남겨두었다. 대부분의 RLP는 온전하게 남아있지만 일부는 아마도 EM 그리드 상의 조건으로 인한 변형의 결과로서 형태가 없어진 것으로 나타났다.

[0219] 니코티아나 벤타미아나 잎에서 로타바이러스 구조 단백질의 발현의 예비 분석

[0220] 이 예비 분석은 예시적 숙주 발현 시스템으로서 니코티아나 벤타미아나 잎에서 로타바이러스 구조 단백질 VP2 (SEQ ID NO:1), VP4 (SEQ ID NO:2), VP6 (SEQ ID NO:3) 및 VP7 (SEQ ID NO:4)의 발현에 초점을 맞추었다. 본원에서 선택된 로타바이러스의 균주는 남아프리카 공화국 및 다른 사하라 사막 이남 영역에서 대부분 순환하는 G9 P[6] 균주였다. 이 균주를 표적화하는 RLP 백신은 사하라 사막 이남 아프리카에서 질환의 부담을 완화하는데 도움을 줄 것이다.

[0221] 아그로박테리움에 의해 매개된 일파성 발현 시스템을 이 분석에서 사용하였다. 일파성 발현은, 유전자 이식 발현과 대조적으로, 숙주 염색체에서 로타바이러스 캡시드 단백질 유전자의 통합 없이, 비교적 짧은 시간에 단백질의 신속한 발현을 허용한다. 대부분의 단백질은 니코티아나 벤타미아나 잎에서 재조합 아그로박테리움 침투의 제3 일까지 검출 가능한 양으로 발현되고 축적되었다. 하기 나타난 바와 같이 여러 로타바이러스 구조 단백질의 성공적인 발현을 관찰하였으며 표 2에서 상세히 설명된 바와 같이 식물 잎 세포 구획에서 VP2, VP4 및 VP6을 포함한다:

표 2

표 2: 다양한 잎 세포 구획에서 로타바이러스 VP 단백질의 발현

캡시드 단백질	잎 세포 구획				0 = 발현 안됨 1 = 발현됨
	아포플라스트	엽록체	세포질	ER	
VP2	0	1	1	1	
VP4	1	0	1	1	
VP6	1	1	1	1	
VP7	0	0	0	0	

[0222]

[0223] 당단백질 VP7의 발현은 아마도 식물 세포에 대한 그것의 독성 효과로 인해 관찰되지 않았다. 이 예비 연구를 위해, 고유 신호 펩티드를 함유하는 VP7이 사용되었다는 것은 주목할 만하다. 동시-발현 시험 중 제3 일에 침투를

또한 시도하였다. 이것을 단백질이 발현되고 VP2 및 VP6과 조립된 직후 RLP를 형성하는지 알아보기 위해 시도하였다. 본 연구에서 관찰된 바와 같이 재조합 VP7의 독성 성질은 이전에 설명되었다 (Williams et al., 1995; McCorquodale, 1987; Arias et al., 1986).

[0224] 유전자 이식 감자에서 과거 VP7 발현 연구가 보고되었다 (Li et al., 2006; Choi et al., 2005; Wu et al., 2003). Choi et. al. (2005)은 원숭이 로타바이러스 VP7을 사용하였고, Li et. al. 및 Wu et. al. (Li et al., 2006; Wu et al., 2003)은 인간 군 A G1 VP7을 사용하였다. 본원에서 설명된 결과는 인간 로타바이러스 G9 VP7을 사용하였다.

[0225] VP2을 발현시켰고 적절한 cDNA를 클로닝할 수 없었기 때문에 아포플라스트를 제외한 모든 구획에 표적화하였고, 시간 제한은 발명자들에게 다른 구조로 진행하기 전에 몇 번의 시도만을 허용하였다. VP2의 발현 수준은 모든 구획에서 상당히 낮은 것으로 주목되었다. Saldana et al. 2006에 의해 인용된 과거 연구에서, 식물에서 발현에 최적화된 서열을 갖는 VP2는, mRNA가 식물 세포에서 검출됨에도 불구하고, 발현될 수 없다는 것으로 결론 내려졌다. 하지만 그들은 합성 DNA를 사용하여 토마토 식물 세포에서 그것들 발현하도록 관리하였다. VP2 발현의 어려움에 대한 이유는 아마도 부적절한 mRNA 번역의 결과 또는 mRNA가 식물 세포를 불안정하게 하는 일부 서열 모티프(motif)를 함유한다는 것이다 (Kawaguchi and Bailey-Serres, 2002). VP6과 비교하여 VP2의 낮은 발현 수준의 증거는 Mena et al. (2006), Saldana et al. (2006), Vieira et al. (2005), 및 Labbe et al. (1991)에 의한 식물 및 곤충-세포 발현 연구 둘 다에서 나타났다.

[0226] 비리온 구조의 표면에서 스파이크를 형성하는 외부 캡시드 단백질, VP4를 발현시켰고 세포질, ER 및 아포플라스트에서 축적을 위해 표적화하였다. 단백질 축적은 엽록체에서는 관찰되지 않았다. VP2에 대하여 관찰된 바와 같이, VP4에 대한 단백질 발현 수준은 웨스턴 블로트에서 VP6에서 보이는 것보다 더 낮았다. 단백질은 두 개의 단백질, VP5 및 VP8을 발생시키는 트립신 분할 부위를 갖는다.

[0227] 니코티아나 벤타미아나 잎에서 국소적 트립신은 단백질의 일부를 분할하는데 그것들이 지정된 구획에서 더 낮은 농도 수준의 축적된, 온전한 VP4를 발생시키면서 생산되기 때문이다. 단백질은 주요 중화 항원인 것으로 나타났지만 백신 개발을 위해 전체 단백질을 클로닝하려는 시도는 적었다 (Khodabandehloo et al., 2009; Mahajan et al., 1995; Nishikawa et al., 1989). 하지만 VP4의 VP5 또는 VP8 서브유닛의 발현을 나타내는 곤충-세포 및 효모 발현 시스템에서 여러 연구가 있었다 (Andres et al., 2006; Favacho et al., 2006; Kovacs-Nolan et al., 2001). 지금까지, 본 연구는 식물 발현 시스템에서 전체 단백질의 발현을 나타내는 첫 번째 연구를 나타낸다.

[0228] VP6을 모든 구획에서 발현시켰으며, 이 구획에서 제1 일 내지 제7 일에 단백질 축적이 관찰되는 세포질에서 과발현이 관찰되었다. 이것은 프로테아제 활성 및 유전자 침묵이 세포질에서 외부 단백질의 축적을 감소시키거나 이를 저해한다고 제안하는 일부 문헌과 대조적이다 (Fischer et al., 2004). 게다가, 올바른 pH 조건이 제공되면, VP6는 튜브형 또는 나선형 입자로 자가-조립하는 것으로 알려져 있으며, 본 연구에서 나타난 입자와 매우 유사하다 (도 9b) (Estes et al., 1987). VP6은 약 50%가 바이러스 코어로 구성되고 그러므로 로타바이러스 백신의 개발에 있어서 주요한 항원이다. 상기 획득한 결과는 발명자들이 세포질에서 VP2, VP6 및 VP4의 동시-발현을 더 조사하는 것을 가능하게 했다.

[0229] 세포질에서 동시-발현될 때, VP2 및 VP6은 조립되어 RLP를 형성한다. 일과성 발현 시스템에서 1.27-1.54 g/kg FW의 매우 높은 단백질 수득율을 관찰하였다. 수크로스 컬럼에서 정제될 때, 보유된 VP의 양은 1.1 g/kg FW였다. 이 수득율은 니코티아나 벤타미아나의 일과성 발현 시스템을 사용하여 항체, IgG의 생산에 대해 얻은 것과 비슷하며, 최대 1.5 g/kg FW의 수득율을 갖는다 (Vezina et al., 2009). Saldana et al. (2006)은 지금까지 유전자 이식 토마토 식물에서 성공적으로 동시-발현된 로타바이러스 VP2 및 VP6을 갖는 것으로 및 대략 1% 전체 수용성 단백질의 수준으로 알려져 있는 유일한 군이었다. 곤충 세포 발현 시스템에서 VP2/6의 조립은 잘 기록되어 있다 (Vieira et al., 2005;; O'Brien et al., 2000). 이 VP2/6 RLP들은 또한 로타바이러스 감염에 대한 보호 면역을 제공하는 것으로 나타났다 (Zhou et al., 2011, Saldana et al., 2006). 따라서 발명자들이 식물 발현 시스템에서 생산한 VP2/6 RLP는 서브유닛 로타바이러스 백신의 개발에 적합한 후보물질이다.

[0230] VP2/6/4를 또한 동시-발현시켰고 검출하였다. 동시-발현된 단백질의 전체 단백질 흡광도 판독에서 볼 수 있는 제1 피크 (도 14, 분획 16)는 조립된 VP2/6/4일 수도 있지만, TEM 하에서 이 분획의 검사에서, RLP가 검출되지 않았다. 관찰된 단백질 피크는 VP4 모노머 또는 그것의 각각의 VP5 및 VP8 서브유닛의 축적으로부터 발생할 수도 있다. 제2 피크 (분획 18)는, TEM 하에서 검사될 때, VP2/6 샘플에서 나타는 것들과 매우 유사한 RLP 구조를 나타냈다. 하지만, Crawford et al은 VP4가 TEM 하에서는 보이지 않고, VP2/6/4 및 VP2/6/4/7 입자가 TEM 하에

서 유사한 구조 및 지름을 갖는다는 것을 이전에 보고했다 (Crawford 1994). 발명자들은 VP2/6/7 RLP, VP2/6/4/7 RLP 및 VP2/6 RLP에 대하여 같은 관찰을 하였으며, 이것은 모두 규칙적인 TEM 하에서 유사하게 보인다.

[0231] 실시예 2

[0232] 구조

[0233] A-2X35S/ CPMV-HT/ RVA(WA) VP2(opt)/ NOS (구조 번호 1710)

[0234] 로타바이러스 A WA 균주로부터 VP2를 암호화하는 최적화된 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 2X35S-CPMV-HT-NOS 발현 시스템으로 클로닝하였다. VP2 암호화 서열을 함유하는 단편을 프라이머 IF-WA_VP2(opt).s1+3c (도 17A, SEQ ID NO:21) 및 IF-WA_VP2(opt).s1-4r (도 17B, SEQ ID NO:22)을 사용하여 증폭하였으며, 주형으로 최적화된 VP2 유전자 서열 (도 19, SEQ ID NO:45)을 사용하였다. 서열 최적화를 위해서, VP2 단백질 서열 (Genbank 수납 번호 CAA33074)을 역변역하여 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 시스템에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 클로닝하였다. 구조 번호 1191 (도 17C)을 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 1191은 CPMV-HT-기반 발현 카세트에서 원하는 유전자의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌(Plastocyanin) 유전자 프로모터 및 종결자 하에서 TBSV P19 침묵의 억제인자의 동시-발현을 위한 유전자 구조를 통합한다. 백본(backbone)은 pCAMBIA 바이너리 플라스미드이고 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 서열이 도 18 (SEQ ID NO:23)에서 제공된다. 결과에 따른 구조는 특정 번호 1710 (도 23, SEQ ID NO:27)이었다. 로타바이러스 A 균주 WA의 VP2 아미노산 서열은 도 20에서 제공된다 (SEQ ID NO:25). 플라스미드 1710의 대표도는 도 21에서 제공된다.

[0235] B-2X35S/CPMV-HT/RVA(WA) VP2(opt)/NOS 내 BeYDV(m)+레플리카제 증폭 시스템 (구조 번호 1711)

[0236] 로타바이러스 A WA 균주의 VP2를 암호화하는 최적화된 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 BeYDV(m)+레플리카제 증폭 시스템을 포함하는 2X35S/CPMV-HT/NOS로 클로닝하였다. VP2 암호화 서열을 함유하는 단편을 프라이머 IF-WA_VP2(opt).s1+3c (도 17 A, SEQ ID NO:21) 및 IF-WA_VP2(opt).s1-4r (도 17B, SEQ ID NO:22)을 사용하여 증폭하였으며, 주형으로 최적화된 VP2 유전자 서열 (SEQ ID NO:45)을 사용하였다. 서열 최적화를 위해서, VP2 단백질 서열 (Genbank 수납 번호 CAA33074)을 역변역하였고 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 카세트에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 BeYDV(m) 증폭 시스템으로 클로닝하였다. 구조 193 (도 22A)를 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 193은 CPMV-HT-기반 발현 카세트 내 BeYDV(m) 증폭 시스템에서 원하는 유전자의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결자 하에서 TBSV p19 침묵의 억제인자의 동시-발현을 위한 유전자 구조를 통합한다. 백본은 pCAMBIA 바이너리 플라스미드이고 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 서열은 도 22B에서 제공된다 (SEQ ID NO:26). 결과에 따른 구조는 특정 번호 1711 (도 23, SEQ ID NO:27)이었다. 로타바이러스 A 균주 WA의 VP2 아미노산 서열은 도 20에서 제공된다 (SEQ ID NO:25). 플라스미드 1711의 대표도는 도 24에서 제공된다.

[0237] C-2X35S/CPMV-HT/RVA(WA) VP6(opt)/NOS (구조 번호 1713)

[0238] 로타바이러스 A WA 균주의 VP6을 암호화하는 최적화된 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 2X35S-CPMV-HT-NOS 발현 시스템으로 클로닝하였다. VP6 암호화 서열을 함유하는 단편을 프라이머 IF-WA_VP6(opt).s1+3c (도 25a, SEQ ID NO:28) 및 IF-WA_VP6(opt).s1-4r (도 25b, SEQ ID NO:29)을 사용하여 증폭하였으며, 주형으로 최적화된 VP6 유전자 서열을 사용한다. 서열 최적화를 위해, VP6 단백질 서열 (Genbank 수납 번호 AAA47311)을 역변역하였고 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 시스템에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 클로닝하였다. 구조 번호 1191 (도 17c)를 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 1191은 CPMV-HT-기반 발현 카세트에서 원하는 유전자의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결자 하에서 TBSV p19 침묵의 억제인자의 동시-발현을 위한 유전자 구조를 통합한다. 백본은 pCAMBIA 바이너리 플라스미드이고 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보

더로의 서열이 도 18에서 제공된다 (SEQ ID NO:23). 결과에 따른 구조는 특정 번호 1713 (도 25c, SEQ ID NO:30)이었다. 로타바이러스 A 균주 WA의 VP6 아미노산 서열은 도 26에서 제공된다 (SEQ ID NO:31). 플라스미드 1713의 대표도는 도 27에서 제공된다.

[0239] D-2X35S/CPMV-HT/RVA(WA) VP6(opt)/NOS 내 BeYDV(m)+레플리카제 증폭 시스템 (구조 번호 1714)

로타바이러스 A WA 균주의 VP6을 암호화하는 최적화된 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 BeYDV(m)+레플리카제 증폭 시스템을 포함하는 2X35S/CPMV-HT/NOS로 클로닝하였다. VP6 암호화 서열을 함유하는 단편을 프라이머 IF-WA_VP6(opt).s1+3c (도 25a, SEQ ID NO:28) 및 IF-WA_VP6(opt).s1-4r (도 25b, SEQ ID NO:29)를 사용하여 증폭하였으며, 주형으로 최적화된 VP6 유전자 서열 (SEQ ID NO:46)을 사용하였다. 서열 최적화를 위해서, VP6 단백질 서열 (Genbank 수납 번호 aAA47311)를 역번역하였고 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 카세트에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 BeYDV(m) 증폭 시스템으로 클로닝하였다. 구조 193 (도 22A)을 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 193은 CPMV-HT-기반 발현 카세트에서 원하는 유전자의 BeYDV(m) 증폭 시스템으로의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결자 하에 TBSV p19 침묵의 억제인자의 동시-발현을 위한 유전자 구조를 통합한다. 백본은 pCAMBIA 바이너리 플라스미드이고 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 서열은 도 22B에서 제공된다 (SEQ ID NO:26). 결과에 따른 구조는 특정 번호 1714 (도 28, SEQ ID NO:32)이다. 로타바이러스 A 균주 WA의 VP6 아미노산 서열은 도 26에서 제공된다 (SEQ ID NO:31). 플라스미드 1714의 대표도는 도 29에서 제공된다.

[0241] C-2X35S/CPMV-HT/RVA(Rtx) VP4(opt)/NOS (구조 번호 1730)

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP4를 암호화하는 최적화된 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 2X35S/CPMV-HT/NOS로 클로닝하였다. VP4 암호화 서열을 함유하는 단편을 프라이머 IF-Rtx_VP4(opt).s1+3c (도 30A, SEQ ID NO:33) 및 IF-Rtx_VP4(opt).s1-4r (도 30B, SEQ ID NO:34)을 사용하여 증폭하였으며, 주형으로 최적화된 VP4 유전자 서열 (도 31B, SEQ ID NO:47)을 사용하였다. 서열 최적화를 위해, VP4 단백질 서열 (Genbank 수납 번호 aEX30660)을 역번역하였고 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 카세트에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 클로닝하였다. 구조 번호 1191 (도 18, SEQ ID NO:23)을 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 1191은 CPMV-HT-기반 발현 카세트에서 원하는 유전자의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결자 하에서 TBSV p19 침묵의 억제인자의 동시-발현을 위한 유전자 구조를 통합한다. 백본은 pCAMBIA 바이너리 플라스미드이고 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 서열이 제공된다 (도 18, SEQ ID NO:23). 결과에 따른 구조는 특정 번호 1730 (도 31C, SEQ ID NO:50)이다. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8]의 VP4 아미노산 서열은 도 32에서 제공된다 (SEQ ID NO:36). 플라스미드 1730의 대표도는 도 33A에서 제공된다.

[0243] E-2X35S/CPMV-HT/RVA(Rtx) VP4(opt)/NOS 내 BeYDV(m)+레플리카제 증폭 시스템 (구조 번호 1731)

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP4를 암호화하는 최적화된 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 BeYDV(m)+레플리카제 증폭 시스템을 포함하는 2X35S/CPMV-HT/NOS로 클로닝하였다. VP4 암호화 서열을 함유하는 단편을 프라이머 IF-Rtx_VP4(opt).s1+3c (도 30A, SEQ ID NO:33) 및 IF-Rtx_VP4(opt).s1-4r (도 30B, SEQ ID NO:34)을 사용하여 증폭하였으며, 주형으로 최적화된 VP4 유전자 서열 (SEQ ID NO:47)을 사용하였다. 서열 최적화를 위해, VP4 단백질 서열 (Genbank 수납 번호 aEX30660)을 역번역하였고 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 카세트에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 BeYDV(m) 증폭 시스템으로 클로닝하였다. 구조 193 (도 22A)를 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 193은 CPMV-HT-기반 발현 카세트에서 원하는 유전자의 BeYDV(m) 증폭 시스템으로의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결자 하에서 TBSV p19 침묵의 억제인자의 동시-발현을 위한 유전자 구조를 통합한다. 백본은 pCAMBIA 바이너리 플라스미드이고 왼

쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 서열은 도 22B에서 제공된다 (SEQ ID NO:26). 결과에 따른 구조는 특정 번호 1731 (도 31, SEQ ID NO:35)이다. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8]의 VP4 아미노산 서열은 도 32에서 제공된다 (SEQ ID NO:36). 플라스미드 1731의 대표도는 도 33B에서 제공된다.

[0245] F-2X35S/CPMV-HT/RVA(Rtx) VP7(opt)/NOS (구조 번호 1733)

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 고유 신호 웨티드를 가진 VP7을 암호화하는 최적화된 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 2X35S-CPMV-HT-NOS 발현 시스템으로 클로닝하였다. VP7 암호화 서열을 함유하는 단편을 프라이머 IF-Rtx_VP7(opt).s1+3c (도 34A, SEQ ID NO:37) 및 IF-Rtx_VP7(opt).s1-4r (도 34B, SEQ ID NO:38)을 사용하여 증폭하였으며, 주형으로 최적화된 VP7 유전자 서열 (SEQ ID NO:54)을 사용하였다. 서열 최적화를 위해, VP7 단백질 서열 (Genbank 수납 번호 aEX30682)을 역번역하였고 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 시스템에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 클로닝하였다. 구조 번호 1191 (도 17C)을 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 1191은 CPMV-HT-기반 발현 카세트에서 원하는 유전자의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결자 하에서 TBSV p19 침묵의 억제인자의 동시-발현을 위한 유전자 구조를 통합한다. 백본은 pCAMBIA 바이너리 플라스미드이고 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 서열은 도 18에서 제공된다 (SEQ ID NO:23). 결과에 따른 구조는 특정 번호 1733 (도 34C, SEQ ID NO:24)이다. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 고유 신호 웨티드를 가진 VP7의 아미노산 서열은 도 35에서 제공된다 (SEQ ID NO:39). 플라스미드 1733의 대표도는 도 36에서 제공된다.

[0247] D-2X35S/CPMV-HT/TrSp-RVA(Rtx) VP7(opt)/NOS (구조 번호 1734)

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 고유 신호 웨티드의 절단된 버전을 가진 VP7을 암호화하는 최적화된 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 2X35S-CPMV-HT-NOS 발현 시스템으로 클로닝하였다. VP7 암호화 서열을 함유하는 단편을 프라이머 IF-TrSP+Rtx_VP7(opt).s1+3c (도 44A, SEQ ID NO:55) 및 IF-Rtx_VP7(opt).s1-4r (도 44B, SEQ ID NO:56)을 사용하여 증폭하였으며, 주형으로 최적화된 VP7 유전자 서열을 사용하였다 (도 44C, SEQ ID NO:57의 nt 88-891에 해당함). 서열 최적화를 위해, VP7 단백질 서열 (Genbank 수납 번호 aEX30682)을 역번역하였고 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 시스템에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 클로닝하였다. 구조 번호 1191 (도 17C)을 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 1191은 CPMV-HT-기반 발현 카세트에서 원하는 유전자의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결자 하에서 TBSV p19 침묵의 억제인자의 동시-발현을 위한 유전자 구조를 통합한다. 백본은 pCAMBIA 바이너리 플라스미드이고 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 서열은 도 18에서 제공된다 (SEQ ID NO:23). 결과에 따른 구조는 특정 번호 1734 (도 44D, SEQ ID NO:58)이다. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 절단된 신호 웨티드를 갖는 VP7의 아미노산 서열은 도 44E에서 제공된다 (SEQ ID NO:59). 플라스미드 1734의 대표도는 도 44F에서 제공된다.

[0249] G-2X35S/CPMV-HT/PDISP/RVA(WA)VP7(opt)/NOS 내 BeYDV(m +래플리카제 증폭 시스템 (구조 번호 1735))

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP7을 암호화하는 서열을 다음 PCR-기반 방법을 사용하여 Plasto_pro/P19/Plasto_ter 발현 카세트를 함유하는 플라스미드에서 2X35S-CPMV-HT-PDISP-NOS 발현 시스템으로 클로닝하였다. his 야생형 신호 웨티드가 없는 VP7 암호화 서열을 함유하는 단편을 프라이머 IF-Rtx_VP7(opt).s2+4c (도 37A, SEQ ID NO:40) 및 IF-Rtx_VP7(opt).s1-4r (도 34B, SEQ ID NO:38)을 사용하여 증폭하였으며, 최적화된 VP7 유전자 서열 (SEQ ID NO:54)을 사용하였다. 서열 최적화를 위해, VP7 단백질 서열 (Genbank 수납 번호 aEX30682)을 역번역하였고 인간 코돈 사용, GC 함량 및 mRNA 구조에 최적화하였다. PCR 생성물을 2X35S/CPMV-HT/NOS 발현 시스템에서 In-Fusion 클로닝 시스템 (Clontech, Mountain View, CA)을 사용하여 알팔파 PDI 신호 웨티드와 인-프레임 클로닝하였다. 구조 1192 (도 38)를 SacII 및 StuI 제한 효소로 분해하였고 선형화된 플라스미드를 In-Fusion 조립 반응에 사용하였다. 구조 번호 1192는 CPMV-HT-기반 발현 카세트에서 알팔파 PDI 신호 웨티드와 인-프레임 원하는 유전자의 "In Fusion" 클로닝을 위해 의도된 수용기 플라스미드이다. 그것은 또한 알팔파 플라스토시아닌 유전자 프로모터 및 종결자 하에서 TBSV p19 침묵의 억제인자의 동

시-발현을 위한 유전자 구조를 통합한다. 백본은 pCAMBIA 바이너리 플라스미드이고 왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 서열은 도 39에서 제공된다 (SEQ ID NO:41). 결과에 따른 구조는 특정 번호 1735 (도 40, SEQ ID NO:42)이다. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 PDISP/VP7 아미노산 서열은 도 41에서 제공된다 (SEQ ID NO:43). 플라스미드 1735의 대표도는 도 42에서 제공된다.

표 3

[0251] RLP의 생산을 위해 합성된 유전자의 설명.

SEQ ID 번호	항원	기원의 균주	서열 타입*	명세서 내 도면
45	VP2	WA	최적화됨	19B
46	VP6	WA	최적화됨	25D
47	VP4	Rotarix	최적화됨	31B
50	VP4	SA11	야생형	43A
51	VP4	SA11	최적화됨	43B
54	VP7	Rotarix	최적화됨	34E
53	VP7	SA11	야생형	43D
52	VP7	SA11	최적화됨	43C

[0252] * 최적화된 서열은 우선적인 인간 코돈의 사용을 선호하고 GC 함량을 증가시키도록 변형되었다.

표 4

[0253] RLP의 생산을 위해 조립된 및 테스트된 구조의 설명.

발현 시스템	증폭 시스템	신호 웨티드	항원 (균주)*	PCR에 사용된 유전자 SEQ ID	구조 번호
CPMV HT	-	-	RVA(WA) VP2 [최적화됨]	SEQ ID NO: 45	1710
CPMV HT	BeYDV(m)+rep	-	RVA(WA) VP2 [최적화됨]	SEQ ID NO: 45	1711
CPMV HT	-	-	RVA(WA) VP6 [최적화됨]	SEQ ID NO: 46	1713
CPMV HT	BeYDV(m)+rep	-	RVA(WA) VP6 [최적화됨]	SEQ ID NO: 46	1714
CPMV HT	-	-	RVA(Rtx) VP4 [최적화됨]	SEQ ID NO: 47	1730
CPMV HT	BeYDV(m)+rep	-	RVA(Rtx) VP4 [최적화됨]	SEQ ID NO: 47	1731
CPMV HT	-	WtSp	RVA(Rtx) VP7 [최적화됨]	SEQ ID NO: 54	1733
CPMV HT	-	TrSp	RVA(Rtx) VP7 [최적화됨]	SEQ ID NO: 54	1734
CPMV HT	-	SpPDI	RVA(Rtx) VP7 [최적화됨]	SEQ ID NO: 54	1735
CPMV HT	BeYDV(m)+rep	WtSp	RVA(Rtx) VP7 [최적화됨]	SEQ ID NO: 54	1736
CPMV HT	BeYDV(m)+rep	TrSp	RVA(Rtx) VP7 [최적화됨]	SEQ ID NO: 54	1737
CPMV HT	BeYDV(m)+rep	SpPDI	RVA(Rtx) VP7 [최적화됨]	SEQ ID NO: 54	1738
CPMV HT	-	-	RVA(SA11) VP4	SEQ ID NO: 50	1760

CPMV HT	BeYDV(m)+rep	-	RVA(SA11) VP4	SEQ ID NO: 50	1761
CPMV HT	-	-	RVA(SA11) VP4 [최적화됨]	SEQ ID NO: 51	1770
CPMV HT	BeYDV(m)+rep	-	RVA(SA11) VP4 [최적화됨]	SEQ ID NO: 51	1771
CPMV HT	-	WtSp	RVA(SA11) VP7	SEQ ID NO: 53	1763
CPMV HT	-	TrSp	RVA(SA11) VP7	SEQ ID NO: 53	1764
CPMV HT	-	SpPDI	RVA(SA11) VP7	SEQ ID NO: 53	1765
CPMV HT	BeYDV(m)+rep	WtSp	RVA(SA11) VP7	SEQ ID NO: 53	1766
CPMV HT	BeYDV(m)+rep	TrSp	RVA(SA11) VP7	SEQ ID NO: 53	1767
CPMV HT	BeYDV(m)+rep	SpPDI	RVA(SA11) VP7	SEQ ID NO: 53	1768
CPMV HT	-	WtSp	RVA(SA11) VP7 [최적화됨]	SEQ ID NO: 52	1773
CPMV HT	-	TrSp	RVA(SA11) VP7 [최적화됨]	SEQ ID NO: 52	1774
CPMV HT	-	SpPDI	RVA(SA11) VP7 [최적화됨]	SEQ ID NO: 52	1775
CPMV HT	BeYDV(m)+rep	WtSp	RVA(SA11) VP7 [최적화됨]	SEQ ID NO: 52	1776
CPMV HT	BeYDV(m)+rep	TrSp	RVA(SA11) VP7 [최적화됨]	SEQ ID NO: 52	1777
CPMV HT	BeYDV(m)+rep	SpPDI	RVA(SA11) VP7 [최적화됨]	SEQ ID NO: 52	1778

[0254] † WtSp: 야생형 신호 펩티드, SpPDI: 식물 기원의 신호 펩티드, 알팔파 단백질 이황화 이소머라제 유전자의 클로닝된 형태, TrSp: 절단된 야생형 신호 펩티드, TrSp는 WtSp의 2번째 Met에서 시작한다 (M30).

[0255] * [최적화됨]은 서열이 코돈 사용, GC 함량 및 RNA 구조에 기초하여 최적화되었다는 것을 의미한다.

실시예 3

유전자 구조의 조립 및 아그로박테리움 형질전환

[0256] 플라스미드 1710, 1713, 1730 및 1734를 포함하는, 모든 플라스미드를 사용하여 아그로박테리움 투메파시엔스 (AGL1; ATCC, Manassas, VA 20108, USA)를 전기천공에 의해 형질전환하였다 (Mattanovich et al., 1989, Nucleic Acid Res. 17:6747). 대안으로, CaC12-제조된 컴피턴트 세포를 사용하는 열 충격 (XU et al., 2008, Plant Methods 4)을 사용할 수도 있다. 생성된 아그로박테리움 투메파시엔스 균주에서 플라스미드의 온전서를 제한 맵핑(restriction mapping)에 의해 확인하였다. 특정 바이너리 플라스미드로 형질전환된 아그로박테리움 투메파시엔스 균주를 AGL1/"플라스미드 번호"로 명명한다. 예를 들어, 구조 번호 1710으로 형질전환된 아그로박테리움 투메파시엔스 균주는 "AGL1/1710"으로 불린다.

[0257] 식물 바이오매스의 제조, 접종, 아그로인필트레이션, 및 수확

[0258] 니코티아나 벤타미아나 식물은 상업적인 퍼트 모스(peat moss) 기질로 채워진 플랫(flat)에서 씨(seed)에서 성장하였다. 식물을 16/8 광주기 및 25°C 낮/20°C 밤의 온도 요법 하에 비닐하우스에서 성장하게 하였다. 파종 후 3주에, 개개의 식물을 골라서, 화분에 옮겨심었고 같은 환경 조건 하에 비닐하우스에서 추가 3주 동안 성장하게 두었다.

[0259] 각 구조로 트랜스펙션된 아그로박테리아를 0.6 내지 2.5의 OD600에 도달할 때까지 식물성 및 10 mM 2-(N-모폴리노)에탄슬론산 (MES) 및 50 µg/ml 카나마이신 pH5.6으로 보충된 LB 배지에서 키웠다. 아그로박테리움 혼탁액을 혼합하여 각 구조에 대하여 적절한 비율에 도달하였고 침투 배지 (10 mM MgCl₂ 및 10 mM MES pH 5.6)로 2.5X OD600으로 된다. 아그로박테리움 투메파시엔스 혼탁액을 4°C에서 하룻밤 동안 저장하였다. 침투의 날에, 배양 배취(batch)를 침투 배지로 2.5배 혼탁액 부피로 회석하였고 사용 전에 따뜻하게 하였다. 니코티아나 벤타미아나의 전체 식물을 밀폐 스테인리스 강 탱크에서 20-40 Torr의 진공 하에 2분 동안 박테리아 혼탁액에 거꾸로 배치하였다. 침투 후, 식물을 수확할 때까지 3-12일 배양 기간 동안 온실에 되돌려 놓았다. 수확된 바이오매스를 입자의 정제에 사용될 때까지 냉동보관하였다 (-80°C).

[0262] 로타바이러스-유사 입자의 추출 및 정제

[0263] 단백질을 3배 양의 추출 버퍼 (TNC: 10 mM Tris pH 7.4, 140 mM NaCl, 10 mM CaCl₂)와 함께 분쇄기에서 기계적 추출에 의해 냉동된 바이오매스로부터 추출하였다. 슬러리(slurry)를 큰 기공 나일론 필터를 통해 여과하여 큰 데브리(debris)를 제거하고 5000 g로 4°C에서 5분 동안 원심분리하였다. 상층액을 수거하여 5000 g로 30분 동안 (4°C) 원심분리하여 추가적인 데브리를 제거하였다. 상층액을 심층 여과 (depth-filter)하였고 한의 여과 (ultra-filter)하였고 여과물을 75 000 g에서 20분 동안 (4°C) 원심분리하여 로타바이러스-유사 입자를 농축하였다. 입자를 함유하는 펠릿을 1/12 부피의 TNC에서 재현탁하였고 불용성 물질을 5000 g에서 5분 동안의 원심분리로 제거하였다. 상층액을 이오디사놀 밀도 구배에 로딩하기 전에 미라클로스에서 여과하였다.

[0264] 밀도 구배 원심분리를 다음과 같이 수행하였다. 5% 내지 45%의 이오디사놀의 단계 구배를 함유하는 튜브를 제조하였고 로타바이러스-유사 입자를 함유하는 여과된 추출물로 오버레이하였다. 구배를 120 000 g에서 4시간 동안 (4°C) 원심분리하였다. 원심분리 후, 1 ml 분획을 바다에서 상단까지 수거하였고 쿠마시-염색된 SDS-PAGE 및 웨스턴 블로트으로 분석하였다. 추가의 분석을 위해 선택된 분획에 대한 이오디사놀을 제거하기 위해, 선택된 분획을 75 000 g에서 20분 동안 (4°C) 원심분리하였고 펠릿화된 입자를 신선한 TNC 버퍼에서 재현탁하였다.

[0265] SDS-PAGE 및 면역블로팅(immunoblotting)

[0266] 단백질 농도를 BCA 단백질 검정 (Pierce Biochemicals, Rockport IL)으로 결정하였다. 단백질을 환원 또는 비-환원 조건 하에서 SDS-PAGE로 분리하였고 쿠마시 블루로 염색하였다. 염색된 겔을 스캔하였고 ImageJ Software (NIH)를 사용하여 농도별 분석을 수행하였다.

[0267] 면역블로팅을 위해, 전기영동된 단백질을 폴리비닐 디플루오라이드 막 (Roche Diagnostics Corporation, Indianapolis, IN)으로 전기 이동시켰다. 면역블로팅 전에, 막을 16-18h 동안 4°C에서 5% 탈지유 Tris-보충된 식염수 (TBS-T)로 차단하였다.

[0268] 면역블로팅을 TBS-Tween 20 0.1% 중 2% 탈지유에서, 2 μg/ml으로, 적합한 항체와의 배양으로 수행하였다 (표 5). 화학 발광 검출에 사용된 2차 항체는 표 5에서 나타난 바와 같으며, TBS-Tween 20 0.1% 중 2% 탈지유에서 지시된 바와 같이 희석하였다. 면역 반응 복합체를 기질로서 루미놀 (Roche Diagnostics Corporation)을 사용하는 화학 발광에 의해 검출하였다. 인간 IgG 항체의 허스래디쉬(horseradish) 퍼옥시다제-효소 컨쥬게이션 (conjugation)을 EZ-Link Plus® Activated Peroxidase 컨쥬게이션 키트 (Pierce, Rockford, IL)를 사용함으로써 수행하였다.

표 5

전기영동 조건, 항체, 및 로타바이러스 항원의 면역블로팅을 위한 희석

로타바이러스 항원	전기영동 조건	1차 항체	희석	2차 항체	희석
VP2	환원	토끼 다클론성 항-VP2 (Koki Taniguchi에 의해 제공됨)	1 μg/ml	염소 항-토끼 (JIR 111-035-144)	1:10 000
VP6	환원	ABIN 308233	1:20000	염소 항-토끼 (JIR 111-035-144)	1:10 000
VP7	비-환원	토끼 다클론성 항-VP7 (Koki Taniguchi에 의해 제공됨)	1:2000	염소 항-토끼 (JIR 111-035-144)	1:10 000

[0270] 항-VP4 효소-결합된 면역 흡착 검정 (ELISA)

[0271] U-바닥 96-웰 미세적정 플레이트를 10 mM PBS pH7.4 (포스페이트-완충된 식염수), 150 mM NaCl에서 1:100000으로 희석된 마우스 단클론성 항-VP4 (Koki Taniguchi 교수에 의해 제공됨)로 4°C에서 16-18시간 동안 코팅하였다. 배양 후, 플레이트를 0.1% Tween-20을 함유하는 10 mM PBS pH7.4, 1 M NaCl로 세 번 세척하였고 0.1% Tween-20을 함유하는 10 mM PBS pH7.4, 150 mM NaCl 중의 5% BSA로 37°C에서 1시간 동안 차단하였다. 차단 단계 후, 플레이트를 0.1% Tween-20을 함유하는 10 mM PBS pH7.4, 1 M NaCl로 세 번 세척하였다. 샘플을 추가하였고 플레이트를 37°C에서 1시간 동안 배양하였다. 그 다음에 플레이트를 0.1% Tween-20을 함유하는 10 mM PBS pH7.4, 1 M NaCl, 1 mM CaCl₂, 0.5 mM MgCl₂로 세 번 세척하였다. 나머지 모든 세척 단계에 대하여, 세척

버퍼는 같은 것을 유지하고 세 번째 세척 중에, 세척 용액을 완전히 제거하기 전에 플레이트를 실온에서 10분 동안 배양하였다. 0.1% Tween-20을 함유하는 10 mM PBS pH7.4, 150 mM NaCl, 1 mM CaCl₂, 0.5 mM MgCl₂ 중의 3% BSA로 1:10000으로 희석된, 로타바이러스에 대하여 발생된 토키 다클론성 항체를 추가하였고 플레이트를 37 °C에서 1시간 동안 배양하였다. 그 다음에 플레이트를 세 번 세척하였고 0.1% Tween-20을 함유하는 10 mM PBS pH7.4, 150 mM NaCl, 1 mM CaCl₂, 0.5 mM MgCl₂ 중의 3% BSA로 1:5000으로 희석된 홀스래디쉬 페옥시다제-컨쥬게이션된(conjugated) 염소 항-토키 항체 (111-035-144, Jackson Immunoresearch, West Grove, PA)를 추가하였고 플레이트를 37°C에서 1시간 동안 배양하였다. 플레이트를 세 번 세척하였다. 최종 세척 후, 플레이트를 SureBlue TMB 페옥시다제 기질 (KPL, Gaithersburg, MD)과 함께 실온에서 20분 동안 배양하였다. 반응을 1N HCl의 추가로 중단하였고 A450 값을 Multiskan Ascent 플레이트 판독기 (Thermo Scientific, Waltham, MA)를 사용하여 측정하였다.

[0272] VP2 및 VP6을 포함하는 로타바이러스-유사 입자의 생산.

[0273] VP2 및 VP6을 포함하는 로타바이러스-유사 입자를 니코티아나 벤타미아나에서 일과성 발현에 의해 생산하였다. 식물에 1:1 비율로 AGL1/1710 및 AGL1/1713의 혼합물을 함유하는 아그로박테리아의 접종으로 아그로-침투하였고 수확 전 7일 동안 배양하였다. 로타바이러스-유사 입자를 재료 및 방법 섹션에서 설명된 방법론을 사용하여 바이오매스로부터 정제하였다. 이오딕사놀 밀도 구배에서 정화된 추출물의 원심분리 후, 튜브의 바닥에서 처음 10개의 분획을 쿠마시-염색된 SDS-PAGE로 분석하였다. 도 45A에서 나타난 바와 같이, 로타바이러스 항원 (VP2 및 VP6)을 밀도 구배의 분획 2 및 3에서 주로 발견하였으며 이오딕사놀의 농도는 대략 35%이며, 로타바이러스-유사 입자가 발견되는 것으로 예상되는 농도이다. 이 분획들에서 식물 단백질에 의한 오염이 거의 발견되지 않았다. 항-로타바이러스 과다면역 토키 혈청 및 다클론성 토키 항-VP2 항체로 분획의 웨스턴 블롯 분석은 밀도 구배 분획에서 VP2 및 VP6의 동일성을 확인하였다 (도 45B 및 45C). 분획 2 및 3을 모아서 이오딕사놀을 고속 원심분리 및 재현탁에 의해 제거하였고, VP2 및 VP6의 로타바이러스 입자와 유사한 입자로의 조립을 확인하기 위해 정제된 입자에 대하여 냉동-전자 현미경 분석 (NanoImaging Services Inc., La Jolla, CA)을 요청하였다. 도 49(왼쪽 패널)에서 나타난 바와 같이, VP2/VP6 입자의 cryoEM 이미지는 항원의 로타바이러스-유사 입자로의 올바른 조립을 확인하였다.

[0274] VP2, VP6 및 VP7을 포함하는 로타바이러스-유사 입자의 생산.

[0275] VP2, VP6 및 VP7을 포함하는 로타바이러스-유사 입자를 니코티아나 벤타미아나에서 일과성 발현에 의해 생산하였다. 식물에 1:1:1 비율로 AGL1/1710, AGL1/1713, AGL1/1734의 혼합물을 함유하는 아그로박테리아의 접종으로 아그로-침투하였고 수확 전 7일 동안 배양하였다. 로타바이러스-유사 입자를 재료 및 방법 섹션에서 설명된 방법론을 사용하여 바이오매스로부터 정제하였다. 이오딕사놀 밀도 구배에서 정화된 추출물의 원심분리 후, 튜브의 바닥에서 처음 10개의 분획을 쿠마시-염색된 SDS-PAGE로 분석하였다. 도 46A에서 나타난 바와 같이, 로타바이러스 항원 (VP2, VP6 및 VP7)을 밀도 구배의 분획 2 및 3에서 주로 발견하였으며 이오딕사놀의 농도는 대략 35%이며, 로타바이러스-유사 입자가 발견되는 것으로 예상되는 농도이다. 이 분획들에서 식물 단백질에 의한 오염이 거의 발견되지 않았다. 분획의 항-로타바이러스 과다면역 토키 혈청 및 다클론성 토키 항-VP7 항체로의 웨스턴 블롯 분석은 밀도 구배 분획에서 VP6 및 VP7의 동일성을 확인하였다 (도 46B 및 46C).

[0276] VP2, VP4, VP6 및 VP7을 포함하는 로타바이러스-유사 입자의 생산.

[0277] VP2, VP4, VP6 및 VP7을 포함하는 로타바이러스-유사 입자를 니코티아나 벤타미아나에서 일과성 발현에 의해 생산하였다. 식물에 1:1:1:1 비율로 AGL1/1710, AGL1/1730, AGL1/1713, AGL1/1734의 혼합물을 함유하는 아그로박테리아의 접종으로 아그로-침투하였고 수확 전 7일 동안 배양하였다. 로타바이러스-유사 입자를 재료 및 방법 섹션에서 설명된 방법론을 사용하여 바이오매스로부터 정제하였다. 이오딕사놀 밀도 구배에서 정화된 추출물의 원심분리 후, 튜브의 바닥에서 처음 10개의 분획을 쿠마시-염색된 SDS-PAGE로 분석하였다. 도 47A에서 나타난 바와 같이, 네 개의 로타바이러스 항원 중 셋 (VP2, VP6 및 VP7)이 가시화되었고 밀도 구배의 분획 3에서 주로 발견하였으며 이오딕사놀의 농도는 대략 35%이며, 로타바이러스-유사 입자가 발견되는 것으로 예상되는 농도이다. 이 분획들에서 식물 단백질에 의한 오염은 거의 발견되지 않았다. 쿠마시-염색된 겔에서 VP4의 검출 가능한 수준의 부재는 같은 분석이 정제된 인간 로타바이러스 비리온에서 수행될 때 VP4가 관찰될 수 없기 때문인 것으로 예상되었다. 분획의 항-로타바이러스 과다면역 토키 혈청 및 다클론성 토키 항-VP7 항체로의 웨스턴 블롯 분석은 밀도 구배 분획에서 VP6 및 VP7의 동일성을 확인하였다 (도 47B 및 47C).

[0278] 이오딕사놀을 고속 원심분리 및 재현탁에 의해 분획 3으로부터 제거하였고, 정제된 입자를 VP4의 존재를 확인하기 위해 ELISA로 분석하였다. 도 48에서 제공된 결과는 ELISA가 VP4를 배경 신호 수준으로만 발생하는 VP2/VP6

및 VP7을 포함하는 음성 대조군 입자로서 특이적으로 인식한다는 것을 분명하게 나타낸다. 반대로, VP2, VP4, VP6 및 VP7 항원을 포함하는 3개의 다른 많은 정제된 입자의 분석은 같은 조건 하에 검정될 때 강하고 균일한 신호를 나타낸다. 정제된 VP2/VP4/VP6/VP7 RLP에 대하여 네 개의 항원의 로타바이러스 입자와 유사한 입자로의 조립을 확인하기 위해 냉동-전자 현미경 분석 (Nanolmaging Services Inc., La Jolla, CA)을 요청하였다. 도 49 (오른쪽 패널)에서 나타난 바와 같이, VP2/VP4/VP6/VP7 입자의 cryoEM 이미지는 항원의 로타바이러스-유사 입자로의 올바른 조립을 확인하였다.

[0279] 표 6은 본 발명의 다양한 구체예에서 제공된 서열을 나열한다.

표 6

서열 식별자에 대한 서열 설명

SEQ ID NO	설명	페이지/도면
1	최적화된 G9P6_VP2의 아미노산 서열	도면 16A
2	최적화됨 G9P6_VP4의 아미노산 서열	도면 16B
3	최적화됨 G9P6_VP6의 아미노산 서열	도면 16C
4	최적화됨 G9P6_VP7의 아미노산	도면 16D
5	프라이머 VP2F	표 1
6	프라이머 VP2R	표 1
7	프라이머 VP4F	표 1
8	프라이머 VP4R	표 1
9	프라이머 VP6F	표 1
10	프라이머 VP6R	표 1
11	프라이머 VP7F	표 1
12	프라이머 VP7R	표 1
13	G9P6_VP2의 뉴클레오티드 서열	도면 16A
14	최적화된 G9P6_VP2의 뉴클레오티드 서열	도면 16A
15	G9P6_VP4의 뉴클레오티드 서열	도면 16B
16	G9P6_VP4의 뉴클레오티드 서열	도면 16B
17	G9P6_VP6의 뉴클레오티드 서열	도면 16C
18	최적화된 G9P6_VP6의 뉴클레오티드 서열	도면 16C
19	G9P6_VP7의 뉴클레오티드 서열	도면 16D
20	G9P6_VP7의 뉴클레오티드 서열	도면 16D
21	프라이머 IF-WA_VP2(opt).s1+3c	도면 17A
22	프라이머 IF-WA_VP2(opt).s1-4r	도면 17B
23	왼쪽t-DNA 보더에서 오른쪽 t-DNA 보더로의 구조 1191 (밀줄 그어져 있음). 플라스토시아닌-P19-플라스토시아닌 침묵 억제자 발현 카세트가 있는 2X35S/CPMV-HT/NOS	도면 18
24	2X35S 프로모터에서 NOS 종결자까지의 발현 카세트 번호 1733. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP7은 밀줄 그어져 있다.	도면 34C
25	로타바이러스 A WA 균주의 VP2의 아미노산 서열	도면 20
26	왼쪽t-DNA 보더에서 오른쪽 t-DNA 보더로의 구조 193 (밀줄 그어져 있음). 플라스토시아닌-P19-플라스토시아닌 침묵 억제자 발현 카세트가 있는 2X35S/CPMV-HT/NOS 내 BeYDV(m)+래플리카제 증폭 시스템	도면 22B
27	2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1710. 로타바이러스 A WA 균주의 VP2(opt)는 밀줄 그어져 있다.	도면 23
28	프라이머 IF-WA_VP6(opt).s1+3c	도면 25a
29	프라이머 IF-WA_VP6(opt).s1-4r	도면 25b
30	2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1713. 로타바이러스 A WA 균주의 VP6(opt)는 밀줄 그어져 있다.	도면 25c
31	로타바이러스 A WA 균주의 VP6의 아미노산 서열	도면 26
32	2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1714. 로타바이러스 A WA 균주의 VP6(opt)는 밀줄 그어져 있다.	도면 28
33	프라이머 IF-Rtx_VP4(opt).s1+3c	도면 30A
34	프라이머 IF-Rtx_VP4(opt).s1-4r	도면 30B
35	2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1731. 로타바이러스 A Rotarix 균주의 VP4(opt)는 밀줄 그어져 있다.	도면 31A
36	로타바이러스 A Rotarix 균주의 VP4의 아미노산 서열	도면 32

37	프라이머 IF-Rtx_VP7(opt).s1+3c	도면 34A
38	프라이머 IF-Rtx_VP7(opt).s1-4r	도면 34B
39	로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP7의 아미노산 서열	Fig 35
40	프라이머 IF-Rtx_VP7(opt).s2+4c	도면 37A
41	원쪽t-DNA 보더에서 오른쪽 t-DNA 보더로의 구조 1192 (밀줄 그어져 있음). 플라스토시아닌-P19-플라스토시아닌 침묵 억제자 발현 카세트가 있는 2X35S/CPMV-HT/PDISP/NOS	도면 39
42	2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1735. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 PDISP/VP7(opt)는 밀줄 그어져 있다.	도면 40A
43	로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 PDISP/VP7의 아미노산 서열	41
44	2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1730. 로타바이러스 A Rotarix 균주의 VP4(opt)는 밀줄 그어져 있다.	도면 31C
45	로타바이러스 A WA 균주의 VP2(opt)를 암호화하는 뉴클레오티드 서열	도면 19
46	로타바이러스 A WA 균주의 VP6(opt)를 암호화하는 뉴클레오티드 서열	도면 25d
47	균주 RVA/백신/USA/Rotarix-A41CB052A/1988/G1P1A[8]의 로타바이러스 A VP4의 최적화된 암호화 서열	도면 31B
48	로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP7을 암호화하는 뉴클레오티드 서열	도면 34D
49	로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 PDISP/VP7(opt)를 암호화하는 뉴클레오티드 서열	도면 40B
50	균주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의 로타바이러스 A VP4의 암호화 서열	도면 43A
51	균주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의 로타바이러스 A VP4의 최적화된 암호화 서열	도면 43B
52	균주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의 로타바이러스 A VP7의 최적화된 암호화 서열	도면 43C
53	균주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의 로타바이러스 A VP7의 최적화된 암호화 서열	도면 43D
54	균주 RVA/백신/USA/Rotarix-A41CB052A/1988/G1P1A[8]의 로타바이러스 A VP7의 최적화된 아미노산 서열	도면 34E
55	프라이머 IF-TrSP+Rtx_VP7(opt).s1+3c	도면 44A
56	프라이머 IF-Rtx_VP7(opt).s1-4r	도면 44B
57	균주 RVA/백신/USA/Rotarix-A41CB052A/1988/G1P1A[8]의 로타바이러스 A VP7의 최적화된 암호화 서열의 뉴클레오티드 서열	도면 44C
58	2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1734. 로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 VP7은 밀줄 그어져 있다.	도면 44D
59	로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의 TrSp-VP7의 아미노산 서열	도면 44E

참고문헌

- [0281] Yang Y M, Li X, Yang H, et al. Immunogenicity and virus-like particle formation of rotavirus capsid produced in transgenic plants. *Sci China Life Sci*, 2011, 54: 82-89.
- [0282] Angel, J., Franco, M.A. and Greenberg, H.B. (2007). Rotavirus vaccines: recent developments and future considerations. *Nature reviews: Microbiology* 5, 529-539.
- [0283] Araujo, I.T., Ferreira, M.S.R. and Failho, A.M. (2001). Rotavirus genotypes [4]G9, P[6]G9, and P[8]G9 in hospitalized children with acute gastroenteritis in Rio de Janeiro, Brazil. *J Clin Microbiol* 39 1999-2001.
- [0284] Arias, C.F., Ballado, T. and Plebafiski, M. (1986). Synthesis of the outer-capsid glycoprotein of the simian rotavirus SA11 in *Escherichia coli*. *Gene* 47, 211-219.
- [0285] Au K.S., Mattion N.M., Estes M.K., (1993). A Subviral Particle Binding Domain on the Rotavirus Nonstructural Glycoprotein NS28. *Virology* 194, 665-67
- [0286] Balen B, Krsnik-Rasol M, (2007). N-glycosylation of recombinant therapeutic glycoproteins in plant

systems. Food Technology and Biotechnology 45 1-10.

- [0288] Bardor, M., Faveeuw, C, Fitchette, A.C., Gilbert,D., Galas, L., Trottein, F., Faye, L. and Lerouge P. (2003). Immunoreactivity in mammals of two typical plant glycoepitopes, core alpha (1,3)-fucose and core xylose. Glycobiology 13 427-434.
- [0289] Berois, M., Sapin, C, Erk, I., Poncet, D. and Cohen, J. (2003). Rotavirus Nonstructural Protein NSP5 Interacts with Major Core Protein VP2. Journal of virology 77, 1757.
- [0290] Bertolotti-Ciarlet, A., Ciarlet, M., Crawford, S.E., Conner, M.E. and Estes, M.K. (2003). Immunogenicity and protective efficacy of rotavirus 2/6-virus-like particles produced by a dual baculovirus expression vector and administered intramuscularly, intranasally, or orally to mice. Vaccine 21, 3885-3900.
- [0291] Chen, J.Z., Settembre, E.C., Aoki, A.T., Zhang, X., Bellamy, A.R., Dormitzer, P.R., Harrison, S.C. and Grigorieff, N. (2009). Molecular interactions in rotavirus assembly and uncoating seen by high-resolution cryo-EM. PNAS 106, 10644-10648.
- [0292] Crawford, S.E., Estes, M.K., Ciarlet, M., Barone, C, O'Neal, CM., Cohen, J. and Conner, M.E. (1999). Heterotypic protection and induction of a broad heterotypic neutralization response by rotavirus-like particles. Journal of Virology 73, 4813-4822.
- [0293] Crawford, S.E., Labbe, M., Cohen, J., Burroughs, M.H., Zhou, Y.J. and Estes, M.K. (1994). Characterization of virus-like paricles produced by the expression of rotavirus capsid proteins in insect cells. Journal of virology 68, 5945-5952.
- [0294] Denisova, E.R., Dowling, W., LaMonica, R., Shaw, R., Scarlata, S., Ruggeri, F. and Mackow, E.R. (1999). Rotavirus capsid protein VP5* Permeabilizes Membranes. Journal of Virology 73 3147-3153
- [0295] Dennehy, P.H. (2007). Rotavirus vaccines-An update. Vaccine 25, 3137-3141.
- [0296] Estes M.K (1996). Rotavirus and their replication. Fields Virology 2, 1625-1655 Fabbretti, E., Afrikanova, I., Vascotto, F. and Burrone, O.R. (1999). Two nonstructural rotavirus proteins, NSP2 and NSP5, form viroplasm-like structures in vivo. J Gen Virol 80 333-9.
- [0297] Favacho, A.R., Kurtenbach, E., Sardi, S.I. and Gouvea, V.S. (2006). Cloning, expression, and purification of recombinant bovine rotavirus hemagglutinin, VP8*, in Escherichia coli. Protein Expression and Purification 46, 196-203.
- [0298] Gentsch, J.R., Laird, A.R., Bielfelt, B., Griffin, D.D., Banyai, K., Ramachandran, M., Jain, V., Cunliffe, N.A., Nakagomi, O., Kirkwood, CD., Fischer, T.K., Parashar, U.D., Bresee, J.S., Jiang, B. and Glass, R.I. (2005). Serotype Diversity and Reassortment between Human and Animal Rotavirus Strains: Implications for Rotavirus Vaccine Programs J Infect Dis. 192, S146-59.
- [0299] Glass, R.I., Parashar, U.D., Bresee, J.S., Turcios, R., Fischer, T.K., Widdowson, MA., Jiang, B. and Gentsch, J.R. (2006). Rotavirus vaccines: current prospects and future challenges. Lancet 368, 323-32.
- [0300] Gonzalez, A.M., Nguyen, T.V., Azevedo, M.S. P., Jeong, K., Agarib, F., Iosef, C, Chang, K., Lovgren-Bengtsson, K., Morein, B. and Saif, L.J. (2004). Antibody responses to human rotavirus (HRV) in gnotobiotic pigs following a new prime/boost vaccine strategy using oral attenuated HRV priming and intranasal VP2/6 rotavirus-like particle (VLP) boosting with ISCOM. Clinical & Experimental Immunology 135 361-372.
- [0301] Gonzalez, R.A., Espinosa, R., Romero, P., Lopez, S. and Arias C.F. (2000). Relative localization of viroplasmic and reticulum-resident rotavirus proteins in infected cells. Archives of Virology 145, 1963-1973.
- [0302] Greenberg, H.B. and Estes, M.K. (2009). Rotaviruses: From Pathogenesis to Vaccination. Gastroenterology 136, 1939-1951.

- [0303] Hoshino, Y., Jones, R.W. and Kapikian, A.Z. (1998). Serotypic characterization of outer capsid spike protein VP4 of vervet monkey rotavirus SA11 strain. *Archives of Virology* 143, 1233-1244.
- [0304] Istrate, C., Hinkula, J., Charpilienne, A., Poncet, D., Cohen, J., Svensson, L. and Johansen, K. (2008). Parenteral administration of RF 8-2/6/7 rotavirus-like particles in a one-dose regimen induce protective immunity in mice. *Vaccine* 26, 4594-4601.
- [0305] Khodabandehloo, M., Shamsi, S.M., Shahrabadi, Keyvani, H. and Bambai, B. (2009). Cloning and Expression of Simian Rotavirus Spike Protein (VP4) in Insect Cells by Baculovirus Expression System. *Iranian Biomedical Journal* 13 9-18.
- [0306] Kim, Y., Nielsen, P.R., Hodgins, D., Chang, K.O. and Saif, L.J. (2002). Lactogenic antibody responses in cows vaccinated with recombinant bovine rotavirus-like particles (VLPs) of two serotypes or inactivated bovine rotavirus vaccines. *Vaccine* 20, 1248-1258.
- [0307] Kovacs-Nolan J., Erika Sasaki, Dongwan Yoo, Yoshinori Mine, Cloning and Expression of Human Rotavirus Spike Protein, VP8*, in Escherichia coli. *Biochemical and Biophysical Research Communications*, 282, 1183-1188.
- [0308] Lawton, J.A., Estes, M.K. and Venkataram, P.B.V. (2000). Mechanism of genome transcription in segmented dsRNA viruses. *Advances in Virus Research* 55, 185-214.
- [0309] Lopez, T., Camacho M., Zayas M., Najera R., Sanchez R., Arias C. F. and Lopez S. (2005). Silencing the Morphogenesis of Rotavirus. *Journal of Virology* 79, 184-92.
- [0310] Lundgren, O. and Svensson, L. (2001). Pathogenesis of rotavirus diarrhoea. *Microbes and Infection* 3 1145-1156.
- [0311] Madore, H.P., Estes, M.K., Zarley, CD., Hu, B., Parsons, S., Digravio, D., Greiner, S., Smith, R., Jiang, B., Corsaro, B., Barniak, V., Crawford, S. and Conner, M.E. (1999). Biochemical and immunologic comparison of virus-like particles for a rotavirus subunit vaccine. *Vaccine* 17, 2461-2471.
- [0312] Marusic, C., Rizza, P., Lattanzi, L., Mancini, C., Spada, M., Belardelli, F., Benvenuto, E. and Capone, I. (2001). Chimeric plant virus particles as immunogens for inducing murine and human immune responses against human immunodeficiency virus type 1. *Journal of Virology* 75, 8434-8439.
- [0313] Matsumura, T., Itchoda, N. and Tsunemitsu, H. (2002). Production of immunogenic VP6 protein of bovine group A rotavirus in transgenic potato plants. *Archives of Virology* 147, 1263-1270.
- [0314] Mena, J.A., Ramirez, O.T. and Palomares, L.A. (2006). Intracellular distribution of rotavirus structural proteins and virus-like particles expressed in the insect cellbaculovirus system. *Journal of Biotechnology* 122, 443-452.
- [0315] Meyer, J.C., Bergmann, C.C. and Bellamy, A.R. (1989). Interaction of rotavirus cores with the nonstructural glycoprotein NS28. *Virology* 171, 98-107.
- [0316] Molinari, P., Peralta, A. and Taboga, O. (2008). Production of rotavirus-like particles in *Spodoptera frugiperda* larvae. *Journal of Virological Methods* 147, 364-367.
- [0317] Nilsson, M., von Bonsdorff, C.H., Wecliewicz, K., Cohen, J. and Svensson, L. (1998). Assembly of viroplasm and virus-like particles of Rotavirus by a Semliki Forest virus replicon. *Virology* 242, 255-265.
- [0318] Nishikawa, K., Fukuhara, N., Liprandi, F., Green, K., Kapikian, A., Chanock, R. and Gorziglia, M. (1989). VP4 protein of porcine rotavirus strain OSU expressed by a baculovirus recombinant induces neutralizing antibodies. *Virology* 173, 631-637.
- [0319] O'Brien, G.J., Bryant, C.J., Voogd, C., Greenberg, H.B., Gardner, R.C. and Bellamy, A.R. (2000). Rotavirus VP6 expressed by PVX vectors in Nicotiana benthamiana coats PVX rods and also assembles into virus-like particles. *Virology* 270, 444-453.

- [0320] Palombo, E.A. (1999). Genetic and antigenic diversity of human rotaviruses: potential impact on the success of candidate vaccines. *FEMS Microbiology Letters* 181, 1-8.
- [0321] Palomares, L.A. and Ramirez, O.T. (2009). Challenges for the production of viruslike particles in insect cells: The case of rotavirus-like particles. *Biochemical Engineering Journal* 45(3), 158-167.
- [0322] Peralta, A., Molinari, P. and Taboga, O. (2009). Chimeric recombinant rotavirus-like particles as a vehicle for the display of heterologous epitopes. *Virology Journal* 6, 192.
- [0323] Ramachandran, M., Kirkwood, CD., Unicomb, L., Cunliffe, N.A., Ward, R.L., Bhan, M.K., Clark, H.F., Glass, R.I. and Gentsch, J.R. (2000). Molecular characterization of serotype G9 rotavirus strains from a global collection. *Virology* 278, 436-444.
- [0324] Ribes, J.M., Ortego, J., Ceriani, J., Montava, R., Enjuanes, L. and Buesa, J. (2011). Transmissible gastroenteritis virus (TGEV)-based vectors with engineered murine tropism express the rotavirus VP7 protein and immunize mice against rotavirus. *Virology* 410 107-118.
- [0325] Rodriguez-Limas, W.A., Tyo, K.E.J., Nielsen, J., Ramirez, O.T. and Palomares, L.A. (2011). Molecular and process design for rotavirus-like particle production in *Saccharomyces cerevisiae*. *Microb Cell Fact.* 10, 33.
- [0326] Saldana, S., Esquivel Guadarrama, F., Olivera Flores Tde, J., Arias, N., Lopez, S., Arias, C., Ruiz-Medrano, R., Mason, H., Mor, T., Richter, L., Arntzen, C.J. and Gomez Lim, M.A. (2006). Production of rotavirus-like particles in tomato (*Lycopersicon esculentum* L.) fruit by expression of capsid proteins VP2 and VP6 and immunological studies. *Viral Immunology* 19, 42-53.
- [0327] Sanchez-Padilla, E., Grais, R.F., Guerin, P.J., Steele, A.D., Burny, M.E. and Luquero, F.J. (2009). Burden of disease and circulating serotypes of rotavirus infection in sub-Saharan Africa: systematic review and meta-analysis. *Lancet Infectious Diseases* 9, 567-576.
- [0328] Steele, A.D., Ivanoff, B. and African Rotavirus Network (2003). Rotavirus strains circulating in Africa during 1996-1999: emergence of G9 strains and P[6] strains. *Vaccine* 21, 361-367.
- [0329] Tian, P., Ball, J.M., Zeng, C.Q.Y. and Estes, M.K. (1996). The Rotavirus Nonstructural Glycoprotein NSP4 Possesses Membrane Destabilization Activity. *Journal of Virology* 70, 6973-6981.
- [0330] Thongprachum, A., Chaimongkol, N., Khamrin, P., Pantip, C., Mizuguchi, M., Ushijima, H. and Maneekarn, N. (2010). A novel multiplex RT-PCR for identification of VP6 subgroups of human and porcine rotaviruses, *Journal of Virological Methods*, 168, 191-196.
- [0331] Trabelsi, A., Peenze, I., Pager, C., Jeddi, M. and Steele, D. (2000). Distribution of rotavirus VP7 serotypes and VP4 genotypes Circulating in Sousse, Tunisia, from 1995 to 1999: Emergence of Natural Human Reassortants. *Journal of Clinical Microbiology* 38, 3415-3419.
- [0332] Varani, G. and Allain, F.H-T. (2002). How a rotavirus hijacks the human protein synthesis machinery. *Nature Structural Biology* 9, 158-160.
- [0333] Vende, P., Taraporewala, Z.F. and Patton, J.T. (2002). RNA-Binding Activity of the Rotavirus Phosphoglycoprotein NSP5 Includes Affinity for Double-Stranded RNA. *Journal of Virology* 76, 5291-5299.
- [0334] Vesikari, T., Karvonen, A., Korhonen, T., Espo, M., Lebacq, E., Forster, J., Zepp, F., Delem, A. and De Vos, B. (2004). Safety and immunogenicity of RIX4414 live attenuated human rotavirus vaccine in adults, toddlers and previously uninfected infants. *Vaccine* 22, 2836-2842.
- [0335] Vezina, L.-P., Faye, L., Lerouge, P., D'Aoust, M.-A., Marquet-Blouin, E., Burel, C., Lavoie, P.-O., Bardor, M. and Gomord, V. (2009). Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnology Journal* 7, 442-455.
- [0336] Zhou, B., Zhang, Y., Wang, X., Dong, J., Wang, B., Han, C., Yu, J., Li, D. (2010). Oral administration of plant-based rotavirus VP6 induces antigen-specific IgAs, IgGs and passive protection in mice.

Vaccine 28, 6021-6027.

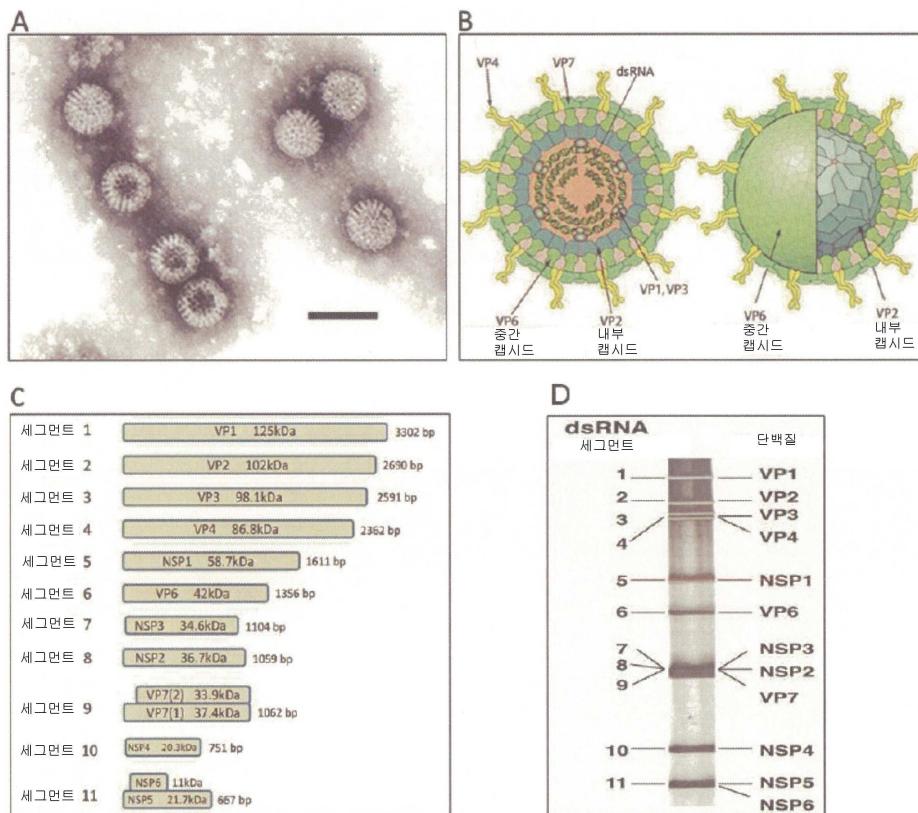
[0337] Zhou, H., Guo, L., Wang, M., Qu, J., Zhao, Z., Wang, J. and Hung, T. (2011). Prime immunization with rotavirus VLP 2/6 followed by boosting with an adenovirus expressing VP6 induces protective immunization against rotavirus in mice. *Virology Journal* 8, 3.

[0338] 인용된 모든 것들은 본원에 참고로 포함된다.

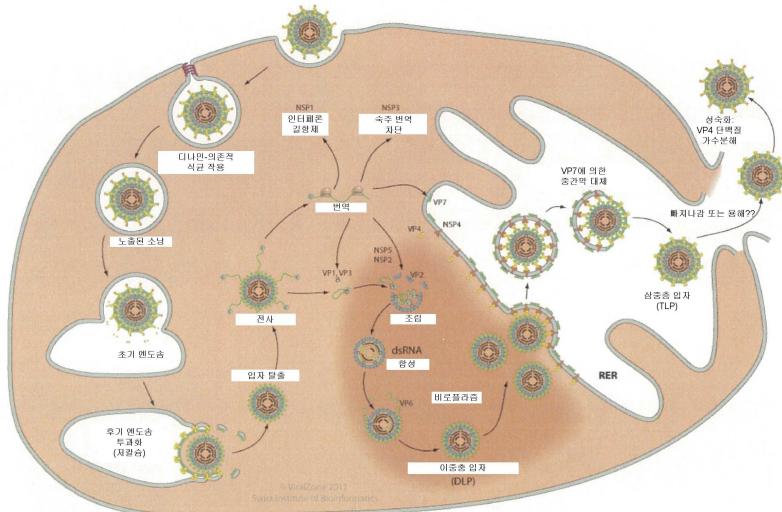
[0339] 본 발명은 하나 이상의 구체예에 관하여 설명되었다. 하지만, 청구범위에서 정의된 바와 같이 본 발명에서 벗어나지 않으면서 많은 변화 및 변형이 이루어질 수 있다는 것이 당업자에게 분명해질 것이다.

도면

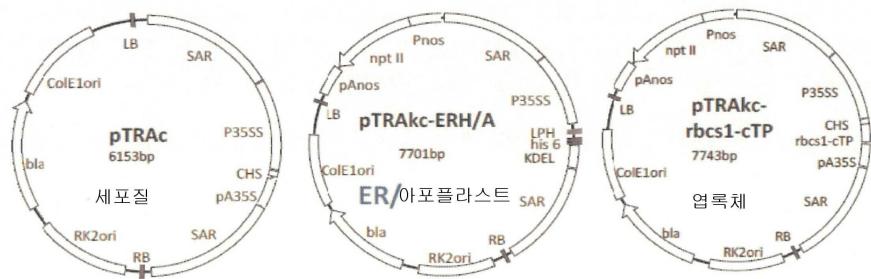
도면1



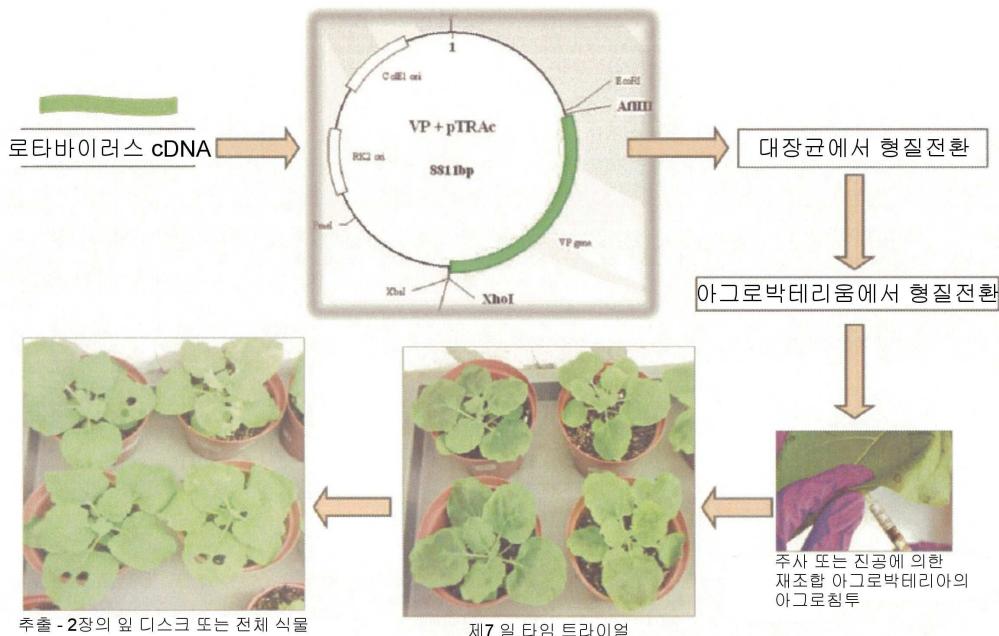
도면2



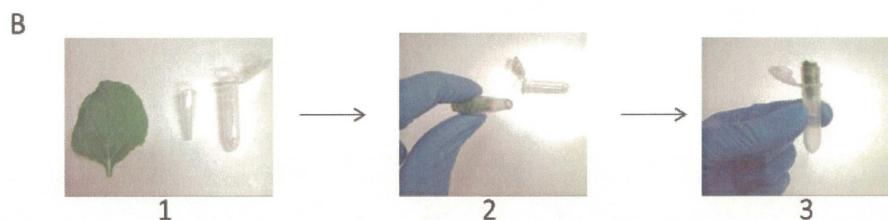
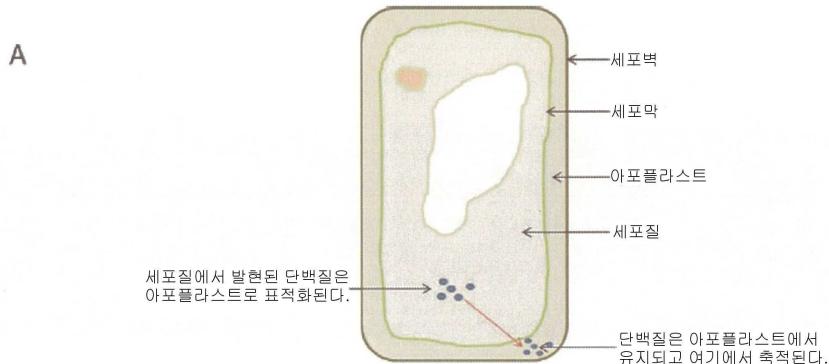
도면3



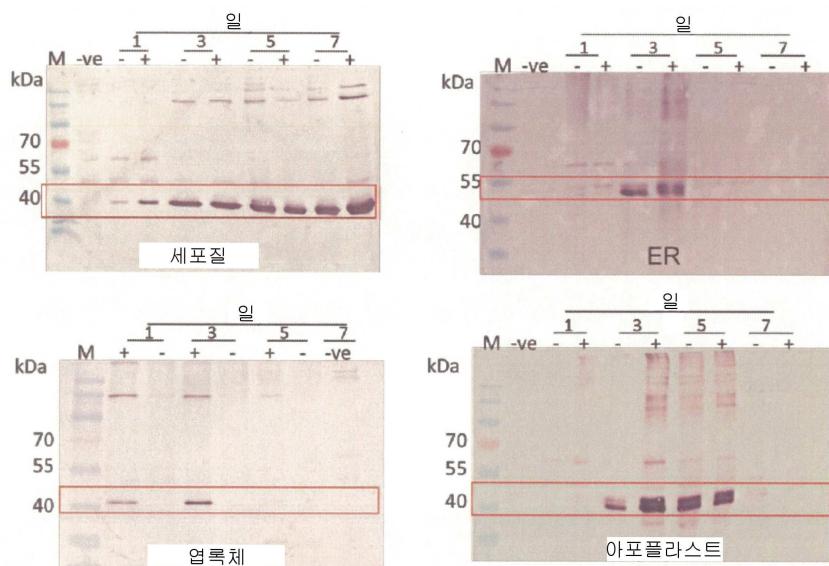
도면4



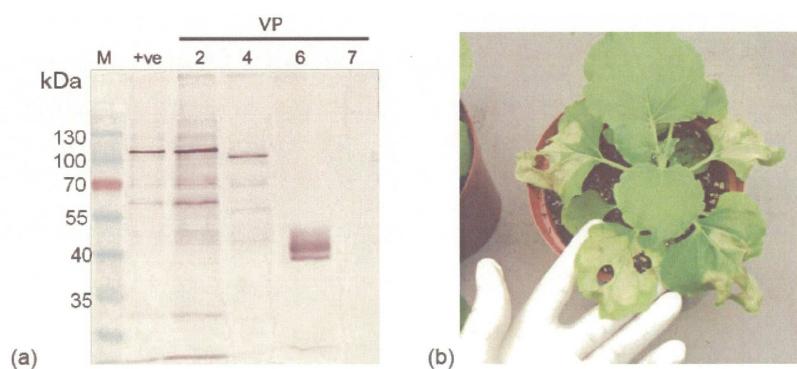
도면5



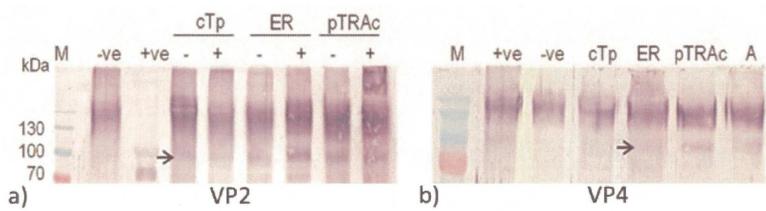
도면6



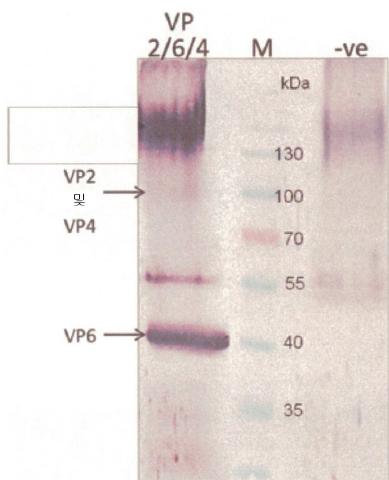
도면7



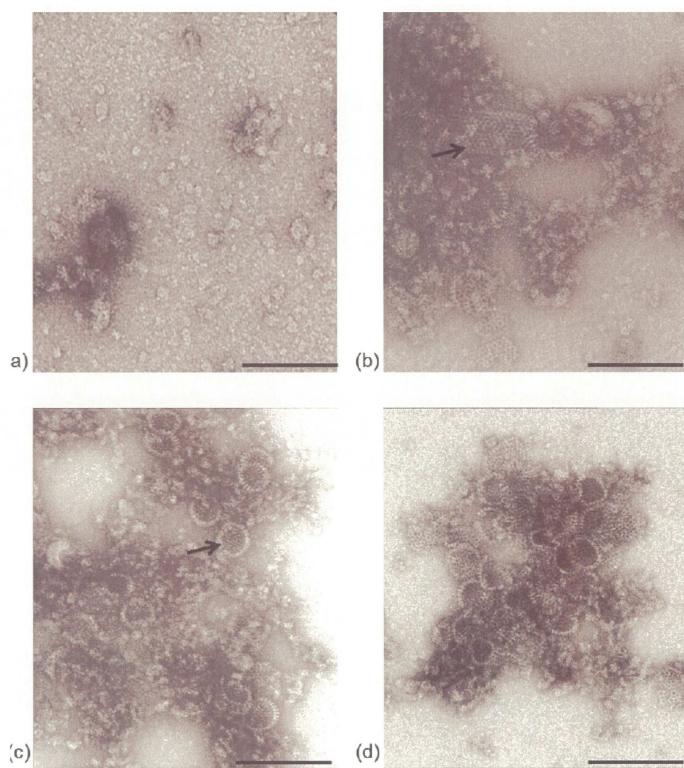
도면8



도면9



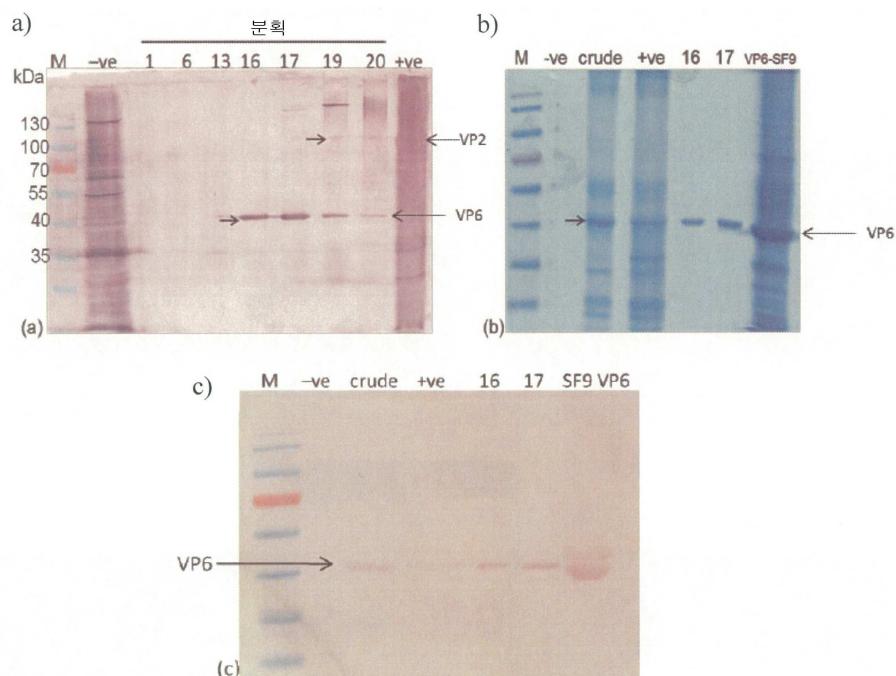
도면10



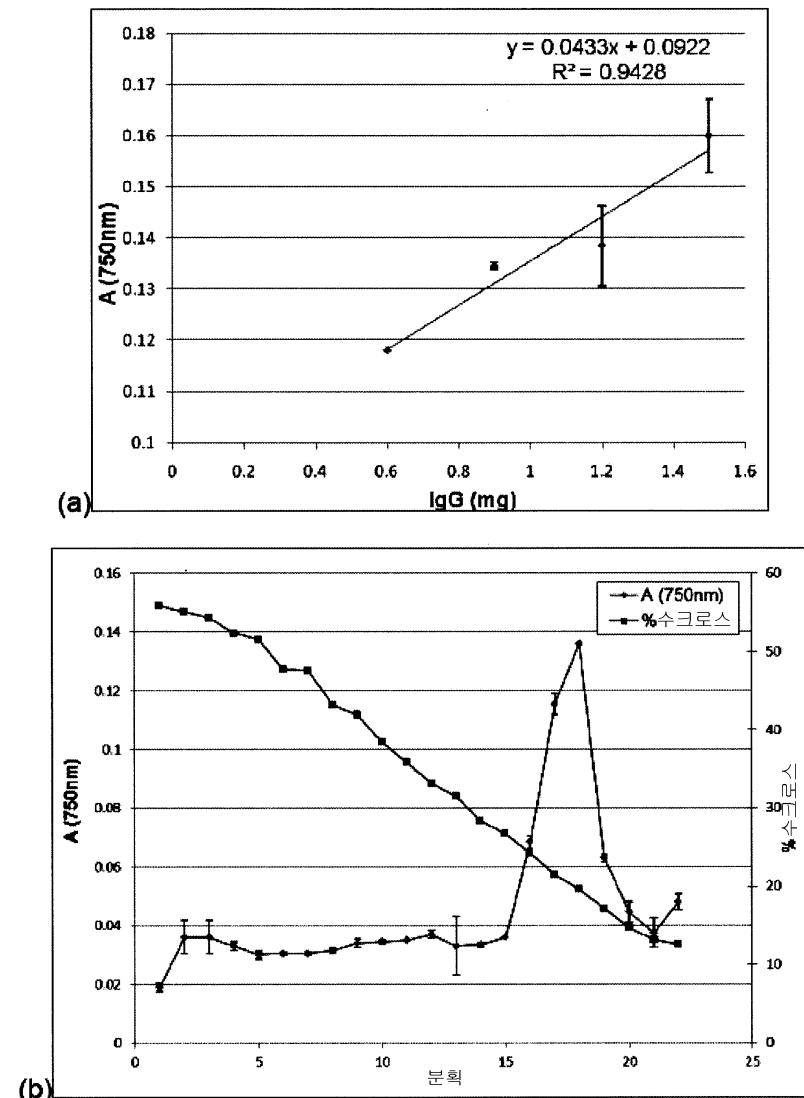
도면11



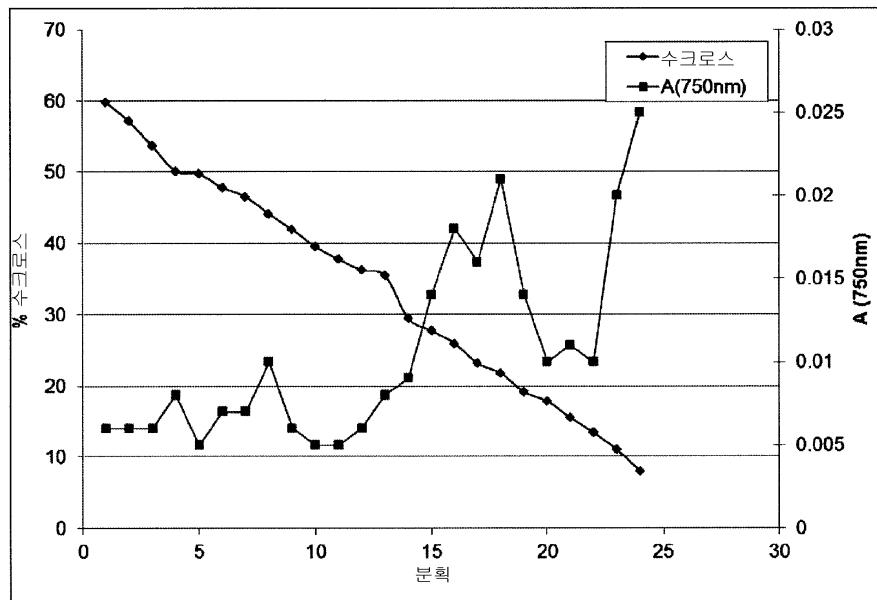
도면12



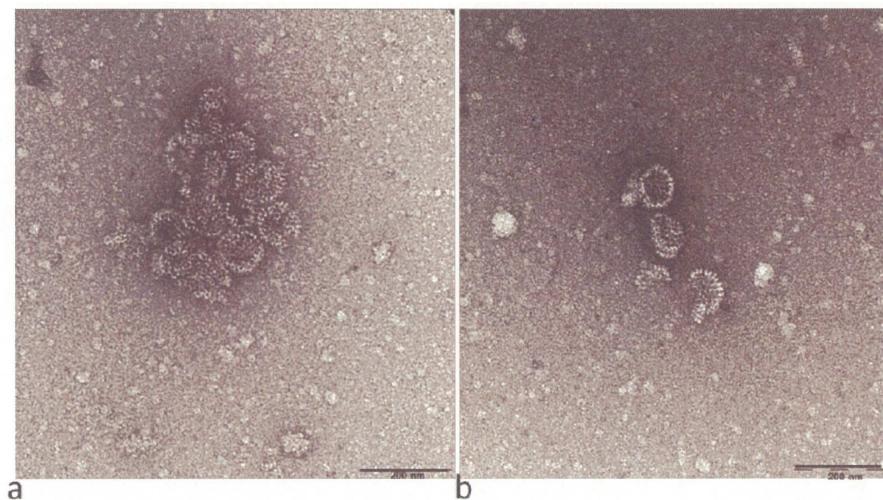
도면13



도면14



도면15



도면16a

로타바이러스 VP2

유전자명: G9P6_VP2

유전자 길이: 2700 bp

발현에 최적화됨: 니코티아나 타바抔

```

GGTACCGAATTGGACGCGTCGTTCCATGGCTTACCGTAAAGGGGTGCTAGGCCTGAA
1  -----+-----+-----+-----+-----+-----+
CCATGGCTTAAGCCTGCGCAAGCAAGGTACCGAATGGCATTTCCCCACGATCCGCACTT
SEQ ID NO:1 G_R_V_R_S_M_A_Y_R_K_R_G_A_R_R_E_
-----+-----+-----+-----+-----+-----+
GCTAACCTCAACAACAACGATAGGATGCAAGAGAAGATCGATGAGAAGCAGGATTCAAAC
61 -----+-----+-----+-----+-----+-----+
CGATTGGAGTTGTTGCTATCCTACGTTCTCTAGCTACTCTCGTCTAAGGTTG
A_N_L_N_N_D_R_M_Q_E_K_I_D_E_K_Q_D_S_N_
-----+-----+-----+-----+-----+-----+
AAGATCCAGCTCTCCGATAAGGTGCTCTCCAAGAAAAGAAGAGATCGTTACTGATTCCCAC
121 -----+-----+-----+-----+-----+-----+
TTCTAGGTCGAGAGGCTATTCCACGAGAGGTCTTCTCTAGCAATGACTAAGGGTG
K_I_Q_L_S_D_K_V_L_S_K_K_E_E_I_V_T_D_S_H_
-----+-----+-----+-----+-----+-----+
GAAGAGGTGAAGGTGACAGATGAGCTTAAGAAGTCCACAAAAGAAGAGTCCAAGCAGCTC
181 -----+-----+-----+-----+-----+-----+
CTTCTCCACTCCACTGTCTACTCGAATTCTCAGGTGTTTCTTAGGTCTCAGGTTCGAG
E_E_V_K_V_T_D_E_L_K_K_S_T_K_E_E_S_K_Q_L_
-----+-----+-----+-----+-----+-----+
CTTGAGGTGCTCAAGACAAAAGAGGAACACCAGAAAGAGATCCAGTACGAGATCCTCAA
241 -----+-----+-----+-----+-----+-----+
GAACCTCCACGAGTTCTGTTCTCCTGTTCTAGGTCTAGGTCTAGGAGGTT
L_E_V_L_K_T_K_E_E_H_Q_K_E_I_Q_Y_E_I_L_Q_

```

도면 16b

AAGACTATCCCAACTTCGAGCCAAAGAGACTATCCTCAGGAAGCTTGAGGATATCCAG
 301 -----+-----+-----+-----+-----+
 TTCTGATAGGGTTGAAAGCTCGGTTTCTCTGATAGGAGTCCTCGAACCTATAGGTC
 K_T_I_P_T_F_E_P_K_E_T_I_L_R_K_L_E_D_I_Q_

CCAGAGCTTGCTAAGAACGAGACTAAGCTCTCAGGATCTCGAGCAAAGCAGCTCCA
 361 -----+-----+-----+-----+-----+
 GGTCTCGAACGATTCTCGTCTGATTCGAGAACGCTTAGAAGCTCGGTTCGAGGGT
 P_E_L_A_K_K_Q_T_K_L_F_R_I_F_E_P_K_Q_L_P_

ATCTACCGTGCTAACGGTAAAGGGAACTTAGAACAGGTGGTACTGGAAGCTCAAGAAC
 421 -----+-----+-----+-----+-----+
 TAGATGGCACGATTGCCACTTCCCTGAAATCCTGTCACCATGACCTTCGAGTTCTC
 I_Y_R_A_N_G_E_R_E_L_R_N_R_W_Y_W_K_L_K_K_

GATACTCTCCCAGACGGTGATTACGATGTGAGAGAGTACTTCCTCAACCTACGATCAG
 481 -----+-----+-----+-----+-----+
 CTATGAGAGGGTCTGCCACTAATGCTACACTCTCATGAAGGAGTTGGAGATGCTAGTC
 D_T_L_P_D_G_D_Y_D_V_R_E_Y_F_L_N_L_Y_D_Q_

GTGCTCACTGAGATGCCAGATTACCTCCTCAAGGATATGGCTGTGGAGAACAGAAC
 541 -----+-----+-----+-----+-----+
 CACGAGTGACTCTACGGTCTAATGGAGGAGGAGTCTATACCGACACCTCTGTTCTG
 V_L_T_E_M_P_D_Y_L_L_L_K_D_M_A_V_E_N_K_N_

TCCAGGGATGCTGGAAAGGTGGATTCCGAGACTGCTTCATCTGTGATGCTATCTC
 601 -----+-----+-----+-----+-----+
 AGGTCCCTACGACCTTCCACCACCTAACGGCTTGACGAAGGTAGACACTACGATAGAAC
 S_R_D_A_G_K_V_V_D_S_E_T_A_S_I_C_D_A_I_F_

CAGGATGAAGAGACTGAGGGTGCTGTGAGGCAGGGTT
 661 -----+-----+-----+-----+-----+
 GTCCTACTTCTGACTCCCACGACACTCCGCAAAGTAACGACTCTACTCCGTCTCCAA
 Q_D_E_E_T_E_G_A_V_R_R_F_I_A_E_M_R_Q_R_V_

CAGGCTGATAGGAACGGTGAACCTACCCATCCATCCTCCACCCAATCGATTACGCTT
 721 -----+-----+-----+-----+-----+
 GTCCGACTATCCTTGCACCACTTGATGGTAGGTAGGAGGTTAGCTAATGCGAAAG
 Q_A_D_R_N_V_V_N_Y_P_S_I_L_H_P_I_D_Y_A_F_

AACGAGTACTTCCTTCAGCACCGAGCTGTGGAGCCACTCAACAACGATATCATCTCAAC
 781 -----+-----+-----+-----+-----+
 TTGCTCATGAAGGAAGTCGTGGTGAACACCTCGGTGAGTTGCTATAGTAGAGTTG
 N_E_Y_F_L_Q_H_Q_L_V_E_P_L_N_N_D_I_I_F_N_

도면 16c

TACATCCCAGAGAGGATTAGGAACGACGTTAACTACATCCTCAACATGGATAGGAACCTC
 841 -----+-----+-----+-----+-----+
 ATGTAGGGTCTCCTAATCCTGCTGCAATTGATGTAGGAGTTGTACCTATCCTGGAG
 Y_I_P_E_R_I_R_N_D_V_N_Y_I_L_N_M_D_R_N_L_

CCATCCACTGCTCGTTACATCAGGCCAACCTCCTCCAGGATAGGCTAACCTCACGAT
 901 -----+-----+-----+-----+-----+
 GGTAGGTGACGAGCAATGTAGTCGGTTGGAGGAGGTCTATCCGAGTTGGAGGTGCTA
 P_S_T_A_R_Y_I_R_P_N_L_L_Q_D_R_L_N_L_H_D_

AACTTCGAGTCCCTGGGATACAATCACTACTTCAACTACATTCTCGCTCGTCCGTG
 961 -----+-----+-----+-----+-----+
 TTGAAGCTCAGGGAGACCCCTATGTTAGTGATGAAGGTTGATGTAAGAGCGAGCAAGGCAC
 N_F_E_S_L_W_D_T_I_T_S_N_Y_I_L_A_R_S_V_

GTTGCCAGATCTCAAAGAACCTCGTGTCCACTGAGGCTCAGATCCAGAACGATGTCAGGAT
 1021 -----+-----+-----+-----+-----+
 CACGGTCTAGAGTTCTTGAGCACAGGTGACTCCGAGTCAGGTCTTACAGGGTCTA
 V_P_D_L_K_E_L_V_S_T_E_A_Q_I_Q_K_M_S_Q_D_

CTCCAGCTTGGGCTCTCACTATCCAGTCCGAGACTCAGTCCCTCACTGGTATCAACTCC
 1081 -----+-----+-----+-----+-----+
 GAGGTCGAACCTCGAGAGTGTAGGTCTGAGTCAGGAGTGAACAGGAGTGAACCATAGTTGAGG
 L_Q_L_E_A_L_T_I_Q_S_E_T_Q_F_L_T_G_I_N_S_

CAGGCTGCTAACGATTGCTCAAGACTCTCATTGCTGCTATGCTCTCCAGAGGACTATG
 1141 -----+-----+-----+-----+-----+
 GTCCGACGATTGCTAACGAAGTTCTGAGAGTAACGACGATAACGAGAGGGTCTCCTGATAC
 Q_A_A_N_D_C_F_K_T_L_I_A_A_M_L_S_Q_R_T_M_

TCCCTCGATTCGTGACTACTAACTATATGTCCTCATCTCCGAATGTGGCTTTGACT
 1201 -----+-----+-----+-----+-----+
 AGGGAGCTAAAGCACTGATGATGATACAGGGAGTAGAGGCTTACACCGAGAACTGA
 S_L_D_F_V_T_T_N_Y_M_S_L_I_S_G_M_W_L_L_T_

GTTGGTGCACACGATATGTTCATCCGTGAGTCCCTTGTGGCTTGCCAGCTCGCTATCGTG
 1261 -----+-----+-----+-----+-----+
 CACCACGGTTGCTATACAAGTAGGCACTCAGGAACACCGAACGGTCGAGCGATAGCAC
 V_V_P_N_D_M_F_I_R_E_S_L_V_A_C_Q_L_A_I_V_

AACACTATCATCTACCCAGCTTCGGAATGCAAAGGATGCACTACCGTAACGGTGTCCA
 1321 -----+-----+-----+-----+-----+
 TTGTGATAGTAGATGGGTCGAAAGCCTACGTTCTACGTGATGGCATTGCCACTAGGT
 N_T_I_I_Y_P_A_F_G_M_Q_R_M_H_Y_R_N_G_D_P_

도면 16d

CAGACTCCATTCCAGATCGCAGAGCAGCAGATCCAGAACTTCCAGGTGGCAAACCTGGCTC
1381 -----+-----+-----+-----+-----+
GTCTGAGGTAAGGTCTAGCGTCGTCTAGGTCTTGAAAGGTCCACCGTTGACCGAG
Q_T_P_F_Q_I_A_E_Q_Q_I_Q_N_F_Q_V_A_N_W_L_

CACTTCGTGAACAACAACCAGTTCAGGCAGGCTGTGATCGATGGTGTGTTGAACCAGGTG
1441 -----+-----+-----+-----+-----+
GTGAAGCACTTGTGTTGGTCAAGTCCGTCCGACACTAGCTACCCACACAACCTGGTCCAC
H_F_V_N_N_Q_F_R_Q_A_V_I_D_G_V_L_N_Q_V_

CTCAACGATAACATCCGTAACGGTACAGTGATCAACCAGCTCATGGAAGCTCTCATGCAA
1501 -----+-----+-----+-----+-----+
GAGTTGCTATTGTAGGCATTGCCAGTGCAGTAGTTGGTCGAGTACCTTCGAGAGTACGTT
L_N_D_N_I_R_N_G_H_V_I_N_Q_L_M_E_A_L_M_Q_

CTCTCCAGGCAGCAGTTCCAACATGCCTATCGATTACAAGCGTCCATCCAGAGGGGA
1561 -----+-----+-----+-----+-----+
GAGAGGTCGTCGTCAAGGGTTGATACGGATAGCTAATGTCGCAAGGTAGGTCTCCCT
L_S_R_Q_Q_F_P_T_M_P_I_D_Y_K_R_S_I_Q_R_G_

ATCCTCCTCCTTCCAACAGGCTTGGACAGCTCGTGGATCTCACTAGGCTCCTCGCTTAC
1621 -----+-----+-----+-----+-----+
TAGGAGGAGGAAAGGTTGTCGAACCTGTGAGCACCTAGAGTGTACCGAGGAGCGAATG
I_L_L_L_S_N_R_L_G_Q_L_V_D_L_T_R_L_L_A_Y_

AACTACGAGACTCTCATGGCTTGATCACTATGAACATGCAGCACGTTCAAGACTCTCACT
1681 -----+-----+-----+-----+-----+
TTGATGCTCTGAGAGTACCGAACGCTAGTGTACTTGTACGTCGTGCAAGTCTGAGAGTGA
N_Y_E_T_L_M_A_C_I_T_M_N_M_Q_H_V_Q_T_L_T_

ACTGAGAAGCTCCAGCTCACTTCCGTGACTTCCCTGCTCATGCTCATCGGAAACGCTACT
1741 -----+-----+-----+-----+-----+
TGACTCTCGAGGTGAGTGAAGGCACTGAAGGGAGACGTACGAGTAGCCTTGCATGA
T_E_K_L_Q_L_T_S_V_T_S_L_C_M_L_I_G_N_A_T_

GATCCCCATCCCCACAGACACTCTTCACTACTACAACGTGAACGTGAACCTCCACTCC
1801 -----+-----+-----+-----+-----+
CACTAGGGTAGGGGTGCTGTGAGAAGGTGATGATGTTGCACTTGCACTTGAAAGGTGAGG
V_I_P_S_P_Q_T_L_F_H_Y_Y_N_V_N_V_N_F_H_S_

AACTACAACGAGAGGGATCAACGATGCTGTGGCTATCATCACTGCTGCTAACAGGCTAAC
1861 -----+-----+-----+-----+-----+
TTGATGTTGCTCTCTAGTTGCTACGACACCGATAGTAGTGACGACGATTGTCCGAATTG
N_Y_N_E_R_I_N_D_A_V_A_I_I_T_A_A_N_R_L_N_

도면 16e

CTCTACCAAAAGAAGATGAAGGCTATCGTTGAGGATTCCCTAAGAGGCTCTACATCTTC
 1921 -----+-----+-----+-----+-----+
 GAGATGGTTTCTTCTACTTCCGATAGCAACTCCTAAAGGAGTTCTCCGAGATGTAGAAG
 L_Y_Q_K_K_M_K_A_I_V_E_D_F_L_K_R_L_Y_I_F__

 GATGTGTCAGGGTGCCAGATGATCAGATGTACCGTCTTAGGGATAGGCTTAGGCTCC
 1981 -----+-----+-----+-----+-----+
 CTACACAGGTCCCACGGCTACTAGTCTACATGCCAGAACATCCTATCCGAAATCCGAGGAG
 D_V_S_R_V_P_D_D_Q_M_Y_R_L_R_D_R_L_R_L_L__

 CCAGTGGAGATCAGAAGGCTCGATATCTTCAACCTCATCCTATGAACATGGATCAGATC
 2041 -----+-----+-----+-----+-----+
 GGTCACCTCTAGTCTCCGAGCTATAGAAGTTGGAGTAGGAATACTTGTAACCTAGTCTAG
 P_V_E_I_R_R_L_D_I_F_N_L_I_L_M_N_M_D_Q_I__

 GAGAGGGCTCCGATAAGATCGCTCAGGGTGTATTATCGCTTACCGTGATATGCACCTT
 2101 -----+-----+-----+-----+-----+
 CTCTCCCGAAGGCTATTCTAGCGAGTCCCACAATAATAGCGAATGGCACTATACGTGGAA
 E_R_A_S_D_K_I_A_Q_G_V_I_I_A_Y_R_D_M_H_L__

 GAGAGGGATGAGATGTACGGATACGTGAACATTGCTAGGAACCTTGAGGGATTCCAGCAG
 2161 -----+-----+-----+-----+-----+
 CTCTCCCTACTCTACATGCCCTATGCACTTGTAAACGATCCTGGAACTCCCTAAGGTCGTC
 E_R_D_E_M_Y_G_Y_V_N_I_A_R_N_L_E_G_F_Q_Q__

 ATCAACCTTGAAGAGCTTATGCCTTCCGGTGATTACGCTCAGATCACTAACATGCTCCTC
 2221 -----+-----+-----+-----+-----+
 TAGTTGGAACCTCTCGAACACGCAAGGCCACTAATGCGAGTCTAGTGATTGTACGAGGAG
 I_N_L_E_E_L_M_R_S_G_D_Y_A_Q_I_T_N_M_L_L__

 ACAACCAGCCAGTGGCTTGTGGTGCTCTCCATTCACTGATTCCCTCGTGATC
 2281 -----+-----+-----+-----+-----+
 TTGTTGGTCGGTCACCGAGAACAAACCACGAGAGGGTAAGTAGTGACTAAGGAGGCACTAG
 N_N_Q_P_V_A_L_V_G_A_L_P_F_I_T_D_S_S_V_I__

 TCCCTCATTGCTAAGTTGGATGCTACTGTGTTGCTCAGATCGTAAGCTCAGGAAAGTG
 2341 -----+-----+-----+-----+-----+
 AGGGAGTAACGATTCAACCTACGATGACACAAGCGAGTCTAGCACTTCGAGTCCTTCAC
 S_L_I_A_K_L_D_A_T_V_F_A_Q_I_V_K_L_R_K_V__

도면16f

GCTAACTACGATTGGGTGCCAACTTCCACTACAAAGGTGTACAAGCAGGTGCCACAGCAG
 2461 -----+-----+-----+-----+-----+
 CGATTGATGCTAACCCACCGTTGAAGGTGATGTTCCACATGTTCGCCACGGTGTGTC
A_N_Y_D_W_V_P_T_S_T_T_K_V_Y_K_Q_V_P_Q_Q

 TTCGATTTCGTAACTCCATGCACATGCTCACTTCAACCTCACTTGACTGTGACTCC
 2521 -----+-----+-----+-----+-----+
 AAGCTAAAGGCATTGAGGTACGTGTACGAGTGAAGGTTGGAGTGAAAGTGACACATGAGG
F_D_F_R_N_S_M_H_M_L_T_S_N_L_T_F_T_V_Y_S

 GATCTCCTCGCTTCTGTCGCTGATACTGTGGAGCCTATCAACGCTGTGGCTTCGAT
 2581 -----+-----+-----+-----+-----+
 CTAGAGGAGCGAAAGCACAGGGACTATGACACCTCGGATAGTTGCGACACCGAAAGCTA
D_L_L_A_F_V_S_A_D_T_V_E_P_I_N_A_V_A_F_D

 AACATGAGGATTATGAACGAGCTTGATGACTCGAGGGATCCTCTAGAGAA
 2641 -----+-----+-----+-----+
 TTGTAECTCTAAATACTTGCTCGAAACTACTGAGCTCCCTAGGAGATCTCTT
N_M_R_I_M_N_E_L_*_* SEQ ID NO: 1

 TTGAGCTC SEQ ID NO: 13
 -----+
 AAGCTCGAG SEQ ID NO: 14

도면 16g

로타바이러스 VP4

유전자명: G9P6_VP4

유전자 길이: 2388 bp

발현에 최적화된: 니코티아나 타바쿰

```

GGTACCGAATTGGACGCGTCGTTCCATGGCTTCCCTCATCTACCGTCAGTTGCTCACT
1   -----+-----+-----+-----+-----+
CCATGGCTTAAGCCTGCGCAAGCAAGGTACCGAAGGGACTAGATGGCAGTCAACGAGTGA
SEQ ID NO 2 G_R_V_R_S_M_A_S_L_I_Y_R_Q_L_L_T_

AACTCCTACACTGTGGAGCTTCCGATGAGATCAACACTATCGTTCCGAGAAGTCCCAG
61  -----+-----+-----+-----+-----+
TTGAGGATGTGACACCTCGAAAGGCTACTCTAGTTGTGATAGCCAAGGCTCTTCAGGGTC
N_S_Y_T_V_E_L_S_D_E_I_N_T_I_G_S_E_K_S_Q_

AACGTGACTATCAACCCAGGACCATTGCTCAGACTAACTACGCTCCAGTGACTTGGTCA
121 -----+-----+-----+-----+-----+
TTGCCACTTCACCTGCTAAGGTGATGATAGCTCGGTACCGAGCTACCTGGTATGGTCGGT
N_V_T_I_N_P_G_P_F_A_Q_T_N_Y_A_P_V_T_W_S_

CACGGTGAAGTGAACGATTCCACTACTATCGAGGCCAGTGCTCGATGGACCATACCAGCCA
181 -----+-----+-----+-----+-----+
GTGCCACTTCACCTGCTAAGGTGATGATAGCTCGGTACCGAGCTACCTGGTATGGTCGGT
H_G_E_V_N_D_S_T_T_I_E_P_V_L_D_G_P_Y_Q_P_

ACTAACTTCAAGCCACCAAACGATTACTGGATTCTCCTCAACCCAACAAACCAGCAGGTG
241 -----+-----+-----+-----+-----+
TGATTGAAGTTGGTGGTTGCTAATGACCTAAGAGGAGTTGGTTGATGGTCGTCCAC
T_N_F_K_P_P_N_D_Y_W_I_L_L_N_P_T_N_Q_Q_V_

GTGCTTGAGGAACTAACAGACTGATATCTGGGTGGCACTCCTTGTGGAGCCAAAC
301 -----+-----+-----+-----+-----+
CACGAACCTCCCTGATTGTTCTGACTATAGACCCACCGTGAGGAAGAACACCTCGGTTG
V_L_E_G_T_N_K_T_D_I_W_V_A_L_L_V_E_P_N_

```

도면 16h

AACAACACTAACAGTGGAGTTCTCGAGATGTTCAGGTCCAACGTGAACGCTGAGTTC
 421 -----+-----+-----+-----+-----+
 TTGTTGTGATTGTTCACCTCAAGAAGCTCTACAAGTCCAGGTTGCAC TGCGACTCAAG
 N_N_T_N_K_W_K_F_F_E_M_F_R_S_N_V_N_A_E_F_

CAGCACAAAGAGGACTCTCACTTCCGATACAAAGCTCGCTGGTTTATGAAGTTCTACAAAC
 481 -----+-----+-----+-----+-----+
 GTCGTGTTCTCCTGAGAGTGAGGCTATGTTGAGCGACCAAAACTCAAGATGTTG
 Q_H_K_R_T_L_T_S_D_T_K_L_A_G_F_M_K_F_Y_N_

TCTGTGTTGACTTTCCACGGTGAAACTCCACACGCTACTACTGATTACTCCTCCACTTCC
 541 -----+-----+-----+-----+-----+
 AGACACACCTGAAAGGTGCCACTTGAGGTGTCGATGACTAATGAGGAGGTGAAGG
 S_V_W_T_F_H_G_E_T_P_H_A_T_T_D_Y_S_S_T_S_

AACCTTCCGAGGTGGAGACTGTGATCCACGTGGAGTTCTACATCATCCAAAGGTCCAA
 601 -----+-----+-----+-----+-----+
 TTGGAAAGGCTCCACCTCTGACACTAGGTGCACCTCAAGATGTAGTAGGTTCCAGGGTT
 N_L_S_E_V_E_T_V_I_H_V_E_F_Y_I_I_P_R_S_Q_

GAGTCTAAGTGCTCCGAGTACATCAACACTGGACTCCCACCAATGCAAACACTAGGAAC
 661 -----+-----+-----+-----+-----+
 CTCAGATTACGAGGCTCATGTAGTTGACCTGAGGGTGGTACGTTGATCCTTG
 E_S_K_C_S_E_Y_I_N_T_G_L_P_P_M_Q_N_T_R_N_

ATCGTGCCAGTGGTTGTCCTCTCGTCCGTGACTTACCAAGAGGCTCAGGTGAACGGAG
 721 -----+-----+-----+-----+-----+
 TAGCACGGTCACCGAAACAGGAGAGCAAGGCACTGAATGGCTCCGAGTCCACTTGCTC
 I_V_P_V_A_L_S_S_R_S_V_T_Y_Q_R_A_Q_V_N_E_

GATATCATCATCTCCAAGACTTCCCTCTGAAAGAGATGCAGTACAACAGGGATATTATC
 781 -----+-----+-----+-----+-----+
 CTATAGTAGTAGAGGTTCTGAAGGGAGACCTTCTACGTATGTTGTCCTATAATAG
 D_I_I_I_S_K_T_S_L_W_K_E_M_Q_Y_N_R_D_I_I_

ATCAGGTTCAAGTTCAACAACTCCATCGTGAAGCTCGGAGGACTCGGATACAAGTGGAGT
 841 -----+-----+-----+-----+-----+
 TAGTCCAAGTTCAAGTTGAGGTGAGGACTTCGAGCCTCCTGAGCCTATGTTCACCTCA
 I_R_F_K_F_N_N_S_I_V_K_L_G_G_L_G_Y_K_W_S_

GAGATCTCCTCAAGGCTGCTAACTACCAAGTACTCCTACCTCAGGGATGGTGAACAGGTG
 901 -----+-----+-----+-----+-----+
 CTCTAGAGGAAGTCCGACGATTGATGGTCATGAGGATGGAGTCCCTACCAACTGTCAC
 E_I_S_F_K_A_A_N_Y_Q_Y_S_Y_L_R_D_G_E_Q_V_

도면 16i

ACAGCTCACACTACTGCTCCGTGAACGGTGTAAACAACCTCTCCTACAACGGTGGTTCC
 961 -----+-----+-----+-----+-----+
 TGTCGAGTGTGATGAACGAGGCACTTGCCACAAATTGTTGAAGAGGGATGTTGCCACCAAGG
 T_A_H_T_T_C_S_V_N_G_V_N_N_F_S_Y_N_G_G_S_

CTCCAACACTGATTTCTCCGTGTCAGGTTACGAGGTGATCAAAGAGAACTCCTACGTTAC
 1021 -----+-----+-----+-----+-----+
 GAGGGTTGACTAAAGAGGCACAGGGCAATGCTTCACTAGTTCTCTTGAGGATGCAAATG
 L_P_T_D_F_S_V_S_R_Y_E_V_I_K_E_N_S_Y_V_Y_

GTGGATTACTGGGATGATCCCCAGGCTTCAGGAACATGGTGTATGTTAGATCCCTCGCT
 1081 -----+-----+-----+-----+-----+
 CACCTAATGACCTACTAAAGGGTCCGAAAGTCCTGTACCACATACAATCTAGGGAGCAG
 V_D_Y_W_D_D_S_Q_A_F_R_N_M_V_Y_V_R_S_L_A_

GCTAACCTCAACTCCGTGAAGTGCTCCGGTGGAAACTACAACCTCCAGATCCCAGTGGGA
 1141 -----+-----+-----+-----+-----+
 CGATTGGAGTTGAGGCACCTCACGAGGCCACCTTGATGTTGAAGGTCTAGGGTCACCC
 A_N_L_N_S_V_K_C_S_G_G_N_Y_N_F_Q_I_P_V_G_

GCTTGGCCAGTGTGTCGGTGGAGCTGTGTCCTCCACTTCGCTGGTGTACACTCTCC
 1201 -----+-----+-----+-----+-----+
 CGAACCGGTCACTACAGACCACCTCGACACAGAGGGTGAAGCGACCACAATGTGAGAGG
 A_W_P_V_M_S_G_G_A_V_S_L_H_F_A_G_V_T_L_S_

ACTCAGTTCACTGATTCGTCCTCAACTCCCTCAGGTTCAAGGTTCTCCCTCACTGTG
 1261 -----+-----+-----+-----+-----+
 TGAGTCAGTCAAGTGACTAAAGCACAGGGAGTTGAGGGAGTCCAAGTCCAAAGAGGGAGTGACAC
 T_Q_F_T_D_F_V_S_L_N_S_L_R_F_R_F_S_L_T_V_

GAAGAGGCCACCATTCTCATCCTCAGGACTAGGGTCCGGACTTACGGACTCCCAGCT
 1321 -----+-----+-----+-----+-----+
 CTTCTCGGTGGTAAGAGGTAGGAGTCTGATCCCACAGGCCTGAAATGCCTGAGGGTCGA
 E_E_P_P_F_S_I_L_R_T_R_V_S_G_L_Y_G_L_P_A_

TTCAACCCAAACAACGGACACGAGTACTACGAGATCGCTGGACGTTCTCCCTATCTCC
 1381 -----+-----+-----+-----+-----+
 AAGTTGGGTTTGTGCTGTGCTCATGATGCTTAGCAGGCTGCAAAGAGGGAAATAGAGG
 F_N_P_N_N_G_H_E_Y_Y_E_I_A_G_R_F_S_L_I_S_

CTCGTGCATCCAACGATGATTACCAAGACTCCAATTATGAACCTCCGTACTGTGAGGCAG
 1441 -----+-----+-----+-----+-----+
 GAGCACGGTAGGTTGCTACTAATGGCTGAGGTTAATACTTGAGGCACTGACACTCCGTC
 L_V_P_S_N_D_D_Y_Q_T_P_I_M_N_S_V_T_V_R_Q_

도면16j

1501 GATCTTGAGAGGCAGCTCGGAGATCTCAGGGAAGAGTTCAACTCCCTCTCCAAGAGATC
 1501 CTAGAACTCTCCGTCGAGCCTAGAGTCCCTCTCAAGTTGAGGGAGAGGGTTCTCTAG
 1501 D_L_E_R_Q_L_G_D_L_R_E_E_F_N_S_L_S_Q_E_I_

1561 GCTATGACTCAGCTCATCGATCTCGCTCTCCTCCACTCGATATGTTCTCCATGTTCTCT
 1561 CGATACTGAGTCGAGTAGCTAGAGCGAGAGGGTGAGCTATAAGAGGTACAAGAGA
 1561 A_M_T_Q_L_I_D_L_A_L_L_P_L_D_M_F_S_M_F_S_

1621 GGTATCAA GTCCACTATCGATGTGGCTAAGTCTATGGTACTAAGGTGATGAAGAAGTTC
 1621 CCATAGTCAGGTGATAGCTACACCGATTCA GATA ACCACTGATTCCACTACTTCTTCAAG
 1621 G_I_K_S_T_I_D_V_A_K_S_M_V_T_K_V_M_K_K_F_

1681 AAGAAGTCCGGACTCGCTACTTCCATCTCCGAGCTTACTGGATCTCTCCAACGCTGCT
 1681 TTCTTCAGGCCTGAGCGATGAAGGTAGAGGCTCGAATGACCTAGAGAGAGGTGCGACGA
 1681 K_K_S_G_L_A_T_S_I_S_E_L_T_G_S_L_S_N_A_A_

1741 TCTTCTGTGCTCTAGGTCCCTCCATCAGGTCCAACATCTCCATCTCAGTGTGGACT
 1741 AGAACACAGATCCAGGAGGGAGGTAGTCCAGGTGTAGAGGAGGTAGAGTCACACCTGA
 1741 S_S_V_S_R_S_S_S_I_R_S_N_I_S_S_I_S_V_W_T_

1801 GATGTGTCCGAGCAGATCGCTGGATCTCCGATTCCGTGCGTAACATCTCCACTCAGACT
 1801 CTACACAGGCTCGTCTAGCGACCTAGAAGGCTAAGGCACGCATTGTAGAGGTGAGTCTGA
 1801 D_V_S_E_Q_I_A_G_S_S_D_S_V_R_N_I_S_T_Q_T_

1861 TCCGCTATCTCCAAGAGGCTTAGGCTCAGAGAGATCACTACTCAGACTGAGGGAATGAAC
 1861 AGGCGATAGAGGTTCTCGAATCCGAGTCTCTAGTGTAGAGTCTGACTCCCTACTTG
 1861 S_A_I_S_K_R_L_R_L_R_E_I_T_T_Q_T_E_G_M_N_

1921 TTGATGATATCTCCGCTGCTGTGCTCAAGACTAAGATCGATAGGTCCACTCACATCTCC
 1921 AAGCTACTATAGAGGCGACACGAGTTCTGATTCTAGCTATCCAGGTGAGTGTAGAGG
 1921 F_D_D_I_S_A_A_V_L_K_T_K_I_D_R_S_T_H_I_S_

1981 CCAGATACTCTCCCAGATATCATCACTGAGTCCTCCGAGAAGTTCATCCAAAGCGTGCT
 1981 GGTCTATGAGAGGGTCTATAGTAGTGACTCAGGAGGCTTCAAGTAGGGTTCGCACGA
 1981 P_D_T_L_P_D_I_I_T_E_S_S_E_K_F_I_P_K_R_A_

도면 16k

2041 TACCGTGTCTCAAGGATGATGAGGTGATGGAAGCTGATGTGGATGGAAAGTTCTCGCT
 -----+-----+-----+-----+-----+
 ATGGCACAAGAGTTCTACTACTCCACTACCTCGACTACACCTACCTTCAGAAGCGA
Y_R_V_L_K_D_D_E_V_M_E_A_D_V_D_G_K_F_F_A_

2101 TACAAAGTGGGAACTTCGAAGAGGTGCCATTGATGTGGATAAGTTCGTGGATCTCGTG
 -----+-----+-----+-----+-----+
 ATGTTTACCCCTGAAAGCTTCTCACCGTAAGCTACACCTATTCAAGCACCTAGAGCAC
Y_K_V_G_T_F_E_E_V_P_F_D_V_D_K_F_V_D_L_V_

2161 ACTGATTCCCCAGTGATCTCGCTATCATCGATTCAAGACTCTCAAGAACCTCAACGAT
 -----+-----+-----+-----+-----+
 TGACTAAGGGTCACTAGAGGCATAGTAGCTAAAGTTCTGAGAGTTCTGGAGTTGCTA
T_D_S_P_V_I_S_A_I_D_F_K_T_L_K_N_L_N_D_

2221 AACTACGGAATCACTAGGTCCCAGGCTCTCGATCTCATCCGATCCAAGGGTGCTC
 -----+-----+-----+-----+-----+
 TTGATGCCCTAGTGATCCAGGGTCCGAGAGCTAGAGTAGGCAAGGCTAGGTTCCCACGAG
N_Y_G_I_T_R_S_Q_A_L_D_L_I_R_S_D_P_R_V_L_

2281 AGGGATTTCATCAACCAGAACAAACCAATCATCAAGAACAGGATCGAGCAGCTCATTCTC
 -----+-----+-----+-----+-----+
 TCCCTAAAGTAGTTGGTCTTGGTTAGTAGTTCTGCTAGCTCGAGTAAGAG
R_D_F_I_N_Q_N_N_P_I_I_K_N_R_I_E_Q_L_I_L_

2341 CAGTGCCGTCTTGATGACTCGAGGGATCCTAGAGAATTGAGCTC SEQ ID NO: 15
 -----+-----+-----+-----+
 GTCACGGCAGAAACTACTGAGCTCCCTAGGAGATCTTAAGCTCGAG SEQ ID NO: 16
Q_C_R_L_*_*_ SEQ ID NO: 2

도면161

로타바이러스 VP6

유전자명: G9P6_VP6

유전자 길이: 1254 bp

발현에 최적화됨: 니코티아나 타바쿰

```

GGTACCGAATTCGGACGCGTTCGTCCATGGATGTGCTACTCCCTCCAAGACTCTC
1   -----+-----+-----+-----+-----+-----+
CCATGGCTTAAGCCTGCGCAAGCAAGGTACCTACACGAGATGAGGGAGAGGTTCTGAGAG
G_R_V_R_S_M_D_V_L_Y_S_L_S_K_T_L_


AAGGATGCTAGGGATAAGATCGTGGAGGGAACCTCTACTCCAACGTTCCGATCTCATC
61  -----+-----+-----+-----+-----+-----+
TTCCTACGATCCCTATTCTAGCACCTCCCTGAGAGATGAGGTTGCAAAGGCTAGAGTAG
K_D_A_R_D_K_I_V_E_G_T_L_Y_S_N_V_S_D_L_I_


CAGCAGTTCAACCAGATGATCATCACTATGAACGGAAACGAGTTCCAGACTGGTGBAAC
121 -----+-----+-----+-----+-----+
GTCGTCAAGTTGGTCTACTAGTAGTGATACTTGCCCTTGCTCAAGGTCTGACCACCTAG
Q_Q_F_N_Q_M_I_I_T_M_N_G_N_E_F_Q_T_G_G_I_


GGAAACCTCCAATCAGGAACCTGGACTTCGATTCGGACTCCTCGGAACACTCTCCTC
181 -----+-----+-----+-----+-----+
CCTTGAGGGTTAGTCCTTGACCTTGAGCTAAAGCCTGAGGAGCCTGATGAGAGGAG
G_N_L_P_I_R_N_W_N_F_D_F_G_L_L_G_T_T_L_L_


AACCTCGATGCTAACTACGTGGAGACTGCTAGAACACTATCGATTACTCGTTGATTTC
241 -----+-----+-----+-----+-----+
TTGGAGCTACGATTGATGCACCTCTGACGATCCTGTGATAGCTAATGAAGCAACTAAAG
N_L_D_A_N_Y_V_E_T_A_R_N_T_I_D_Y_F_V_D_F_


GTGGATAATGTGTGCATGGATGAGATGGTCGTGAGTCCCAGAGGAACGGAATTGCTCCA
301 -----+-----+-----+-----+-----+
CACCTATTACACACGTACCTACTCTACCAAGCACTCAGGGCTCCTTGCCCTAACGAGGT
V_D_N_V_C_M_D_E_M_V_R_E_S_Q_R_N_G_I_A_P_


CAGTCCGATTCCCTCAGGAAGCTCTCCGGTATCAAGTTCAAGAGGATCAACTCGATAAC
361 -----+-----+-----+-----+-----+
GTCAGGCTAAGGGAGTCCTCGAGAGGCCATAGTCAGTTCTCCTAGTTGAAGCTATTG
Q_S_D_S_L_R_K_L_S_G_I_K_F_K_R_I_N_F_D_N_


TCCTCCGAGTACATCGAGAACCTGGAAACCTCCAGAACAGAAGGCAGAGGACTGGATTCACT
421 -----+-----+-----+-----+-----+
AGGAGGCTCATGTAGCTTGCACCTGGAGGTCTGTCTCCGTCTCCTGACCTAACGAGTGA
S_S_E_Y_I_E_N_W_N_L_Q_N_R_R_Q_R_T_G_F_T_

```

도면 16m

TTCCACAAGCCAAACATCTTCCCATACTCCGCTTCCTCACTCTAACAGGTCCCAGCCA
 481 -----+-----+-----+-----+-----+
 AAGGTGTCGGTTGAGAAGGGTATGAGGCAGAGGAAGTGAGAGTTGTCAGGGTCGGT
F_H_K_P_N_I_F_P_Y_S_A_S_F_T_L_N_R_S_Q_P

GCTCACGATAACCTCATGGAACTATGGCTAACGCTGGTCTGAGATCCAGGTGGCA
 541 -----+-----+-----+-----+-----+
 CGAGTGCTATTGGAGTACCCCTGATACACCGAGTTGCGACCAAGACTCTAGGTCCACCGT
A_H_D_N_L_M_G_T_M_W_L_N_A_G_S_E_I_Q_V_A

GGATTGATTACTCCTGCGCTATCAACGCTCCAGCTAACACTCAGCAGTCGAGCACATC
 601 -----+-----+-----+-----+-----+
 CCTAAGCTAATGAGGACCGATAGTTGCGAGGTGATTGTGAGTCGTCAAGCTCGTAG
G_F_D_Y_S_C_A_I_N_A_P_A_N_T_Q_Q_F_E_H_I

GTTCAGCTCAGAAGGGTGTCACTACTGCTACTATCACTCTCCTCCCAGATGCTGAGAGG
 661 -----+-----+-----+-----+-----+
 CAAGTCGAGTCCTCCACGAGTGATGACGATGATAGTGAGAGGAGGGTACGACTCTCC
V_Q_L_R_R_V_L_T_T_A_T_I_T_L_L_P_D_A_E_R

TTCTCCTCCCAAGGGTGTCAACTCCGCTGATGGTGTACTACTTGGTACTTCAACCCA
 721 -----+-----+-----+-----+-----+
 AAGAGGAAGGGTCCCACTAGTTGAGGCAGTACACAGATGATGAACCATGAAGTTGGGT
F_S_F_P_R_V_I_N_S_A_D_G_A_T_T_W_Y_F_N_P

GTGATCCTCAGGCCAACACAGTGGAGGTGGAGTTCTCAACGGACAGATCATCAAC
 781 -----+-----+-----+-----+-----+
 CACTAGGAGTCCGGTTGTTGCACCTCACCTCAAGGAAGAGTTGCCTGTCTAGTAGTTG
V_I_L_R_P_N_N_V_E_V_E_F_L_L_N_G_Q_I_I_N

ACTTACCAAGGCTCGTTCGGAACTATCGTGGCTAGGAACATTGATACAATCAGGCTCTCC
 841 -----+-----+-----+-----+-----+
 TGAATGGTCCGAGCAAAGCCTTGATAGCACCGATCCTGAAGCTATGTTAGTCCGAGAGG
T_Y_Q_A_R_F_G_T_I_V_A_R_N_F_D_T_I_R_L_S

TTCCAGCTTATGAGGCCACAAACATGACTCCATCCGTGGCTGCACTCTCCAAACGCA
 901 -----+-----+-----+-----+-----+
 AAGGTGAAACTCCGGTGGTTGACTGAGGTAGGCACCGACGTGAGAAGGGTTGGT
F_Q_L_M_R_P_P_N_M_T_P_S_V_A_A_L_F_P_N_A

CAGCCATTGAGCACACGCTACTGTGGACTCACTCTAACAGATCGAGTCCGCTGTG
 961 -----+-----+-----+-----+-----+
 GTCGGTAAGCTCGTGGTGCATGACACCCCTGAGTGAGAGTTCTAGCTCAGGCACACAG
Q_P_F_E_H_H_A_T_V_G_L_T_L_K_I_E_S_A_V_C

도면 16n

1021 GAGTCTGTGCTCGCTGATGCTTCGGAGACTATGCTCGCTAACGTGACTTCTGTGAGGCAA
 CTCAGACACGAGCGACTACGAAGGCTCTGATACGAGCGATTGCACTGAAGACACTCCGTT
E_S_V_L_A_D_A_S_E_T_M_L_A_N_V_T_S_V_R_Q_

1081 GAGTACGCTATCCCAGTGGGACCAGTGTTCACCAGGAATGAACGGACTGATCTCATC
 CTCATGCGATAGGGTCAACCTGGTCACAAAGGTGGCCTTACTTGACCTGACTAGAGTAG
E_Y_A_I_P_V_G_P_V_F_P_P_G_M_N_W_T_D_L_I_

1141 ACTAACTACTCCCCATCCAGAGAGGATAACCTCCAGAGGGTGGTCACTGTGGCTTCATC
 TGATTGATGAGGGTAGGTCTCTCTATTGGAGGTCTCCCACAAGTGACACCGAAGGTAG
T_N_Y_S_P_S_R_E_D_N_L_Q_R_V_F_T_V_A_S_I_

1201 CGTTCCATGCTCGTGAAGTGATGACTCGAGGGATCCTCTAGAGAA
 GCAAGGTACGAGCACTCACTACTGAGCTCCCTAGGAGATCTCTT
R_S_M_L_V_K_*_*_ SEQ ID NO: 3

TTCGAGCTC SEQ ID NO: 17

AAGCTCGAG SEQ ID NO: 18

도면16o

로타바이러스 VP7

유전자명: G9P6_VP7

유전자 길이: 1041 bp

발현에 최적화됨: 니코티아나 타바쿰

```

GGTACCGAATTCCGACGCGTCGTTGCATGTACGGAATCGAGTACACTACTATCCTCACT
1   -----+-----+-----+-----+-----+
CCATGGCTTAAGCCTGCGCAAGCAACGTACATGCCCTAGCTCATGTGATGATAGGAGTGA
SEQ ID NO:4 G_R_V_R_C_M_Y_G_I_E_Y_T_T_I_L_T

TTCCCTCATCTCCATCGTGCCTCAACTACATCCTCAAGTCCCTCACTCCGCTATGGAT
61  -----+-----+-----+-----+-----+
AAGGAGTAGAGGTTAGCACGAGGAGTTGATGTAGGAGTCAGGGAGTGAAGGCATACCTA
F_L_I_S_I_V_L_L_N_Y_I_L_K_S_L_T_S_A_M_D

TTCATCATCTACCGTTCCCTCCTCATCGTGATCGCTCCCCATTGTTAACAGACTCAG
121 -----+-----+-----+-----+-----+
AAGTAGTAGATGGCAAAGGAGGAGTAGCACTAGCGAAGGGTAAGCAATTCTGAGTC
F_I_I_Y_R_F_L_L_I_V_I_A_S_P_F_V_K_T_Q

AACTACGGTATCACCTCCAATCACTGGATCTATGGACTTGCTTACGCTAACCTCC
181 -----+-----+-----+-----+-----+
TTGATGCCATAGTTGGAGGGTTAGTGACCTAGATACTATGACGAATGCGATTGAGGAGG
N_Y_G_I_N_L_P_I_T_G_S_M_D_T_A_Y_A_N_S_S

CAGCAAGAGACTTCCCTACTCCACACTCGCTCTACTACCAACTGAGGCATCCACA
241 -----+-----+-----+-----+-----+
GTCGTTCTCTGAAAGGAGTGAAGGTGTGAGACGGAGATGGGTTGACTCCGTAGGTGT
Q_Q_E_T_F_L_T_S_T_L_C_L_Y_Y_P_T_E_A_S_T

CAGATCGGAGATAACAGAGTGGAGGATACTCTCTCCAGCTTCCCTCACTAAGGGATGG
301 -----+-----+-----+-----+-----+
GTCTAGCCTCTATGTCTCACCTCTATGAGAGAGGGTCGAGAAGGAGTGATTCCCTACC
Q_I_G_D_T_E_W_K_D_T_L_S_Q_L_F_L_T_K_G_W

CCAACTGGTTCCGTGACTTCAAAGAGTACACTGATATCGCTTCCCTCATCGATCCA
361 -----+-----+-----+-----+-----+
GGTGACCAAGGCACATGAAGTTCTCATGTGACTATAGCGAAGGAAGAGGTAGCTAGGT
P_T_G_S_V_Y_F_K_E_Y_T_D_I_A_S_F_S_I_D_P

```

도면 16p

CAGCTCTACTGCGATTACAACGTGGTGCTTATGAAGTACGATCCACTCTTGAGCTTGAT
 421 -----+-----+-----+-----+-----+
 GTCGAGATGACGCTAAATGTTGCACCAACGAATACTTCATGCTAAGGTGAGAACTCGAAGTA
 Q_L_Y_C_D_Y_N_V_V_L_M_K_Y_D_S_T_L_E_L_D_

ATGTCGAGCTTGCTGATCTCATCCTAACGAGTGGCTTGCAACCCAATGGATATCACT
 481 -----+-----+-----+-----+-----+
 TACAGGCTCGAACGACTAGAGTAGGAGTTGCTCACCGAGACGTTGGGTTACCTATAGTGA
 M_S_E_L_A_D_L_I_L_N_E_W_L_C_N_P_M_D_I_T_

CTCTACTACTACCAGCAGACTGATGAGGCTAACAAAGTGGATCTATGGGACAGTCCTGC
 541 -----+-----+-----+-----+-----+
 GAGATGATGATGGTCGCTTGACTACTCCGATTGTTCACCTAGAGATAACCTGTCAGGACG
 L_Y_Y_Y_Q_Q_T_D_E_A_N_K_W_I_S_M_G_Q_S_C_

ACTATCAAAGTGTGCCACTAACACTCAGACTCTCGGAATCGGATGCATCACTAAC
 601 -----+-----+-----+-----+-----+
 TGATAGTTTCACACGGGTGAGTTGAGTCTGAGAGCCTAGCCTACGTAGTGATGATTG
 T_I_K_V_C_P_L_N_T_Q_T_L_G_I_G_C_I_T_T_N_

ACTGCTACTTCGAGGAAGTGGCTACTCCGAGAACGCTCGTGTGATGACTGATGTGGTGGAT
 661 -----+-----+-----+-----+-----+
 TGACGATGAAAGCTCCCTCACCGATGAAGGCTTCAGCAGACTAGTGACTACACCACCTA
 T_A_T_F_E_E_V_A_T_S_E_K_L_V_I_T_D_V_V_D_

GGTGTAAACCACAAGCTCGATGTGACTACTAACACATGCACAAATCAGGAACATGCAAGAAG
 721 -----+-----+-----+-----+-----+
 CCACAATGGTGTGAGCTACACTGATGATTGTTACGTGTTAGTCCTGACGTTCTTC
 G_V_N_H_K_L_D_V_T_T_N_T_C_T_I_R_N_C_K_K_

CTCGGACCAAGGGAAAACGTGGCTATCATCCAAGTGGGAGGTTCCGATGTGCTCGATATC
 781 -----+-----+-----+-----+-----+
 GAGCCTGGTCCCTTTGCACCGATAGTAGGTTCACCCCTCCAAGGCTACAGGAGCTATAG
 L_G_P_R_E_N_V_A_I_I_Q_V_G_G_S_D_V_L_D_I_

ACTGCTGATCCAACACTGCTCCACAGACTGAGAGGATGATGAGGGTGAACATGGAAGAAG
 841 -----+-----+-----+-----+-----+
 TGACGACTAGGTTGATGACGAGGGTGTGACTCTCCTACTACTCCCACCTGACCTTCTTC
 T_A_D_P_T_T_A_P_Q_T_E_R_M_M_R_V_N_W_K_K_

TGGTGGCAGGTTTCTACACTGTGGGATTACATCAACCAGATGTTAGGTGATGTCC
 901 -----+-----+-----+-----+-----+
 ACCACCGTCCAAAAGATGTGACACCACCTAATGTAGTTGGTCTAGCAAGTCCACTACAGG
 W_W_Q_V_F_Y_T_V_V_D_Y_I_N_Q_I_V_Q_V_M_S_

도면 16q

AAGAGGTCCCGTTCTCAACTCCGCTGCTTCTACTACCGTGTGATGACTCGAGGGA
 961 -----+-----+-----+-----+-----+
 TTCTCCAGGGCAAGAGAGTTGAGGCGACGAAAGATGATGGCACACACTACTGAGCTCCCT
 K_R_S_R_S_L_N_S_A_A_F_Y_Y_R_V_*_* SEQ ID NO: 4

TCCTCTAGAGAATCGAGCTC SEQ ID NO: 19
 1021 -----+-----+
 AGGAGATCTCTTAAGCTCGAG SEQ ID NO: 20

도면17

도 17

A-2X35S/CPMV-HT/ RVA(WA) VP2(opt)/ NOS (구조 번호 1710)

도 17A, SEQ ID NO: 21

IF-WA_VP2(opt).s1+3c

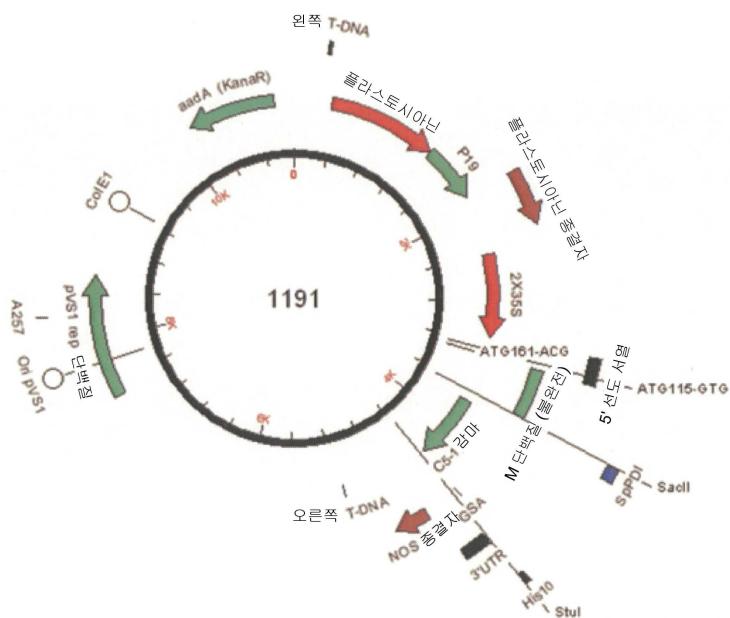
AAATTGTCGGGCCATGGCATACCGGAAGAGAGGAGCAAAGCGCGAA

도 17B, SEQ ID NO: 22

IF-WA_VP2(opt).s1-4r

ACTAAAGAAAATAGGCCTTAAAGCTCGTTATTATCGCATATTGTCGA

도 17C



도면 18a

도 18, SEQ ID NO: 23

원쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 구조 1191 (밀줄 그어져 있음).
플라스토시아닌-P19-플라스토시아닌 침목 억제자 발현 카세트가 있는
2X35S/CPMV-HT/NOS

도면 18b

CACACAGCTAGACGCACCCCCGGGAGGAGCAGTCAACAGCACTTCCGCTCAGTCAGTGAACCTC
CCATCATGCAACCAGACTGGCTCAATGGCAAGGGAGGCGATCGCTCACCATCACCATCACCAT
CACCTAAAGGCCATTITCTTAGTGTAACTGATTTACTGTATTTCGGTGTGCAATTCTATGTTGGTA
GGCGTTTCTGTGCTCAGATGTTATTATTGTAACTTAAATTCTTGTGAGCTCTGTAGCAG
GTCTGTCCTCTAGCAAGGGACACAAAAGATTTAATTAAAAAAAAGACCGG
GAATTCTGATATAACGCTTACAGCTGACCTGCAAGATCTGCAACATTGGCAATAAGATCTTAAAGATIG
AACTCTGTTGGGCTTCTGGATGATTATCATATAATTCTGTGATTAACCTGGATCTAGTGAATAAT
TAACATGTAATGCTGACGTTATTATGAGATGGGTTTTATGATTAGACTCCCGCAATTATACATT
AAATACGGCTAGAAAAGAAAATATAGCCGCCAACATAGGATAAAATTCTGCGCAGGGCTGTCTATCA
GTTACTAGATCTAGAGCTCAAGCTTGGGCCCAAGCTGACTAGTGGCACTGGCGCTGTCTATTAC
AACGTCGTGACTGGGAAACCTGGCGTACCCAACTTAATCGCCTTGACGACATCCCCCTTCGCC
AGCTGGCTAATAGCGAAGAGGCCGACCGATGCCCTTCCAAAGCTTGGCGACGCTGAATGGCG
AATGCTAGAGCTGACTGGATCAGTGTGTTCTCCGCACTGTTAACATGCTTAAACATCAGTGT
GACAGGATATACTGGGGTAAACCTAAAGGAAAAGGGCTTA

도면19

도 19 로타바이러스 A WA 군주의 뉴클레오티드 서열 암호화 VP2(opt)
(SEQ ID NO: 45)

```

ATGGCATACCGGAAGAGAGGAGCAAAGCGCAAACCTGCCAACAGAACGAGAGA
CTGCAAGAAAAAGAGATAGAGAAAGATGTCGACGTAACAATGAAAACAAGAATAAC
AATAGGAAACAAACAGCTGTCGACAAAGTCTGTCCCAGAAGGAGGAATTATCACT
GACGCCAGGAGCATATAAAGCCGAGAAATAAGAAGAGCTGAAAGAGAAGAA
TCTAACACAGCTGTCGAAATCTGAAAACAAAGACCATCAGAAAGAGATTCAA
TATGAAATTTCGCAAAAACAATACCTACATTGAGTCAAAGAAAGTATCCTCAAG
AAGCTGAAGACATAAGACGGAGCAGGCAAAAAACAGATGAAACTCTTCGCAATT
TTCGAGCAAACAGCTCCCTATATATCGCGCAATGGCAGAAGGAGCTACGCAAC
CGGTGGACTGGAAGATTGAAAAAGACACCCCTGCCAGATGGAGATTATGACGTCCGG
GAGTATTTCCTCAATCTCTATGATCAGATCCTCATCGAAATGCCGACTATCTGCTC
CTCAAGGACATGCCGTGGAGAACAAAAATAGCAGAGCAGCCGAAAGTTGTCGAC
TCTGAGACTGCCAATATTGATGATGCCATCTCCAGGATGAGGAGACCGAGGGAGTC
GTCCGTAGATTCTCGTATGCGCAACAGGTCCAGGCTGATCGTAACATTGTC
ATTACCCCTCCATCCTCACCCATTGATCATGCATTCAATGAGTATTCTAAC
CACCACTGGTGGAGCCGCTGAACATGAGATAATCTCAATTACATACCAGAGAGG
ATAAGGAATGACGTGAATTACATCCTGAAACATGGATATGAATCTGCATCTACAGCC
AGGTATATCAGGCCAAACTTGTGAGGATAGACTGAATCTTCAGATAATTTGAG
TCCCTGTGGGATACCACACAACTACATTCTGGCAGGTCCGTGTTCCC
GATTGAAAGGAGAAGGAGCTGGTCTCACCAGAACAGATCCAGAAAATGAGCAG
GACCTGCAGCTGGAGGCCCTCACTATTGAGGAGCACACAGTTTAGCCGGGATT
AACAGTCAGGCTCCAATGATTGTTCAAGACCCCTCATAGCCGCACTGCTCTCAA
AGAACCATGCTTGGACTTTGAGCACCAGAACTATATGAGCCTAATCTCCGGAATG
TGGACTTACAGTGAATTCCACGATATGTTCTCCGGGAGTCAGTAGTGGCTGT
GAGCTGGCAGTCATCAACACCATCGTGTATCCAGCATTGGAATGAGAATGCAT
TACCGGAATGGCAGCCCTCAGACACCCCTCAGATCGCAGAACAGCAGATCCAGAAT
TTCCAGGTGGGAACGGCTCCATTATAACAATAACAGATTAGGCAAGTTGTG
ATTGATGGAGTTCTGAATCAGACTCTGAACGACAATATACGGAATGGACAGGTCA
AACAGCTGATGAAAGCATTGACTGAACTCAGCAGAACAGCAGCTCCACGATGCC
GTGGATTACAAACGGGAGCATCCAACGGGATTCTGCTCTCCAAATAGGCTGGG
CAGCTTGTCGACTTAACCCGACTGGTCTCTATAACTACGAGACGCTAATGGCTGT
GTGACCATGAAACATGCAGCACGTGCAAACCCCTGACAACTGAGAAGTTGAGCTCA
TCTGTGACTTCGTTGTATGTTATTGTTAACACAACCGTGTACGGCTCCCCACAG
ACACTGTTCACTACTACAACATCAACGTGAATTCCACTCCAATTATAATGAGGG
ATCAACGACGCCGCGCATAATTACCGCAGCAAATAGGCTGAATCTTATCAGAAA
AAAATGAAGTCCATAGTGGAGACTTCTGAAACGGCTCCAGATTTCGACGTACCA
CGAGTGCTGACGACCAAATGTACGGCTGAGGGATGCCCTCGCTCTACCGTT
GAAAGGAGACGGCTGACATATTCAACTTGATCTGTATGAAATATGGAGCAGATCGA
CGCCTCTGATAAGATTGCTCAGGGGTTATCATCGCATAACCGAGATATGAGCTG
GAACCGCAGAGATGTACGGATATGTTAATATTGACCGGAATCTGATGGCTACCA
CAAATTAACCTGGAGGAACCTATGCGCACCGGTGATTACGGACAAATTACGAAACATG
CTTCTCAACAATCAACCGTTGCCCTTGTTGCGATTGCCCTCGTTACGGACTCA
TCCGTGATCAGCTAATGCCAACGCTCGACGCAACCGTCTCGCTCAGATAGTGAAG
CTCAGGAAAGTTGACACACTGAAAGCCATACTGTACAAAATAACTCGGATTCCAAT
GACTTTACCTTGCCAACTACGACTGGATCCCCACAAGTACAACATAAGGTCTAC
AAACAGGTGCCACAACCATTGCACTTAGAGGCCAGCATGCACATGCTGACTTCAAC
CTTACGTTACCGTCACTCTGACCTACTGTCATTGTTAGCGGACACGGTAGAG
CCCATTAACGCACTGCAGTCGAAATATGCGAATAATGAACGAGCTTAA

```

도면20

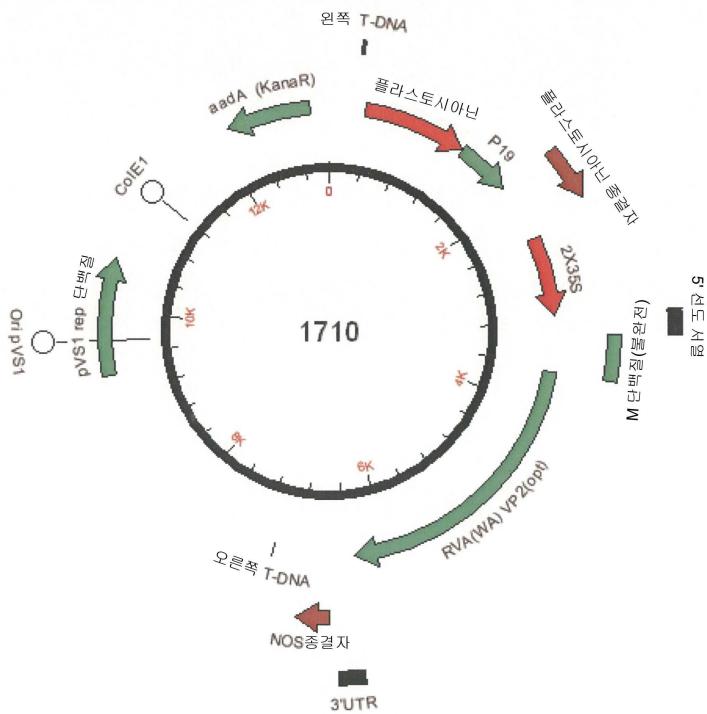
도 20, SEQ ID NO: 25
로타바이러스 A WA 군주의 VP2의 아미노산 서열

```

MAYRKRGAKRENLPQQNERLQEKEIEKDVDTVMENKNNRKQQLSDKVLSQKEEITDAQDDIKIAGEIK
KSSKEESQLLEILKTKEDHQEIQYELQKTIPTFESKESILKKLEDIRPEQAKKQMKLFRIFEPKQLPIYRA
NGEKELRNRYWVWLKKDTLPDGYDVREYFLNLYDQILIEMPDYLKLKMAVENKNSRDAGKVVDSE
TANICDAIFQDEETEGVVRRFIADMRRQVQADRNIYNPSILHPIDHAFNEYFLNHQLVEPLNNEIIFNYIPE
RIRNDVYILNMMDMNLPSLARYIRPNLLQDRNLHDNFESLWDTITSNYILARSVVPDLKEKELVSTEAQ
IQKMSQDLQLEALTIQSETQFLAGINNSQAANDCFKTLIAAMLSQRMSLDFVTNYSMSLISGMWLLTVIPN
DMFLRESLVACELAINTIVYPAFGMQRMHVRNGDPQTPQIAEQQQIQNQFQVANWLHFINNNRFRQVVID
GVLNQTLNDNIRNGQVINQLMEALMQLSRQQFPMPDYKRSIQRGLLSSNRLGQLVDLTRLVSYNYET
LMACVTNMQHVQTLTTEKLQLTSVSLCMLIGNTTVPSPQTFLHYHNVNHFHSYNERINDAVAITA
ANRLNLYQKKMKSVIEDFLKRLQIFDVPRVPDDQMYRLRDRRLLPVERRLDIFNLLJLMNMEQIERASD
KIAQGVIIAYRDMQLERDEMVGYVNIARNLDGYQQINLEELMRTGDYQQTNMLLNNPVALVGALPFV
TDSSVISLIAKLDAVFAQIVKLKRVDTLKPILYKINSDSNDFYLVANYNDWIPSTTKVYKQPQPDFRAS
MHMLTSNLTVSDLSFVSAADTVEPINAVAFDNMRIMNEL

```

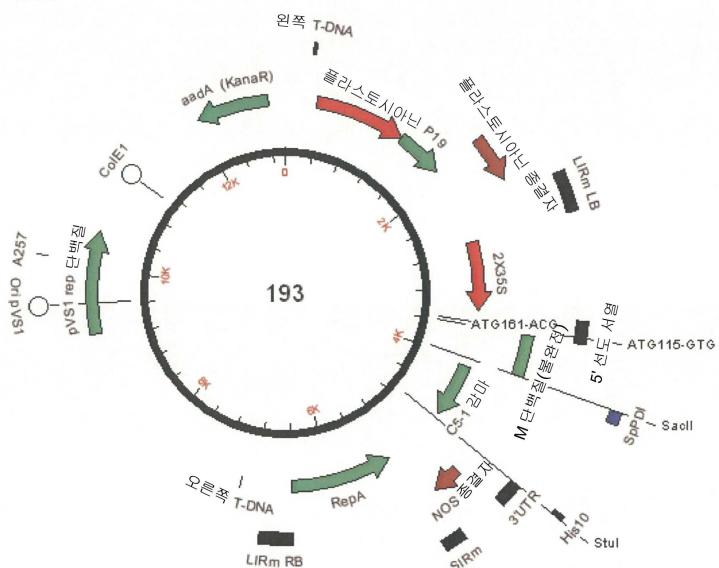
도면21



도면22a

B-2X35S/CPMV-HT/RVA(WA) VP2(opt)/NOS 대
BeYDV(m)+레플리카제 증폭 시스템(구조 번호 1711)

도 22A



도면22b

도 22B, SEQ ID NO: 26

원쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 구조 193 (밑줄 그어져 있음).
2X35S/CPMV-HT/NOS 내 플라스토시아닌-P19-플라스토시아닌 침목 억제자
발현 카세트가 있는 BeYDV(m)+레플리카제 증폭 시스템

```

TGGCAGGATATAATTGGTGTAAACAATTGACGCTTAGACAACCTAATAACACATTGGGACGTT
TTATGACTGAATTACCGCGAACCCGGCTGGTATATTATGTGCAATAACTCAAAAACC
ATAAAGTTAACTTAGCAAGTGTACATTTCATTGAAACAAAATTCACCTACTACTGTATA
AATCATATTAAACATTAGATAAGAAATATGGATGATAAGAACAGAGTAGTGATATTGACAA
CAATTGTTGCAACATTGGAGAAAATTTGTTCTCTTCTTGTCAAAAACAATAGAGAG
AGAAAAAGGAGGAGGAGAATAAAAACAATATGTGAGTAGAGAGAGAAAGTGTACAAAAGTTG
TACCAAATAGTGTACAAATATCATGGAGATTGACAAAAGCTACACAAATAAGGTTAATTG
TGTAATAAATAAGGATGACGATAGAGAGATGACATTAGAGAATTGGCAAGTCATTA
AGAAGAATAAATTATTTAAATAAAAGTGTGAGTCATTGATTAACATGTGATTATTAATGAA
TTGATGAAAGAGTGGATTAAAGTGTATTAGATAATTAGAATTGGTGTCAAAATTAAATTGACATT
GATCTTCTCTATATTGCCCATAGACTGCTTAACCTTATTTCTATAGATCAAATAAGA
GAATAACGGTATAATCCTCCAAAAAAACGGTATTACTAAAGTCTAAAGCTAAGCCACG
TAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCAAACCAATCACAAATCCTGATG
AGATAACCCACTTAAGGCCACGCCATCTGGCACATCTATTCTAAATCACACATTCTCCACA
CATCTGAGCCACACAAAACCAATCCACATCTTATCACCCATTCTATAAAAACACACTTTGTGAG
TCTCATTTGATTGGTTCAGGGATTACTATGAGCTTCAGCTTCTGGGTTTCGACCAGATCGGATGT
AGAAAATGGAACAGACTATACAAGGAAACGCGCTAGGGAAACAAGCTAACAGTGAACGTGGATG
GAGGATCAGGAGGATCACCCTCCCTCAAACTTCTGACGAAAGTCTCGAGTTGAGTGGCG
GCTACATAACGATGAGGAGAACATCTGAAGATAATCCCTTGTGTTCAAGGAAAGCTGGGTTTC
GGGAAAGTGTATTAGAGATACTCTAGAGATCACGAGCAGGAAAGCTTCACTGCACAGAGTCTG
GATCTTGACGGGAGATTGGTTAACTATGAGCTTCAGCTTCGATTTCTGGGTTTCGACCAGATCGGATGT
ACCTATAGTATTGGTTCAGGGAGTTAGTATCAGCTTCAGCTGGGGGTGCGGAACACTTCAGCATCT
CTGTGAGATGGCAATTGGTCTAAGGAAACTCTGCTACAGCTTGGGCAACTCGAAGTGGAAAGTAA
GTATCAAGAGGATGCCCTGAGGTACTCAACACCTCGAAAAGAAAGCGAGTAAGTAAATGCTT
CTTCGCTCTTATTATAATATGGTTATTGTTATTGTTAGAAGAGCTTAAATTAAATCG
TGTTGTTATTGAAATACTATTGTGAGATGACTGGTGTAAATGTTACAGTAAAGTGGAGT
CAGAATCAGAATTTCTCCATAACTAACTAGACATGAAGACCTGGCGCGTACAATTGTTATATT
TGAAACAATAAATGAACTATGCAACTTAAAGTGTAAATAGCTCAAATATATGTT
CAAGTCAATAGATTAAATGAAATATCAGTATCGAACTTAAACATCAACCTAACGTTATT
AACTACTAATTATATCATCCCTTGTAAATGATAGTACACCAATTAGGAAGGAGCATGCTGCC
TAGGAGATTGTCGTCGGCCCTTCACTGGTCAAGCTGCTAGCCGTAGGCAATACCGAACCG
CTCTCCCGCGCGTGGAAATTACTAGCGCTGTCAGACGCGTTGTTGTGACTCCGAGGGGTG
CCTCAAACCTATCTTATAACCGGGCTGGAGGCATGGAGGGTATTGGTCAATTAAAGT
AGTGGAAAATGGTGAATTAAAGACGAAGTCTTGCACAAAGGGGGGCCACCCGAA
TTAATTACCGGGCGTGGCCCCCTTATCGCGAGTTGCTTAGCACGAGCGGTCCAGATTAAAGTA
AAAATTCGGCCACTAGGGTAAAGGTGTCAACTATAAAAGCATATACGATGATGGTATT
TGTCGACAAGCTGCAATGGCAGTGGGAGGACACACTTGTCACTCCAAAGGAAACCTCTGG
AAAGATAACGATCTCAGAAGACCAAGGGCAATTGAGACTTCAACAAAGGGTAATATCCGAAAC
CTCTGGATTCCATGCCAGCTACTGTCATTGTAAGATGTTGAAAGAAGTGGAAAAGGAAGGTGGCT
CTACAAATGGCATCTGGGAAAGGGCATGGGAAAGGAGCATCTGGAAAGACGTTCCAGCTGGTCCC
AAAGATGGACCCCCACCCACGGAGGAGCATCTGGAAAAAGAACGTTCCAACACGTCTCAAAG
CAAGTGGATTGATGTGATAACATGGTGGAGCAGCACACACTTGTCACTCCAAAATATCAAAGATA
CAAGTCTCAGAAGACCAAGGGCAATTGAGACTTCAACAAAGGGTAATATCCGAAACCTCTGG
ATTCCATTGGCCAGCTCTGTCACTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTTACAAA
TGCCATATTGCGATAAAGGAAAGGCCATCTGGAAAGATGCTCTGGCGACAGTGGTCCAAAGATG
GACCCCAACCCACGGAGGAGCATCTGGAAAAAGAACGTTCCAAACACGTTCTCAAAGCAAGTGG
ATTGATGTGATATCTCACTGACGTAAGGGATGACGCAAACTCCACTATCTTCGCAAGACCTTCC
TCTATATAAGGAAAGTCTTCTTGTGAGGAGTAAATTAAATCTTAAAGGTTTGTAAAGCGAAC
GTGGGAAACCCGAACCAAACCTCTTCAAAACTCTCTCATCTCTTAAAGCAAAACTTCTCTT
GTCTTCTTGTGAGGATCTTCAACGTTGAGCTGCTTGGGACCCAGTACAACGTTCTTIC
ACTGAAGCGAAATCAAAGATCTTGTGAGCTGCTTGGGACCCGTTAAATAACGTTACTGT
CCTATTCTTGTGAGGATCTTGTGACTTCCGGGGTGCATAATCTCTACTCTGCTGACGAGGTATTGTTG
CTGTAATTCTCTCTCTCTGTGATGGTCTATAAGAAATCTAGTATTCTGTGAAACAGAG
TTTCCCGGTGGTTTCGAACCTGGAGAAAAGATTGTTAGCTTCTGTATATTCTGCCAAATTGTCG
GCCCCGGGATGGGAAAAGCGTGGGATTTCGGCTTAAAGGTTCTCTGTGTTGGTCTCTTC
AGATCTCGCTCGAGGCTCTCAGGAAAACGACACCCCATCTGCTATCCACTGGCCCTGGATC
TGCTGCCAAACTAACTCCATGGTGAACCTGGGATGCTGGTCAAGGGTATTCCCTGAGCCAGTG
ACAGTGAACCTGGAACTCTGGATCCCTGTCAGCGGTGACACCTCCAGCTGCTGAGTGA

```

도면22c

CCTCTACACTCTGAGCAGCTCAGTGACTGTCCCTCCAGCACCTGGCCCAGCGAGACCGTCACCTGC
 AACGTTGCCACCCGGCCAGCAGCACCAAGGTGACAAGAAAATTGTGCCCAAGGGATTGGTGTGTA
 AGCCTGCATATGTCAGCTCAGAAGTATCATCTGICATCTCCCCCAAAGGCCAAGGATGTG
 CTCACCATACTCTGACTCCTAAGGTACCGTGTGTTGAGACATCAGCAAGGATGATCCGAGG
 TCCAGITCAGCTGGTTGTAGATGATGAGGTGCACACAGCTCAGACGCACCCGGGAGGAGA
 GTTCAGACTCAGCACTTCCGCTCAGTCAGTGAACTCCCCATCATGCACCCAGGACTGGCTCAATGCCAAG
 GAGGATCGCTACCATCACCATCACCATCACCATAAAGGCTATTTCTGTCAGAGTGTGTTATT
 TTACTGTTATTGGTGTGCAATTCTATGTTGGTAGCGGTTCTGTCAGAGTGTGTTATT
 TGTAATTAAATTCTTGAGCTCTGAGCTCAGCAAGGACACAAAAAGATT
 TAATTAAATTAAAAAAAGACCGGGAAATCGATATAAGCTTACGACATGATTATCAGATCAGACATC
 GTTCAACACATTGCAAAAGTTCTAAGATTGAATCCTGTTGGCTTGTGATGATTATCATA
 TAATTCTGTTGAATTACGTTAACATGTAATGTCATGACGTTATTATGAGATG
 GGTTTATGATTAGACTCCGCAATTATACATTAAACCGGATAGAAAACAAAATAGCGGCG
 AACTAGGATAAATTATCGCGCGGTGTCATCTGTTACTAGATCTAGCTCAAGCTTGGCG
 GCCATAAAATGATTATTATGAAATTATTCATGTCAGGATAGAAATTACATATGTTACATA
 ACACACGAAATAACAAAAAAAGACAATCCAACAAACACCCAAAAAAATAATCACTTACA
 TAAACTCGTATGAGGAGAGGCACGTCACTGACTGACGATTCCGAGCAAAAGTCTCCCGC
 ACACATATAGGGTGCACCAATTATCTTAAGGTAATCCTCTGACTGTCATTGATAACATCC
 AGTCTCGTCAAGGATGAAATTATAGAAGGGATCCACCTTTATTCTCTTCTTCCATATT
 TAGGGTTGACAGTGAATCAGACTGGCAACCTTAAATTGCTTCCAATGGGAGCAACTGAGGG
 GATGTCGTCGATGATTATAGTGGCGTTCATGTCAGTTGGTGAATCTGATGGTACCGTGTCC
 AGTTGTTGTCGCGAGACTCTAGCCCAGGTCTTCCGGTAGCAGTTGTCAGATGAG
 GCTGGGGTGTGCGGATTCTCATGTCCTGGTTAAATCGCCATCCATICAAGGTCAGATTG
 AGCTTGTGGATGAGACAGGATGATGTCAGTTAGGTTACGTTAGCTTACATGGTATAGATGGTTT
 CCCCTCAGGAGTGTAGATCTCGTGGCAGCGAACATCTGATTGTAAGGGGCAACACATACGGTTC
 AGGGTGTGGAGGGAAATAATTGTTGGCTGAATATTCCAGCCATTGAGTTGTTGCCCATTGAG
 GGAATTCTCTCTGATCATGTCAGAATCTCTTACGCTGAGCTGGGATCTGGGAAATAGTTCTCC
 CTGCGTCAGATTGCGAGGAGACCTTATGATCTCGGAAATCTCTCTGGTTAAATATCTCC
 CTTGATATGTAATCAAGGACTGTTAGGTTAGCTGGCTGGATATTAGGGTGAATTCTCTCAA
 AATCGAAAAAAAGAAGGATCCCTAAACAGGTTTTTATCAAGCTGGAGAAAGAGCATGATAAGTGG
 TAGTGCCTATGATGAGCTCAGAACACCCAAGGAAGAAAATAAGGAGGTGTGAGTTCT
 CCCAGAGAAACTGGAATAATCATCTTGTGAGATGAGCTGGGATAGGTAAGGAAACATATT
 AGATGGAGCTGAGGTTCTACTAGCAGAACGGCATGTTGACTCCGAGGGGTTGCCCTAAC
 TCTATTTATAACCCGGTGGAGGCATGGAGGCAAGGGGTTATTGGTCATTAAATAGATAGTGGAA
 AATGACGTTGAATTACTTAAAGACGAAGTCTTGCAGAACGGGGGCCACGCCAATTAAAT
 TACCGGCGTGGCCCCCTTATCGCGAGTGCCTTAGCACGAGGGTCCAGATTAAAGTAGAAAATT
 TCCCGCCCACTAGGGTTAAAGGTGTTACACTATAAAAGCATATACGATGTGATGGTATTGACTAG
 TGCGACTGGCGTCGTTTACAACGCTGTGACTGGGAAACCCCTGGCGTACCCAACTTAAATGCC
 GCACACATCCCCCTTCGCGAGCTGGTAATAGCGAAAGGGCCGCACCGATGCCCTTCCAAAC
 AGTTGCGCAGCCTGAATGCCAATGCTAGAGCACCTGAGCTTGGATCAGATTGTCGTTTCCCGCTT
 CAGTTAAACTATCAGTGGTAGAGGATATTGGCGGTAACCTAAGGAAAGAGCGTTA

도면23a

도 23, SEQ ID NO: 27

2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1710.
로타바이러스 A WA 균주의 VP2(opt)는 밀줄 그어져 있다.

```

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAAATATCAAAGATACTGTCAGAACAGCC
AAAGGGCAATTGAGACTTTCAACAAAGGGTAATATCGGAAACCTCCTGGATTCATGCCAGC
TATCTGTCACTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCTACAAATGCCATATTGCGAT
AAAGGAAAGGATCGCTGGAAAGAGACGTTCCAACCACGTCTTCAAAGCAAGTGGAATTGATGTATAAC
ATGGTGGAGCACGACACACTTGTCTACTCCAAAATATCAAAGATACTGTCAGAACAGCCAAAGG
GCAATTGAGACTTTCAACAAAGGGTAATATCGGAAACCTCCTGGATTCATGCCAGCTATCTG
TCACTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCTACAAATGCCATATTGCGATAAAAGGA
AAGGCCATCGTGAAGATGCCCTGGCACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGC
ATCGTGGAAAAGGAAAGAGACGTTCCAACCACGTCTTCAAAGCAAGTGGAATTGATGTATACTCCACTG
ACGTAAGGGATGACGACAATCCCACATCTTGCAAGACCCCTCTTATATAAGGAAGTTCATTT
CATTTGGAGAGGTTAAAATCTTAAAGGTTGATAAAAGCGAACAGTGGGAAACCCGAACCAA
CCTCTTCAAACCTCTCTCATCTTAAAGCAAACCTCTCTTGTGCTTCTTGCAGGATCT
TCAACGTTGAGCTGAGCTGCGACCAGTACAACGTTTCTTACTGAGCGAAATCAAAGATC
TCTTGTGGACAGTAGTGGGGCCATTAAATACTGTTACTTGTCTTATCTGTGCGTGTGCT
TGGGAAAAGGAGCTGCTGGAGGCTGCTTACGCCCCATACATTGTTACGATTCTGCTGACT
TTGGGGGGTGCAATACTCTACTTGTGAGGAGTATTGCTGCTGACTTCTTCTTCTTC
TTGCTGATTGTTCTATAAGAAATCTAGTTTCTTGAAGGAGCTTCTTGTGAGGAGTCT
GGAGAAAAGATGTTAAGCTCTGTATATTCTGCCAATTGTCGGGGCATGGCATACCGGAAAGAG
AGGAGCAAAAGCGCAGGAAACCTGCGCAACAGAACAGAGACTGCAAGAAAAGAGATAGAGAAAAG
ATGCGAGCTAACATGAAAACAAGAAATAACATAAGGAAACAACAGCTGCGACAAAGTCTGT
CCCAAGGAGGAAATTACTGACGCCAGGAGCATTTAAAATTGCGGAGAAATAAGAGA
GCTGAAAGAAGAAATCTAACAGCTGCGAAATTCTGAAAAGAACAGGACATCGAGAAAGAGA
TTCAATGAAATTGCAAAACATACTTACATTGAGTCCAAAGAAAGATATCTCTAAAGAAGT
TGAAGACATAAGACGGAGCAGGCAAAAACAGATGAAACTCTTGTGATTTCAGGCAAACAA
GCTCCCTATATCGGCCAATGGCGAGAGGAGACTACGCAACCGGTGGTACTGGAAGTGGAAA
AGACACCTGCGAGATGGAGATTGACGTCGGGAGTATTCTCAATCTATGATGATCTCTA
TCGAATGCCGAACTATGCTCTCAAGGACATGGCGTGGAGAACAAAATAGCAGAGACGCC
GCAAGTGTGCGACTCTGAGACTGCCAATTTGATGCGCATCTTCCAGGATGAGGAGACCGAGGG
AGTCGCGTAGATTCGCTGATAGCGCAACAGTGGCTGAGCTGATCGAACATGTCATATAC
CCTCCATCTCTCCACTTATGATCATGCAATTGAGTTTCTTAAACACCAGTTGGTGGAGCC
GCTGAACAATGAGATACTTCAATTACAGGAGAGGATAAGGAATGACGTGAAATTACCTG
AACATGGATATGAATCTGCCATCAACGGCTGATATCAGGCCAAACTGTGCGAGGATAGCTGA
ATCTTCACGATAATTGAGTCCTGTGGGATACCATCACAACATCCAACATACATCTGCCAGGTG
GTCGTTCCGAGTTGAGGAGAGGAGCTGGTCTCCAGGAAGCACAGATCCGAGAAAATGAGCAG
GACCTGCGAGCTGGGGCCTCACATTGAGCGAGACACAGTTTGTGCGGAGTAAAGCTCAGG
CTGCAATGATTGTTCAAGACCCCTCATAGCGCCATGCTGCTCAAAGAACATGCTTGGACTT
GTGACCCAGAACTATAGGCCATACTCCGGAAATGTCGACTTACAGTGGATTCACCGATATGTT
CCTCCGGAGGACTACTGTCGCTGTGAGCTGGCGATCATCACACCCATGTGATTCAGCATCGGA
ATGCAAGAGAATGCAATTCCGGAATGGCGACCCCTAGACACCCCTCCAGATCGCAGAACAGCAGATCC
AGAAATTGCGGATGGCAAGGGAGCTGGTCTCCAGGAAGCACAGATCCGAGAAAATGAGCAG
GACCTGCGAGCTGGGGCCTCACATTGAGCGAGACACAGTTTGTGCGGAGTAAAGCTGAG
CTGCAATGATTGTTCAAGACCCCTCATAGCGCCATGCTGCTCAAAGAACATGCTTGGACTT
GTGACCCAGAACTATAGGCCATACTCCGGAAATGTCGACTTACAGTGGATTCACCGATATGAA
CCTCCGGAGGACTACTGTCGCTGTGAGCTGGCGATCATCACACCCATGTGATTCAGCATCGGA
ATGCAAGAGAATGCAATTCCGGAATGGCGACCCCTAGACACCCCTCCAGATCGCAGAACAGCAGATCC
AGAAATTGCGGATGGCAAGGGAGCTGGTCTCCAGGAAGCACAGATCCGAGAAAATGAGTGA
TGGAGTTCTGAATCAGACTCTGAGCAATAACGGGAATGGAGCGGATCATCAACCCAGTGTGAG
GCATGATGCAACTCAGCAGACAGCAGTTCCCGACGATGCGCTGTGAGTACAAACGGAGCATCCAAC
GGGGCATTCGCTCTGCTTACCGCTGAGCTGGGAGACGGCTTGCACATTTAACCTTGTGATGAA
TACAGAGCGCTAATGGCTTGATGACCATGCAACATGCAACAGTGTGAAACCTGACAATGAGAAGT
TGCAGCTACTTGTGACTTGTCTTGTGTTAATGGTAACACAACCGTGTGCTTCCGCTCCACAG
ACACTGTTGCAACTTCAACGCGTAACTTCAACGCAATTCTCAACGAAATGAGCGGATCAACGAG
CCGTCGCCATAATTACCGCAGAAATTAGGCTGAATTCTTATCAGAAAAAAATGAAGTCCATAGTGA
AGACTTCTGAACGGCTCCAGATTGAGCTGACCGTACCCAGTGGCTGACGACCAAATGTCAGCGT
AGGGATCGCTCTGCTTACCGCTGAGCTGGGAGACGGCTTGCACATTTAACCTTGTGATGAA
ATATGGAGCAGATGCAACCGCTTGTGAGTGGGTTTACATCCGATACCGAGATAT
GCAGCTGGAACCGCAGGAGATGACGGATATGTTAATGCAAGGAATCTGAGTGGCTACCGCAA
ATTAACCTGGAGGAACACTGCGCAGGGCTGATTACGGAACATGCTTCTCAACATC
AACCCGGTGGCCCTTGTGGGTGCTTGTGCGACTCATCCGTGATCGTCAATCGCCAAG
CTCGAGCAACCGCTCTGCGTCAAGATGTAAGCTGAGCGAAAGTGTGACACACTGAGCCCATACTGT
ACAAATAAAACTCGGACTTCAACGACTTACCTTACCTGACTTACAGGCAACATGCTGACTTCT
AACTAAGGTACAAACAGGTCGCCACACATTGCACTTACGACTTACGACTGAGTCCACAAAGTAC
AACCTTACGTTACCGCTACTCTGACCTACTGTTAGTGTGAGCGGACACGGTAGAGCCCATAA
CGCAGTCGCAATTGCAAAATATGCAAAATGAGCGAGCTTAAAGGCTTATTTCTTGTGAGTGGTGAATT
TACTGTTATTGGTGTGCAATTGATGTTGGTGAAGCGGTTTCTGTCAGAGTGTGTTTATTTAT

```

도면23b

```

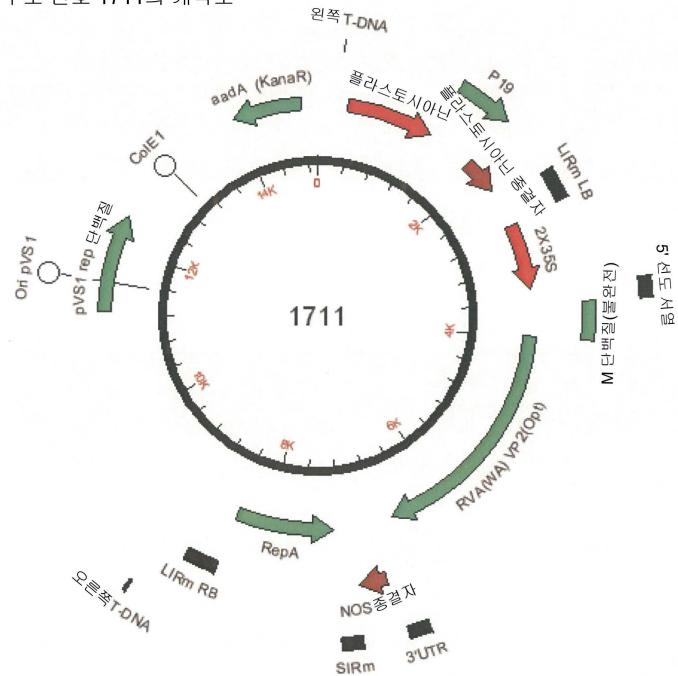
GTAATTTAATTCTTGTGAGCTCTGTTAGCAGGTGTCCTTCAAGCAAGGACACAAAAAGATTT
AATTTTAATTTAAAAAAAGACGGGAATTGATATCAAGCTTATGACCTGAGTC
GTTCAACACATTGGCAATAAGTTCTTAAGGTTAATCTGCTGTCGGCTTGGATGATTATCATA
TAATTCTGTGAAATTACGTTAAGCATGTAATAATTAACATGTAATGCGATGACGTTATTGAGATG
GGTTTATGATTAGAGTCCGCAATTACATTAATACGCGATAGAAAACAAATAGCGCGCA
AACTAGGATAAATTACGCGCGGTGTCATATGTTGAGCGGTTTCTGTCAGAGTGTGTTTATTTAT

```

도면24

도 24

구조 번호 1711의 개략도



도면25aC-2X35S/CPMV-HT/RVA(WA) VP6(opt)/NOS (구조 번호 1713)

도 25a, SEQ ID NO: 28

IF-WA_VP6(opt).s1+3c

AAATTGTCGGGCCATGGAGGTCTTATAGTCTCTCCAAAACGCTGA

도 25b, SEQ ID NO: 29

IF-WA_VP6(opt).s1-4r

ACTAAAGAAAATAGGCCTACTTGATCAACATACTCCGGATAGAGGCCA
CA

도 25c, SEQ ID NO: 30

2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1713.

로타바이러스 A WA 균주의 VP6(opt)는 밀줄 그어져 있다.

GTCAACATGGAGCACACTTGTCTACTCCAAAATATCAAAGATACTGTCAGAACACC
AAAGGCAATTGAGACTTTCAACAAAGGGTAATATCGGAAACCTCCTCGATTCCATGCCAGC
TATCTGTCACTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCTACAAATGCCATATTGCGAT
AAAGGAAAGGCCATCGTGAAGATAGTGGAAAAGGAAGGTGGCTCTACAAATGCCATATTGCGAT
AGGAGCATCGTGGAAAAGAACGCTTCAACCCACGCTTCAAGCAAGTGGATTGATGTATAAC
ATGGTGGAGCACGACACACTTGTCTACTCCAAAATATCAAAGATACTGTCAGAACGACAAAGG
GCAATTGAGACTTTCAACAAAGGGTAATATCGGAAACCTCCTGGATTCCATGCCAGCTATCTG
TCACCTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCTACAAATGCCATATTGCGATAAGGA
AAGGCCATCGTGAAGATGCCCTGCGCAGCTGGTCCCAAAGATGGACCCCCACCCAGGAGGAGC
ATCTGGGAAAAGAACGCTTCAACCCAGCTTCAAGCAAGTGGATTGATGTATACTCCACTG
ACGTAAGGGATGACGCCACAATCCCACTATCTTCCAAGACCTCTCTCTATAAAAGGAAGTTATT
CATTTGGAGAGGTATTAAATCTTAAATAGGTTTGTATAAAAGCGAACGTTGGGGAAACCGAACCAA
CCTCTTAAACTCTCTCATCTCTTAAAGCAAAACTCTCTTGTCTTCTTGTGAGGGATCT
TCACCGTGTGAGATCGTGTGCGCACCGTACAACCGTTCTTCACTGAAGCGAACATCAAAGATC
TCTTGTGGACACGTAGTGTGCGGCCATTAAATAACGTGTACTTGTCTTATTGTGCGTGTGGTCT
TGGGAAAAGAACGCTTCAAGGAGCTGAGGGACAAGATCGTGGAGGTACACTTATAGCAATGTCA
ACCTAATACAGCAGTTAAATCAAATGATCGTACAATGAATGGGAATGATTCCAACTGGCGTAT
TGGTAATCTGCCCCTGAGGAACCTGGACATTTCGATTTCGCTGGGACGACTCTCTTAAACTCTG
ATGCAAATTATGTAGAAAACGCCAACGATTATCGAGACTTATCGATTTCATTGATAACGTTGT
ATGAGATGAGATGGCCCGCGAGTCACAACCGAACGGAGTTGCTCCTCACAGTCCGAGGCCCTCGAAA
CTCGCCGGCATTAAGTCAGCGTATTAAATTGACAACCTCCCTCCGAATATAAGAGAACGAACTT
GCAGAACATCGTGCAGAGAACCGGCTTCCGTTGAGGCTTAAACCTAATATCTTCCGTATAGCGCTCAT
TCACCTGAATAGAGTCAGCCCATGTCAGCACACCTCATGGTACAATGTTGCGTAAATGCCGAG
TGAATACAGTCGCCGGTTGCAATTTCGCTGCAATTAAATGCAACCGCAACATCAGCAGTC
GAACATATCGTCAACTAACGCGCTCTCACGACCGCAGAACATTACACTCTGCCGACGCCGAGC
GCTCTCTTCCCGCGTAACTCAACTCAGCTGATGCCGACCACTTGGTTCTCAACCTGTTATAT
TGCCTCCTAACAAACGTTAGAGGTGAGTTCTCTTAAACGGACAGATCATCAATACTTACCAAGCCAG
GTTGGCAGCAGTATTGCAAGGAAGAACCGCTGAGGCTGCTTCCAGCAGCACCGCACCGTCGG
TTGACTCTCGAATAGAGAGCGCCGCTCGGAATCAGTGTGGCAGACGCCAACGAGACGCTGCTGG
CAAACGTTACCGCCGTCGGCAAGAGTATGCCATCCCAGTAGGGCTGTGTTCCACCGGCCATGAA
CTGGAATGAACTAAACTAACTATAGCCCATCCAGAGAACGACAACCTGAGCGGGTCTTCACTGT
GCCCTATCCGGAGTATGTGATCAAGTAGAGGGCTTATTTCTTGTGTTGAATTACTGTTATCCGGT
GTGCAATTCTATGTTGGTGAAGCGTTCTGTGCTCAGAGTGTGTTATTTATGTAATTAAATTCT
TTGAGCTCTTGTGAGCAGGTCGCTCTCAGCAAGGACACAAAAGATTAAATTAAATTTAAAAA
AAAAAAAGAACCGGAAATCGATATCAAGCTTATCGACCTGAGATCGTCAACACATTGGC
AATAAAGTTCTTAAGGATGAATCTGTGCGGCTCTGCGATGATTATCATATAATTCTGTTGAATT
ACGTTAAGCATGTAATAATTAAACATGTAATGCACTGAGCTTATTTATGAGATGGGTTTATGATTAGA
GTCCCGCAATTATACATTAAACCGCAGTAGAAAACAAAATAGCGCAGCAACTAGGATAAATTAT
CGCCGCGGGTGTATCTATGTTACTAGAT

도면25b

도 25d

로타바이러스 A WA 군주의 뉴클레오티드 서열 암호화 VP6(opt) (SEQ ID NO: 46)

```
ATGGAGGTCCTTATAGTCTCCAAAACGCTGAAGGACGCTAGGGACAAGATCGTG  
GAGGGTACACTTATAGCAATGTCAGCAGCTAATACAGCAGTTAACAAATGATC  
GTTACAATGAATGGGAATGATTCCAACCTGGCGGTATTGTAATCTGCCGTGAGG  
AACTGGACATTCGATTCGGCTGCTGGCACGACTCTCCTAACCTCGATGCAAAT  
TATGAGAAAACGCCAGAACGATTATCGAGTACTTTATCGATTTCATTGATAACGTT  
TGTATGGATGAGATGGCCCGAGTCACAACCGGAACGGAGTTGCTCACAGTCCGAG  
GCCCTGGAAACTCGCCGGATTAAGTTCAAGCGTATTAAATTCGACAACCTCC  
GAATATATAGAGAACTGGAACITGCAGAACGAGAACCGGCTTCGTGTT  
CATAAACCTAATATCTTCGCTATAGCGCCTCATTCAACCTGAATAGGAGTCAGCCC  
ATGCACGACAACCTCATGGGATCAATGTGGCTGAATGCCGGAGTGAACACAGTC  
GCCGGGTCGATTAACCTCTGTGCCATTATGCACCCGAAACATCCAGCAGTTGCAA  
CATATCGTCAACTAACGACGGCTCTCACGACCGCGAACATTACACTCCTGCCGAC  
GCCGAGCGCTCTCCTTCCCCCGTAATCAACTCAGCTGATGGCGCCACCACTTGG  
TTCTCTCAACCCCTGTTATATTCCGCCCTAACACGCTAGAGCTGGAGTTCTCTTAAAC  
GGACAGATCATCAATACCTACCAACGCTAGAGCTGGCAGATTATTGCAAGAAATTTC  
GACGCTATCAGGCTGCTCTCCAACTGATGAGGCCCCCAATATGACTCCCGCTGTG  
AACGCTTGTTCGCAGGCTCAGCCTTCCAGCACACGCCACCGCTGGCTTGA  
CTTCGAATAGAGAGCGCGGTCTCGGAATCAGTGTGGCAGACGCCAACGAGACGCTG  
CTGGCAACGTTACCGCGTGCAGAACAGAGTATGCCATCCAGTAGGGCTGTGTT  
CCACCCGGCATGAACGGACTGAACATAACTAATAGCCATCCAGAGAAAGAC  
AACTGCAGCGGGCTTCACTGTGGCTCTATCCGGAGTATGTTGATCAAGTAG
```

도면26

도 26, SEQ ID NO: 31

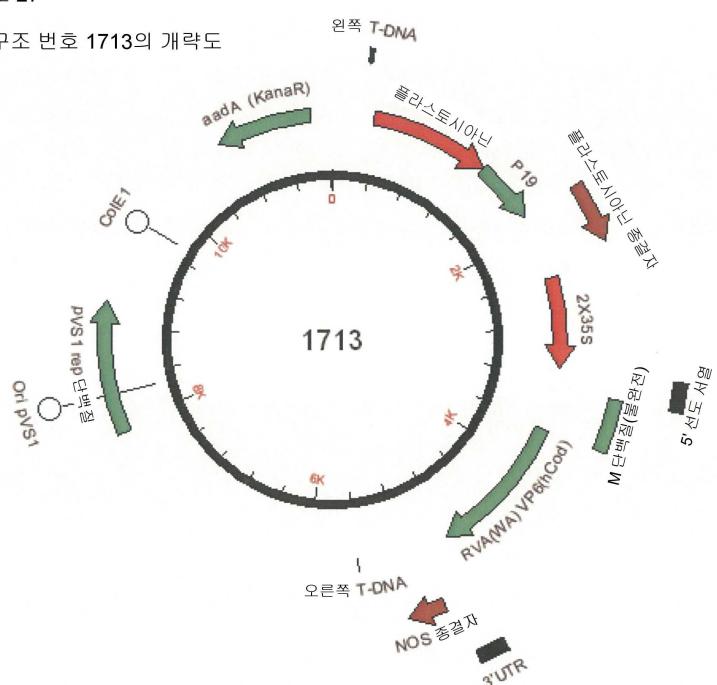
로타바이러스 A WA 군주의 VP6의 아미노산 서열

```
MEVLYSLSKTLKDARDKIVEGTLYSNVSDLIQQFNQIVTMNGNDQTGGIGNLPVRNWTDFGLLGTTL  
I.NLDANYVENARTIEYFIDFDVNVCMDMARESQRNGVAPQSEALRKLAGIKFRNFDNSSEYIENWNL  
QNRRQRTGFVFHKPNIFPYSASFTLNRSQPMHDNLMTWLNAGSEIQVAGFDYSCAINAPANIQQFEHI  
VQLRRLTTATILLPDAERFSFPRVINSADGATTWFNPVILRPNNVEFLLNGQINTYQARFGTIARN  
FDAIRLLFQLMRPPNMTPAVNALFPQAOPFQHATVGLTLRIESAVCESVLADANETLLANVTAVRQEYA  
IPVGPFVPPGMNWTELITNYSPSREDNLQRVFTVASIRSMLIK
```

도면27

도 27

구조 번호 1713의 개략도



도면28

D-2X35S/CPMV-HT/RVA(WA) VP6(opt)/NOS 내 Beydv(m)+레플리카제 증폭 시스템 (구조 번호 1714)

도 28, SEQ ID NO: 32

2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1714.

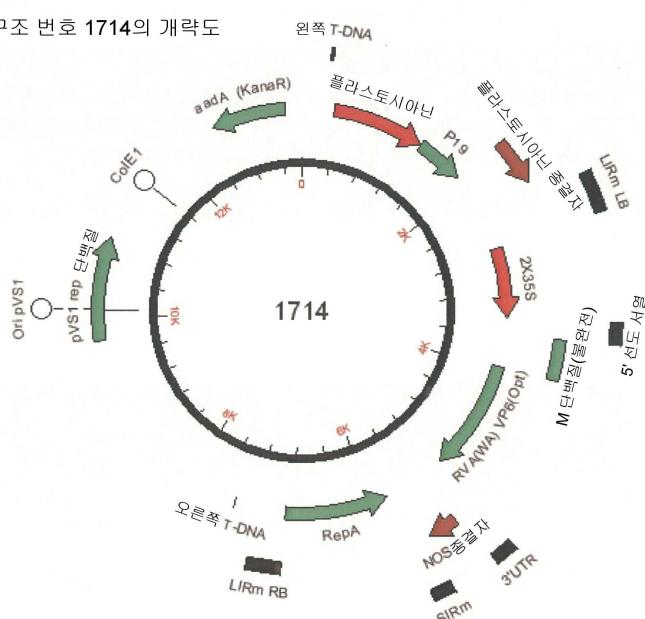
로타바이러스 A WA 균주의 VP6(opt)는 밀줄 그어져 있다.

GTCAACATGGTGGAGCACGACACACTTGCTACTCCAAAAATATCAAAGATACTGTCAGAACGACC
AAAGGGCATGGAGCTTTCAACAAAGGTAATATCCGAAACCTCTCGGATTCATGCCAGC
TATCTGTCACCTATTGTAAGAGATGCCTGCGACAGTGGTCCAAGAGATGGACCCCCACCCACG
AAAGGAAAGGCCATCGTTGAAGATGCCTGCGACAGTGGTCCAAGAGATGGACCCCCACCCACG
AGGGCATGTGGAAAAGGCCATCGTCAACAAAGGTAATATCCGAAACCTCTCGGATTCATGCCAGC
ATGGTAGGAGACGACACACTTGCTACTCCAAAATCAAAGATACTGTCAGAACGACC
GCAATTGAGACTITCAACAAGGTAATATCCGAAACCTCTCGGATTCATGCCAGC
TCACCTTATGTGAAGATGTGGAGGTGGCTCTACAAATGCCATATTGCGATAAAAGGA
AAGGCCATCGTTGAAGATGCCATCGCGACAGTGGTCCAAGATGGACCCCCACCCACGAGGAGC
ATCGTGGAAAAGAACGCTTCAACACAGCTCTCAAAGCAAGTGGATTGATGTATATCTCAGT
ACGTAAGGGATGGCAGCACAATCCACTCTCGCAAGGACCTCTCTATAAAGGAAGGTTCATT
CATTTGGAGAGGTATTAAAATCTTAAGGTTTGATAAAAGCGAACAGTGGGAAACCCGAACCAA
CCTCTCTAAACTCTCTCATCTCTTAAAGCAAACCTCTCTTGTCTTCTGCGTAGCGATCT
TCAACGTTGAGCTGAGCTGCTTGCACAGTACACGTTTCACTGAGCGAAATCAAAGATC
TCTTGTGGACAGTACTGCGGCCATTAAATAACGTGTACTTGTCTATTCTTGTGGTGTGGTCT
TGGGAAAAGAACGCTGGAGGTGCTTCAAGCCCACATACATTACTGTACGATCTGCTGACT
TTCGCGGGTAACTCTACTCTGCTIGACGAGGATATGTCGCTGACTCTTCTCTTC
TTGCTGATTGGTCTATAAGAACATCTAGTATTTCGAAACAGAGTTCCCGTGGTTTCGAACCT
GGAGAAAGATTGTTAAGCTCTGTATTCGAAACAGGTTCTCGCAAGGAGGTCTTATAG
CTCTCCAAACCGCTGAAGGACGCTAGGGACAAGATCGTGGAGGGTACACTTATAGCAATGTCAGCG
ACCTAACACAGCTTAACTAACATGATGTTACATGAATGGGAATTTCCAACACTGGCGTAT
TGTAAATCTGCCGTGAGGAACCTGACATTCTGGCTGCTGGCACACTCTCTTAATCTCG
ATGCAAATTATGTAAGAAAAGCGAACAGATTATCGAGTACTTATCGATTTCATTGATAACGTTGT
ATGGATGAGATGGCCGAGTCACACCGAACGGAGTTGCTCCAGCTCGAGGCGGCCCTCGGAAA
CTCGCCGCAATTAGCTCAAGCTTAATTCTGCAACACTCTCCGAATATAGAGAACGACTGGAACCT
GCAGAACATCGTCAGAGAACCGCCTCTGTTGCTTCCATAACCTAATATCTTCCGTTAGCGCCTCAT
TCACCTGAAATAGGAGTCGGCCATTCTACGACACCACTCATGGTACAATGTCGAATGCGGGAG
TGAACATACAGGTGCCGGTTCTGATTACTCTGTGCCATTAAATGCAACCGCAGAACATCAGCTTC
GAACCATATCTGCAACTAAAGCAGGCTCTACGACAGCCGACAAATACACTCTGCCGACGCCGAGC
GCTTCCGCTAACACGTTAGAGGTGAGTTCTCTAAACGGCAGAGTATCATCAACACTACCAAGCCAG
GTTCCGACGATTATGCAAGAAATTTCGACGCTACAGGCTGCTCTCCAACTGATGAGGCCCCCA
ATATGACTCTCCGCTGTAACGGCTTGTGCTCCGAGGCTCAGCTTCCAGCACCACGCCACCGTCCG
TTGACTCTTCGCAATAGAGAGCGGGTCTGCAATCTGCTGAGCTTCCACCCGGCATGAA
CAAACGTTACCGCCGTGGCGCAAGAGTATGCCATCCAGTAGGGCCTGTTCCACCCGGCATGAA
CTGGACTGAACTAATTACTAACTATAGCCCACCCAGAGAACACTTGCAGCGGGTCTTCACTGTG
GCCCTCATCCGGAGTATGTCGATCAAGTAGAGGCTATTTCAGTTAGGTTGAATTACTGTTACCGT
GTGCAATTCTATGTTGGTAGCGGTTCTGTCAGAGTGTGTTTATTTATGTAATTAAATTCT
TGTGAGCTCTGTTAGCAGGTCCTCTCAGCAAGGGACACAAAAGATTAAATTAAATTAAAAA
AAAAAAAAAAAAGACCGGGAAATTGATATCAAGCTTATCGACTCGAGATCGTTAAACATTG
AATAAAAGTTCTAAGATTGAATCTGTTGCCGTTGCGATGATTATCATATAATTCTGTAATT
ACGTTAACATGATAATTACATGTAATGTCAGCTTATGAGATGGTTTATGATTAGA
GTCGCCAATTATACATTAATACCGCATAGAAAACAAAATAGCGCAGAACACTAGGATAAATTAT
CGCGCGGGTGTATCTATGTTACTAGAT

도면29

도 29

구조 번호 1714의 개략도



도면30

E-2X35S/CPMV-HT/ RVA(Rtx) VP4(opt)/ NOS (구조 번호 1731)

도 30A, SEQ ID NO: 33

IF-Rtx VP4(opt).s1+3c

AAATTGTCGGGCCATGGCTAGCCTGATCTACAGACAACCTTGACCAAT
TC

도 30B, SEQ ID NO: 34

IF-Rtx_VP4(opt).s1-4r

ACTAAAGAAAATAGGCCTTCAGAGTTACATTGCAAGGATTAATTGCTCAA
TCCTA

도면31a

도 31A, SEQ ID NO: 35

2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1731.

로타바이러스 A Rotarix 균주의 VP4(opt)는 밑줄 그어져 있다.

도면31b

CATTTCGGCCGCGCTAAACCAAAATCGATATGCTACTCGAGTCGGCAAGAACACTCTGCC
GATATCGAACCGAACGCCCTCGAAAAGTTATCCTTAAGGGCAGCTACGAATATTGAAAGATGACG
AGGTACGGAGATCACACAGAAGGGAAGTCTTCGCTTATAAGTACAACCCCTTGAGGAGTTCC
GTTTCAGCTAACATAAGTTGCAAGCTCGAGCATGTCAGTGTCTGCCTCATGACTTCA
AGACTTTGAAAGAACCTGAAAGACAACATATGGAATAAACCGGACCGAAGCGTTGAAACCTCATAAGTC
CAATCCCAAATGTCGCAATTCTAAACAGGAACTTCAACAACTGAAAGATAATTAGGTGGACAA
TTAATCTGCAACTGAAACTCTGAAGGCCATTCTTTCTTAACTTACTGTTAATTCGTTGCT
TCTCATGTTGGTGAGCGGTTTCTGCTCAGAGTGTGTTTATTTATGTAATTAACTTCTTGTGA
GCTCTCTGTTAGCAGGGTAATCCCTCAAGGACACAAAAGTTTAAATTAAAAAAA
AAAAAAAAGGGCTATCGATATCAACGTTATCGACCTGCGAGATCGTCAAACATTGGCAATAA
AGTTCTTAAGATGTAATCTGTGCGGTCTGCGATGATTATCATATAATTCTGTTGAAATCGTT
AAGCATGTAATAATTAAACATGTAATGCACTGACGTTATTATGAGATGGTTTTATGATTAGAGTCCC
GCAATTATACATTTAAACCGCATAGAAAACAAATATAGCGCGCAAACCTAGGATAAATTATCGCGC
GGCGTCTCATCTATGTTACTAGAT

도면31c

도 31B

균주 RVA/Vaccine/USA/Rotarix-A41CB052A/1998/G1P1A[8]의
로타바이러스 A VP4의 최적화된 임호화 서열(SEQ ID NO: 47)

```

ATGGCTAGCCTGATCTACAGACAACCTTGCACCAATTCTATTCGTGGATCTCATGACGAAATCGA
GCAGATTTGGTCCGAGAACGCCAGAACGTCGACCATCAACCCCTGGACCTTTGCTCAGACCCGCTAT
GCCCTGTAATGGGATCACGGAGAACATCAACCGACAGTACGACCGTCGAACCCATTCTGGACGGG
CATACCAACCCACCCACCTCACCCACCTAATGATTATTGGATTAAATCAACTCCAACACAAAACGGA
GTGGCTACGAGTCCAATAACTCCGATTTGGACCGCCGTTGAGCCATCGAGCCACACGTCAA
TCCTGTCGATCGCCAGTATAATGATATTCGGCAGTCAAACAGTTAACGTTAACGTTCAATGACAGCAAC
AAATGGAAGTTCTGGAGATGTTCGCAGCTCTCAGAACGAAATTCTATAATAGACGGACCCCTA
CTTCGATACAGACTCGGGTATTTAAAGTACGGCAGGGGTGAGGACATTTCACGGTGAAC
CTTCGAGCAACCCACTGACTCCAGTAGCCTGCAACAACTGAACAAATATCTATTACCATCCACAGC
GAATTCACATAATCCAAGATCTCAGGAAAGTAAGTGTAAACGAATATATCAACAAACGGACTCCCC
CAATCAGAAACACCGAACGTCAGTCTCCCAACTCAGTTCTCGTCTATCCAGTATAAGAGAGC
ACAAGTGAATGAGGACATTATTGTGAGCAAGACTAGCCTTGAAAGAAATGCACTACAAACAGAGA
CATTATCATCGGTTAAGTTGGAACTCTATCGTAAGATGGCGGCCTGGGGTACAAATGGTCA
GAATCTCATAAAGCGGCAACACTACAGTAAACTACTTGAGAGACGGCGACAGGTAAACCGCCC
ACACAAACATGCTGTCACGGGTTAAATAACTTTAGCTACACGGAGGCTTCTCCACCGAACCTC
GGTATCAGCCGGTATGAAGTCATCAAGGAAAATTCTATGTTGAGCTAGATTACTGGGATGATAGCA
AAGCGTAACTACAGATCTGGGAGGCTGGCTGCTAACTCAATTCAGTTGAAAGTGTACTGGT
GGATCATATTATTCATCCGCTGGGGCTTGGCAGTCAGTAAATGGGGGAGTCTCCCTCA
TTTGCTGGCGTGAAGTGTGAGCACTCAGTTACCGATTCTGTGTCCTGAGCTCCGAGGTTGGGTT
TCCCTTACTGTCGAGGAGCCCCATTCAAGCATTCTGCGTACAAGAACACTGTCACCTCTACGGGTTAC
CTGCGCGGAATCAAACACGGCAATGAATACTATGAAATTCCGGCCGTTCTTTGATAAGTCT
GGTACCAACTAATGACGACTATCAGACACCCATCATGAACAGCGTGAECTGTCAGACAGGACCTGGA
AAGACAACCTACAGATCTGGGAGGAAATTCAACTCTCAGTCAGGAGATTGCAATGGCCAATTG
ATAGATCTGCCCTACTGCCCTCGATATTTAGTATGTTCTCCGGCATCAAATCAACTATAGATCT
GACAAGAGCATGCTACTCTGTGATGAAGAAGTCAAGGAAATCAAACATTGCCACGAGCATATCA
GAATGACGAACTCTGAGTGTGAGCATTCAGCGTCAACCGTTCCATTGGTCAATC
TCAGCGCCATCAGCAACTGGACAAACGTGTCCAACGGACGTCAAGTGACCAACTCCITGAACGA
TATTCTACCCAGACGTCAACGATCAGTAAGAAACTCCGTTGAAAGAAATGATCACCCAGACTGAG
GGAATGTCCTTGACGACATTCCGCCGCGTGTAAAACCAAATCGATATGTCACTCAGATCG
GCAAGAACACTGCCGATATCGTAACCGAACGCTCCGAAAAGTTATCCCTAACGCGAGCTACAG
AATATTGAAAGATGACGAGGTATGGAGATCAAACACAGAACGGAGTTCTCGCTTATAAGATCAA
CACCTTGACGAGGTTCCGTTGACGTCAATAAGTTGCAAGAGCTGACAGATAGTCCAGTGATT
CTGCCATCATTGACTTTAAGACTTGTGAAAGAACCTGACGACAACATGGATAAACCGGACCGAAGC
GTGAAACCTCATTAAGTCCAATCCAATATGTTGCCAATTCTATTAACCAGAACATCCAATCTAA
GAAATAGGATTGAGCAATTAACTCGCAATGTAACCTG

```

도면31d

도 31C (SEQ ID NO: 44)

2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1730.
로타바이러스 A Rotarix 균주의 VP4(opt)는 밑줄 그어져 있다.

```

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAAATCAAAGATACTGCAGAACGACC
AAAGGCAATGGAGACTTTCAACAAAGGTAATATCGGAAACCTCTCGGATTCATGCCAGC
TATCTGTCACTTTATGTGAAGATAGTGGAAAAGGAAGGTGGCTCTACAAATGCCATATTGGCAT
AAAGGAAAGGCCATCTGGAGATGCTCTGCCGACAGTGCTCCAAAGATGGACCCCCACCCACG
AGGAGCATCTGGAAAAAGAACGCTTCAACCACGCTTCAAGCAAGTGGATTGATGTATAAC
ATGTGGAGCACGACACACTTGTCTACTCCAAAATCAAAGATACTGCAGAACGACCAAAAGG
GCAATTGAGACTTTCAACAAAGGTAATATCGGAAACCTCTCGGATTCATGCCAGCTATCTG
TCACITATTGGAAGTAGTGGAAAAGGAAGGTGGCTCTACAAATGCCATATTGCGATAAAGGA
AAGGCCATCTGGAGATGCTCTGCCGACAGTGCTCCAAAGATGGACCCCCACCCACGAGGAGC
ATCGTGGAAAAGAACGAGCTCTGCCAACCGCTTCAAGCAAGTGGATTGATGTATACTCCACTG
ACGTAAGGGATGACGCCAACATCCCATCTCGCAAGACCCCTCTCTATATAAGGAAGTTCATT
CATTTGGAGAGGTATAAAATCTTAAGGTTTGATAAAAGCGAACGTTGGGGAAACCCGAACCAA
CCTCTCTTAACTCTCTCATCTCTTAAAGCGAACCTCTCTTGCTTCTTCTTGCGTGAAGCGATCT
TCACCGTGTGAGCTGCTGCCGACAGTACAACGTTCTTCACTGAAGCGAACATCAAAGATC
TCTTGTGGACAGTAGTGGCGCCATTAAACGTTGACTTGTCTTACCTTGCGTGTGCT
TGGAAAAGAACGCTGCTGGAGCTGCTGTCAGCCCCATACATAACTTGTACGATCTGCTGACT
TTGCGGGTGGCAATCTACTCTGCTGACGAGGTATTGTTGCTTACTCTTCTTCTCTC
TTGCTGATTGTTCTATAAGAAACTCTAGTTTCTTGTGAAACAGAGTTTCCCGTGGTTTCTGAACTT
GGAGAAGATGTTAGCTGCTGTTGACGCTTCCAAATTGCTGGCCCATGGCTAGCGCTGACTAC
AGACAACTCTGACCAATCTCATTTCTGGAATCTCATGACGAAATCGAGCAGATTGGTCCGAGA
AGACCCAGAACGTAACGATCAACCCGGACCTTGTCTAGACCCGCTATGCCCTGTGAATIGGA
TCACGGAGAACATCAACGAGTACGACGGCTGAACCCATTCTGACGGGGCATACCAACCCACACC
TTACCCACCTAATGATTGATTAACTCAACACAAACGGAGTGGCTACGACTCCAC
TAATAACTCCGATTGGACCGCCGTTGAGCATCGAGCCACAGCTCAATCTGCGATGCCAGT
ATATGATACTGGAGGAGTAACTGAGTAACTGTTCCAAATGCTAGCAGAACAAATGGAAGTTTCTGGA
GATGTTGCGCAGCTCTCTGACGAACTCTATAATAGACGGACCCCTACCTCCGATACACGACTCG
TGGGATTTTAAGTACGGCGCAGGGTGTGGACATTICACGGTAAACCCCTGAGCAACCACTGA
CTCAGTACGACTGAACTCTGAAACATACTTATGATTACCATCACAGGAAATCTACATAATCCA
AGACTCAGGAAAGTAAGTGAAGAATATAACAAACGGACTCCCCAATTGAGAACACCG
AACGTTGGCCTCTCCACTCAGTCTGGTCTATCCAGTATAAGAGAGCACAAGTGAATGAGGACA
TTATGAGGAAAGTGGAGGAGTGGAAAGAAATGCAAGTACAACAGAGACATTATCATCCGGTTAA
GTTGGGAACTCTATCGTGAAGATGGCGGGCTGGGTACAAATGGTCAAGAACATCTATATAAGCC
GCCAACTATCAGTATAACTACTGAGAGACGGCAGCAGGTAACCGCCACACAACATGCTGTCA
ACGGGTTATAACTTAACTGAGGAGCTGGAGGACTCTGGTATCAGCCGGTATGAA
AGTCACTGAGGAAATTCTATGTGACTGAGTAACTGGGATGATAGCAAAGCGTTCGGCAACATG
GTGATGTTAGGAGCCTGGTCTGATAATCTCAATTCTGTAAGTGTACTGGTGGATCATATTATCTC
AATTCCGTTGAGGAGTCTGGCTGCTGAGTAACTGCGGGAGTCTCCCTCATTITGCTGGCTGACGT
TGACGACTCAGTTACCGATTCTGCTGACTGAACTCTGGTGGCTTCTCCTACTGTGAGC
AGCCCCCATTGAGCATCTGCTGACAAGAACTGTCACACCTCTACGGGTTACCTGCCGCAATCCAA
CAACGGCATGAATGAAATTTCGGGGCTCTTCTTGTATAAGTCTGGTACCAACTAATGAC
GACTATCAGACACCCATCATGACACGGTCACTGAGGAGATTGCAATGGCCAATTGATGATCTGCCACT
CTGGGGAGAATTCATCTCAGTCACTGGTCAAGGAGCTGGAAAGACAACATTACAGAT
GCCCTGCTGATATGTTAGTGTGTTCTCCGGCATAAATCAACTATAGATCTGACAAAGAGCATGGCTA
CTTCTGTTGAGAAGTGGAGGAGTGGAGGAGTCTCGCTTAAAGTCAACACCCCTGGACGAGCT
GAGTGTGAGCATCGCTGAGCATCTGCTGAGGAGTGGCTGAGGAGTGGAGGAGTGGCT
TGGACAAACGTCACACGGTCAACGTCAGCAACGTAACCTCTGGTGAACGCTATTCTACCCAGACGT
CAACGATCAGTAAAGAAACTCCGCTGAAAGAAATGATCACCCAGACTGAGGGAAATGCTTGCACGA
CATTTCCGCGCGCTGCTAAAACCAAATCGATATGCTACTCTGAGATCGGCAAAGAACACTCTCCG
GATATCGTAAACCGAAGCCTCGAAAGGGAAAGTTTACCTTAAGGGCAGCTACAGAATATGAAAGATGACG
AGGTGATGGAGATCAACACAGAGACAGGAGTCTCGCTTAAAGTCAACACCCCTGGACGAGGTTCC
GTTGACGTCATAAGTTGAGGAGCTGAGCAGATAGTCCAGTGTGATTCTGCGCATATTGACTTTA
AGACTTGAAGAACCTGAAACGACAATGGAATAACACGGACCGAAGCGTGAACCTCTTAAAGT
CAATCCCAATATGTCGCAATTCTCATTAAGGCTTATTCTTAACTGTTAATTACTGTTATTGCGAT
TTAATCTGCAATGAAACTCTGAGGGCTTATTCTGCTGAGGAGTGTGTTATTGAAATTCTTGTG
TCTCTATGTTGAGCGGTTTCTGCTGAGGAGTGTGTTATTGAAATTCTTGTGAAATTCTTGTG
GCTCTGTTAGCGGCTCCCTCAGCAAGGACACAAAAGATGTTTAAAGGAGTGTGTTATTGAA
AAAAAAAGACGGGAAATCTGATATAAGCTTACGCTTACGCTGAGATCTGACGAACTCTGGCAATAA
AGTTCTTAAGGTTGAATCTGTCGGGTCTGGCTGATTATCATATAATTCTGTTGAATTACGTT
AAGCATGATAATTAAACATGTAATGCGATGACGTTATTATGAGATGGGTTTATGATTAGAGTCCC
GCAATTATACATTAAACGCGATAGAAAACAAATAGCGCAGAACACTAGGATAAATTATCGCGC
GCGGTGTCATCTATGTTACTAGAT

```

도면32

도 32, SEQ ID NO: 36

로타바이러스 A Rotarix 균주의 VP4의 아미노산 서열

```

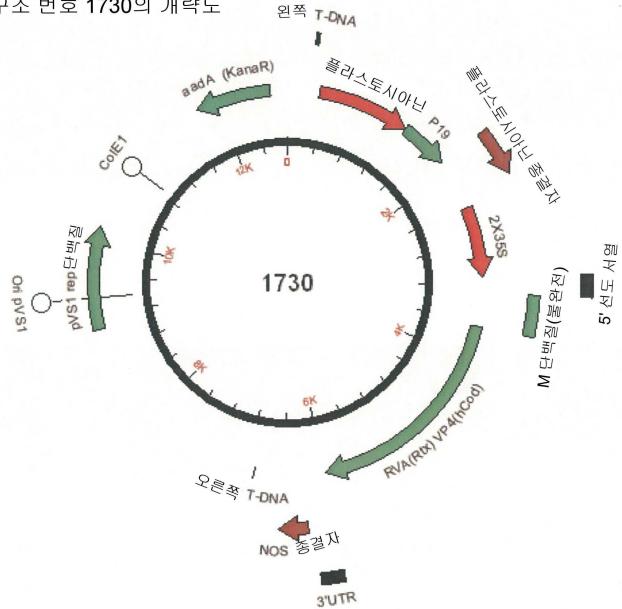
MASLIYRQLLTNSYSSVLDHDEIEQIGSEKTNQVNTINPGPFAQTRYAPVNWDHGEINDSTTVEPILDGPYQPT
MASTPPNDYWLINSNTNGVYESTNNSDFWTAVVAIPEPHVNVPDROQMIFGESKQFNVSNDNSNKWKFLE
MFRSSQQNEFYNRRLTSSTDTRLVGFIFYGRVRWFHGETPRATTDSSTANLNNISITIHSEFYIIPRSQESK
CNEYINNGLPPIONTRNVPLPLSSRSIQYKRAQVNEDIIVSKTSWIKEMQYNRDIIRPKFGNSIVKMGL
GYKWEJSYKAANYQNYLRDGEQVTAHTCSVNGVNNFSYNGFLPTDFGIRSYEVIKENSYVVVDY
WDDSKAFRNRMVYVRSLAANLNSVKCTGGSYFSIPVGAWPVMNGGAVSLHFAGVTLSFTDFVSLNS
LRFRISLTVDEPPFSILRTTVNLVLYGLPAANPNNGNEYYEISGRFSLISLVPNTDDYQTPIMNSVTVRQDLE
RQLTDLREEFNSLSQEJAMAQLIDLALLPLDMFMSMFSGKSTIDLTKSMATSVMKKFRKSKLATSISEMTNS
LSDAASSASRNVSIRSNLNSAISNWNTVNSDVNSVNTSNSLDISTQTSTISKLLRKEMITQTEGMSPFDISAA
VLKTKIDMSTOIQGKNTLPDIVTEASEKFIPKRSYRILKDEVMIEITEKFFAYKINTFDEVVPFDVNKFEL
VTDSPVISAIDFKTLKNLNDNYGITRTEALNIKSNPMLRNFINQNPNPIIRNRIEQLILQCKL

```

도면33a

도 33A

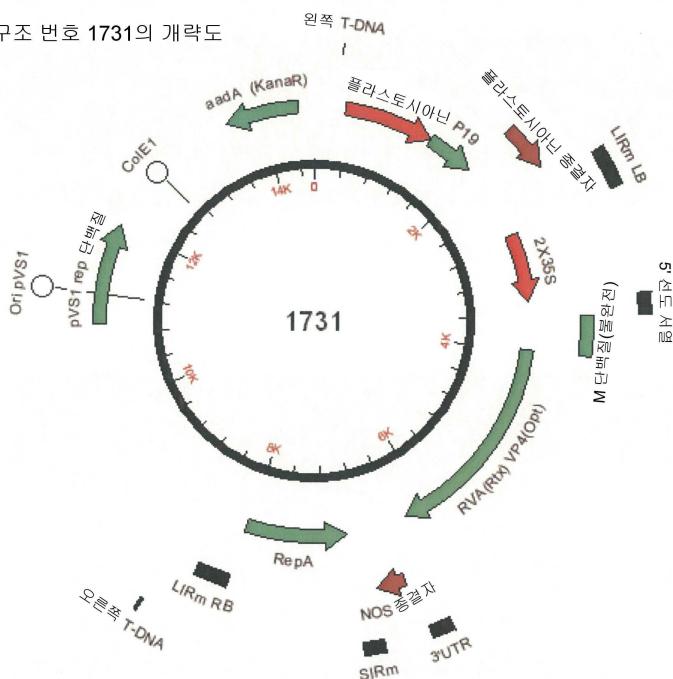
구조 번호 1730의 개략도



도면33b

도 33B

구조 번호 1731의 개략도



도면34a

F-2X35S/CPMV-HT/RVA(Rtx) VP7(Opt)/NOS (구조 번호 1733)

도 34A, SEQ ID NO: 37

IF-Rtx_VP7(opt).s1+3c

AAATTTGTCGGGCCATGTACGGCATCGAGTATAACAACAATTAAATTTC

도 34B, SEQ ID NO: 38

IF-Rtx_VP7(opt).s1-4r

ACTAAAGAAAATAGGCCTCTAACCGCGATAATAGAAGGCTGCTGAGTTCA
GGGA

도 34C, SEQ ID NO: 24

2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1733.

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의

VP7은 밀줄 그어져 있다.

GTCAACATGGTGGACGACACACTTGTCTACTCCAAAAATATCAAAGATAACAGTCTCAGAAAGACC
AAAGGCAATGGACTTCAACAAAGGTAATATCCGAAACCTCCTCGGATTCCATTGCCAGC
TATCTGTCACTTATGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATATCGAT
AAAGGAAAGGCCATCGTGAAGATGGCTCAGTGGCTCCAAAGATGGACCCCCACCCACG
AGGAGCATCGGAAAAGAAGACGCTTCAAAAGCAAGTGGATTGATGTGATAAC
ATGGTGGAGCACGACACACTTGTCTACTCCAAAATATCAAAGATAACAGTCTCAGAAGACCAAAG
GCAATTGAGACTTCAACAAAGGTAATATCGGAAACCTCTCGGATTCCATTGCCAGTATCTG
TCACTTATGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATATTGCGATAAAAGGA
AAGGCATCGTGAAGATGCCCTCGGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGC
ATCGTGGAAAAGAAGACGTTCACCAACACCGTCTCAAAGCAAGTGGATTGATGTGATATCTCACTG
ACGTAAGGGATGACGCACATCCACTATCCTCGCAAGACCCCTCTCTATAAAGGAAGGTTCATT
CATTGGAGAGGTTAAATCTTAAAGGTTGATAAAAGCGAACGTTGGGAAACCCGACCAA
CCTCTCTAAACTCTCTCATCTCTTAAAGCAAACCTCTCTCTGCTTCTTGCGTGGCGATCT
TCAAGTGTGAGCTGCTTCGACCAGTACAAACGTTTCTTCACTGAAAGCGAAATCAAAGATC
TCTTGTGGACGTTAGTGGCGCCATTAAATAACGTGTTACTTGTCTAATCTTGTGGTGTGCT
TGGGAAAAGAAGCTGCTGGAGGCTGCTTCAAGCCCCATACATTACTGTTACGATTGCTGACT
TTCGGCGGGTGAATATCTACTCTGCTGCTGCTGAGGAGTATTGTTGCCIGTACTCTTCTTCT
TTGCTGATTGGTCTATAAGAAATCTAGTTTTCTTGAAGAACAGAGTTTCCGGTGGTTTCA
GGAGAAAGATTGTTAAGCTCTGTATATCTGCCAAATTGCGGGCCATGTAACGGCATCGAGTAT
ACAACAAATTAAACTACATCTTGTAAACTACATCTTAAAGTCCGTGACCGAAATT
ATGGAATTATATTATCTATCGTAGCTCCCTCATCTACGGCCCTTTTGCCCTGACCGAGGG
CTATGGCCTGAACCTACCAATCACCGGTTCAATGGATAACCGTTACGCTAATTCAACTCAAGAGGG
ATATTCTGACAAGAACCTGCTGCTGTTATTACACAGAACGCTGACCCAGATCAATGATGGGG
AGTGGAAAGGATAGTCTCTACAGATGTTCTAACCCAAGGGCTGCCAACCGGTTCCGCTACTCTAA
GGAATACTCTAGTATTGTCGACTCTCAGTTGACCCCCAGTTTATTGCGACTACAACCTGGTACTTA
TGAATACGACCAAGAACCTGGAGCTGGATGCTGGTGACCTGATCTCAATGAGTGGCT
GTGCAACCCAGACATGGACATCACATTATTTACTACCCAGCTGAGAATCCAACAAGTGGATCAGT
ATGGCTCAAGTGTGACCGTGAAGGTGTTGCCCTGAAACACCCAAATGCTGGCATGGTCA
CAACTAATGGAATTCTGTTAGGGCTTATTGTTAGTTGAATTTACTGTTATTGCTGCA
CTATGTTGGTGAAGCGGTTTCTGCTCAGAGTGTGTTATTGTTATGTAATTAAATTCTTGTGAGC
TCCTGTTAGCAGGTGCTGCCCTCAGCAAGGACACAAAAAGATTAAATTATTA
AAAAAGACCGGAAATTCGATATCAGCTTACGCTTACGTTACGTTACGTTAA
TTCTTAAGATTGAATCTGTTGCCGTCTGCGATGATTATCATATAATTCTGTTGAATTACGTTAA
GCATGTAATAATTAAACATGTAATGCATGACGTTATTGAGATGGGTTTATGATTAGAGTCCGC
AATTATACATTAAACCGCATAGAAAACAAAATAGCGCGAAACTAGGATAATTATCGCGC
GGTGTCACTATGTTACTAGAT

도면34b

도 34D

USA/Rotarix-A41CB052A/1988/G1P1A[8]의 VP7을 암호화하는
뉴클레오티드 서열 (SEQ ID NO: 48)

```
ATGTACGGCATCGAGTATAACAACAATTAACTTCTGATTTCCATCATTCTGTTAAACTACATCCTT
AAGTCCTGACCAGAATTATGGATTATATTATCTATCGTAGCCTCTCATCTACCGTGGCCCTTTTGC
CCTGACCCGGCAGAACATGGCTGAACCTTACCAATCACCGGTTCAATGGATACCGTTACGCT
AATTCCACTCAAGAGGGATATTCTGACAAGTACCCCTGTGCGTGTATTATCCAACAGAAGCCTAC
CCAGATAATGATGGGAGTGGAGAGTAGTCTCTCACAGATGTCCCTAACCAAGGGCTGCCACC
GGTCCGCTACTTCAAGGAATACTAGTATTGCGACTTCAAGTGAACCCCAGCTTATTCGGA
CTACACCTGACTTATGAAATACGACCCAGAACCTGGAGCTGAGATATGCGAACCTGACCTG
ATCCTAATGAGTGGCTGTGCAACCCCCATGGACATCACATTATTAACCTACAGCAGTCTGGAGAAT
CCAACAAGTGGATCAGTATGGCTCAAGTGCACCGTGAGGGTGTGCTCTGACACCCAAATGCT
GGCATTGGTTGCAAGACAACAAATGTTGAAATGGTAGCCGAAACAGAGAAGCTGGCT
ATAGTGGACGTAGTCGATGGGATTAAACCAAGATCAATCTGACTACCAACACTGTACCATCAGAA
ACTGTAAGGCTGGGAGAAGTGGGAGAAGCTGGGAGAAGCTGGGAGGCAATGTGCTCG
ACATTACTGCCGACCCCTACCAACATCACAGACGGAGTGTGAGAGTCAACTGGAAAGAAT
GGTGGCAGGTCTTATACCATGTGGACTACATTAACCAAGATTGTGCAAGTCATGAGTAAACGGTC
CAGATCCCTGAACTCAGCAGCCCTATTATCGCGTTAG
```

도 34E

균주 RVA/Vaccine/USA/Rotarix-A41CB052A/1988/G1P1A[8]의
로타바이러스 a vp7의 최적화된 암호화 서열

```
ATGTACGGCATCGAGTATAACAACAATTAACTTCTGATTTCCATCATTCTGTTA
AACTACATCCTTAACTCGTAGCAGAACATTGGATTATATTATCTATCGTAGCCTC
CTCATCTACGTGGCCCTTTTGCCCTGACCAAGGGCCAGAACATGGCCTGAACCTA
CCAATCACCGGTTCAATGGATACCGTTACGCTAATTCACACTCAAGAGGGATATT
CTGACAAGTACCCCTGTGCTCTTATTATCCAACAGAAGCCTAACCCAGATCAATGAT
GGGGAGTGGAAAGGATAGTCTCACAGATGTTCTAACCAAGGGCTGGCCACCGGT
TCCGCTACTTCAAGGAATACTCTAGTATTGCGACTTCTCAGTTGACCCCCAGCTT
TATTGCGACTAACACCTGGTACTTATGAAATACGACCCAGAACCTGGAGCTGGATATG
TCCGAGCTGGCTGACCTGATCCCTCAATGAGTGGCTGTGCAACCCCCATGGACATCACA
TTATTAACCTACAGCAGTCTGGAGAATCCAACAAGTGGATCAGTATGGGCTCAAGT
TGCACCGGTGAAGGTGTGCCCCCTGAACACCCAAATGCTGGGATTGGTTGTCAGACA
ACTAATGTGGATTGTTGAAATGGTAGCCGAAACAGAGAAGCTGGCTATAGTGGAC
GTAGTCGATGGGATTAAACCAAGATCAATCTGACTACCAACACTTGTACCATCAGA
AACTGTAAGGCTGGCCCCGGAGAACGTCGCCGTGATCCAGGTGGGGGGAGC
AATGTGCTCGACATTACTGCCGACCCCTACCAACATCCACAGACGGAACGGATGATG
AGAGTCAGTGGAAATGGTGGCAGGTCTTATACCATGTGGACTACATTAAC
CAGATTGTGCAAGTCATGAGTAAACGGTCAGATCCCTGAACTCAGCAGCCTTCTAT
TATCGCTTTAG
```

도면35

도 35, SEQ ID NO: 39

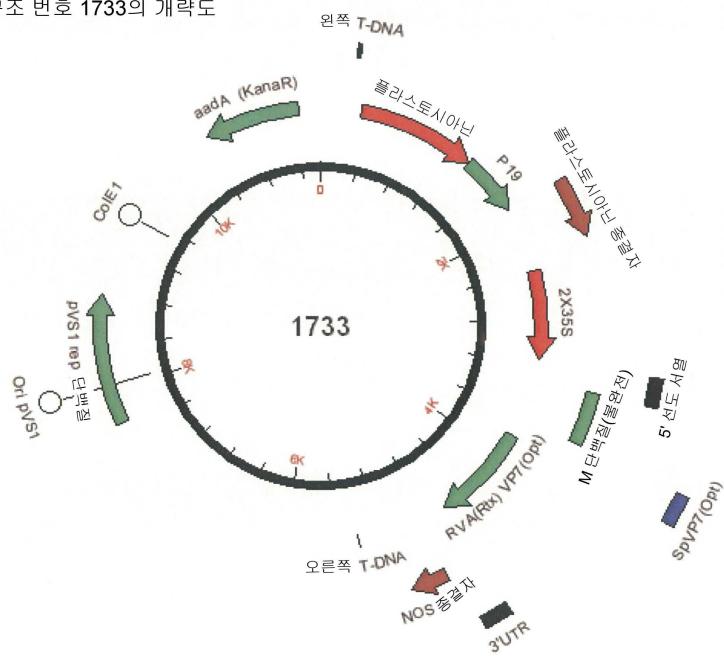
로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의
VP7의 아미노산 서열

```
MYGIEYTTILFLISIILLNYILKSVTRIMDYIIYRSLLIYVALFALTRAQNYGLNLPITGSMDTVYANSTQEGL
FLTSTLCLYYPT EASTQINDGEWKDSLQMFPLTKGWPTGSVVFKEYSSIVDFSDPQLYCDYNLVLMKYD
QNLELDMSELADLILNEWLCNPMDITLYYYQQSGESNKWISMGSCTVKVCPNTQMLGIGCQTTNVDS
FEMVAENEKLAIVDVVDGINHKINLTTCIRNCKKLGPRENVAIVQVGGSNVLDTADPTTNPQTERM
MRVNWKWWQFYITVDYINQIVQVMSKRSRSLNSAAFYRV
```

도면36

도 36

구조 번호 1733의 개략도



도면37

G-2X35S/CPMV-HT/ RVA(Rtx) VP7(opt)/NOS (구조 번호 1735)

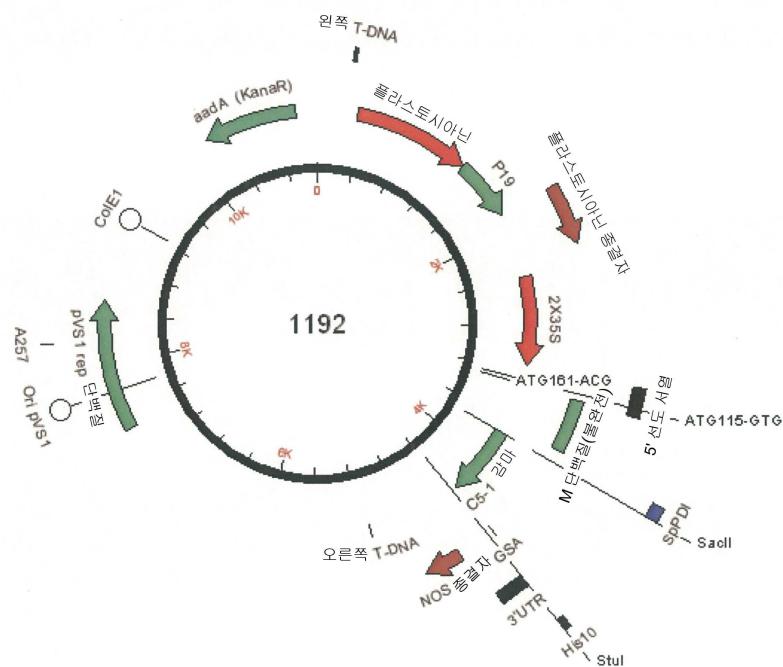
도 37, SEQ ID NO: 40

IF-Rtx_VP7(opt).s2+4c

TCTCAGATCTTCGCCAGAACTATGGCCTGAACCTACCAATCACCGTTCA
ATGGATACC

도면38

도 38

구조 번호 1192의 개략도. 플라스미드 선형화에 사용된
SacII 및 StuI 제한 효소 부위는 도면에 주석이 달려 있다.

도면39a

도 39, SEQ ID NO: 41

왼쪽 t-DNA 보더에서 오른쪽 t-DNA 보더로의 구조 1192(일줄 그어져 있음).

플라스토시아닌-P19-플라스토시아닌 침묵 억제자 발현 카세트가 있는

2X35S/CPMV-HT/PDISP/NOS

```
TGGCAGGATATAATGGGTGTAACAAATTGACCGCTTAGACAACACTAATAACACATTGCGGACGTT
TTAATGACTGAATTAACGCCGAATCCCGGGCTGGTATTTATATGGTCAAAATAACTCAAAAACC
ATAAAAGTTAAGTCAAGTGTCACATTACTGAAACAAAATTACCTACTACTGTATA
AACATTAAACATTAGAGAAGAAAATTGGTACAGTGTGATAAGACAAGAGTAGTGATATTGACAA
CAATTGTCACATTGAGAAAATTGGTCTCTCTTGTCAAAAACAAATAGAGAG
AGAAAAAGGAAGGGAGAAGAAAACATAATGGTACAGTGTGAGATAGAGAGAAGTGTACAAAGTG
TACAAAATAGTTGACAAATATGGAGAATTGACAAAAGCTACACAAATAAGGGTTAATTG
TGAAATAAAATAAGGATGACGCTAGAGAGAGATGACAGTGTGAGAATTGGCAAGTCATTAA
AGAAAAGATAATTGTTAAAGATGAGTCAATTGATTAACATGTGATTATTAA
TTGATGAAAGAGTTGGATTAAAGTGTAGTAGTAAATTGAGAATTGGTGTCAAATTAAATTGACATT
GATCTTCTTATATTGCCCCATAGCTCACTCATTTTATTCATAGTCACAAATAGA
GAAATAACCGTATATTAACTCCCAAAAAAAACGGTATTTACTACAAAAAACTAACGCCACG
TAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACAACTCACAAATCCGTATG
AGATAACCCACTTAAAGGCCACGATCTGGCACATCTACATATCTAAATCACACATCTTCCACA
CATCTGAGCCACACAAAACCACTTACACATCTTACCCATCTTACACACTTGTGAG
TCTACACTTGTACCTTCAACACATACAAAGAGAAGAGACTAATTAAATTAA
AGAATGGAACGAGCTACAGGAAACGACGCTAGGAAACAGCTAACAGTGAACGTTGGAGAT
GAGGATCAGGGATGACCTTCTCCCTAAACTCTGTACGAAAGTCCAGTTGACTGAGTGGG
GCTACATAACGATGAGACGAATTCGAATCAAGATAATCCCTTGGTCAAGGAAAGCTGGGTTTC
GGGAAGGTTAAAGGATACCTGCAGATACAGGACAGGAGCGAAGCTTCACTGACAGAGTCTTGT
GATCTGGACGGGAGATTGCTTAACATGAGCATCTGGTCAAGTGTGACAGATCGGATGT
ACCTATAGTATTGGTTGAGGAGTTAGTACCCGTTCTGGAGGGTGCAGACTCTTCAGCATCT
CTGGAGATGGCAAGGACTGCTACAGTGTGCCCCAACTGAAGTGGAAAGTGAAT
GTATCAAGAGGATGCCCTGAAGGTCACAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAATGTT
CTTCGTCCTCATTATAATATGGTTGTTATTGTTAATTGTTCTGTAGAAGAGCTTAAATACG
TGTGTTGTTAAACTATTTGAGATGAGTGAACCTGGTGTAAATTCAATTGAGTGGAGT
CGAAATCAGAATGTTCTCTTACAACTAACAGACATGAGCATGAGACCTGCCGCTACAAATTGCTTAA
TGAACAACAAATTGAACATCTTGCACAACTTATAAGTGGTTAATATAGCTAAATATGGT
CAAGTCAAGATAATTAAATGGAAATATCAGTTACAGTACAACTTACAACTAACGTTATT
AAACTAAATTATATCATCCCTTGTATAAGTATAGTACACCAATTAGGAAGGAGCATGCTGCC
TAGGAGATTGTCGTTCCCGCCTCAGTTGCAAGCTGCTCTAGGGTGTAGCCAATACGCAACCGC
CTCTCCCGCCTGGAAATTCTGCAGATGCTGACAGGCTGTCAGCTGCCAGATGGTGGAG
CACACACACTTGTCACTTCCAAAATATCAAAGATACTGCTCAGAAGACCAAAGGGCAATTGAG
ACTTTCAACAAAGGGATAATCCGAAACCTCTCGGATTCTTGCAGCTATGCTCACTTAT
TGTGAGAAGTGTGAGGAAAGGGCTCTAAAGTCAACTGCTACATTGCGATAAAGGGAAAGGCAT
CGTGAAGATGCCCTGCGGACAGTGGCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCAC
AAAGAAGACGTTCAACACCGTCTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCAC
ACACITGCTACTCTCAAAGATACTGCTCAGAAGACCAAAGGGCAATTGAGACTTTC
AAACAAGGGATAATCTGGAAACCTCTCGGATTCTTGCAGTCTGCAACTGCTATCTGACTTATTGGAAG
ATAGTGGAAAAGGAAGGGTCTCTCAAATGCACTATTGCGATAAAGGGAAAGGCCATGTTGAA
GATGCCCTGCGGACAGTGGCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCAC
GACCTTCTGCGGACAGTCTGCGGACACTTCTGCGGATTCTTGCAGTCTGCGGATCTGCTGAG
TACAGTGGGAAACCTCTGCGGACAGTCTGCGGACACTTCTGCGGACAGTCTGCGGACAG
TAGTGGCGGCCATTAAACACGTGACTGTCTTCTGCGGCTTGTGCGGAAAGAAGC
TTGCTGGAGGCTGCTGTTGAGCCCATACATTACTGTTACGTTCTGCTGACTTCTGGCGGGTGC
TATCTCTACTCTGCTGAGCAGGTTGTTGCGCTACTCTCTCTCTGCTGATTTGGTCT
TAAAGAAATCTGTTGTTGAAACAGAGTTTCCGCGGATTTGCGAACTTGGAGAAAGATGTT
AAGCTTCTGTTGTTGCGGCTCTCTGAGATCTTCTGCGGCTCTCAGCCTAACAGACCCCCCA
TCTGCTATCCACTGGCCCTGATCTGCTGCCAAACTAACCTGCTGTTGCGGACTTGGGATGCC
CAAGGGCTTCTGCGGAGCTGACAGTGGCTGACAGTGGCTGAGCTTCTGCGGCTGCG
ACCTCCCGAGCTGCTGAGCTGCTACACTGAGCAGCTAGTGAAGTGTGCTGACTTCTGCGG
CTGCGGCCAGCGAGACCGTCACTGCAACGTTGCCACCCGCCAGCAGCACCAAGGGAGAAGAA
AATTGCGGAGGAGTTGTTGAGTGAACCTGCGGATTTGCGAAGGATATCTGCTTCA
TCTTCCCCCAAAAGGCAAGGATGCTCACCAACTCTGACTCTAACGGTACAGTGTGAG
GACATCAGCAAGGATGATCCGAGGTCCAGTCACTGGTTGAGATGATGAGGTCACACAG
CTAGCAGCAACCCGGAGGAGCAGTCAACAGCACTTCCGCTAGTCACTGAACCTCCATCAT
```

도면39b

```
GCACCAAGGACTGGCTCAATGGCAAGGAGCGATGCTCACCATCACCATCACCATCACCAT
AAAGGCCATTCTTAGTTGAATTACTGTTACCTGCTGCAATTCTATGTTGGTGAAGCGGTT
TCTGCTCAGAGTGTGTTATTATGTAATTAAATTCTCTGAGCTCTGTTAGCAGGTGCTCC
CTTCAGCAAGGACACAAAAGATTTAAATTAAAAAAAGACGGGAAATTCG
ATATCAAGCTTACGACCTGCGAGATCGTCAACACATTGGCAATAAAAGTTCTAAGATTGATCTG
TTGCGGCTCTGCGATGATTATCATATAATTCTGTTGAAATTGTTGAGCTTACGATGTAATAATTACATG
TAATGCGATGACGTTATTATGAGATGGGTTTTATGTTGAGCTTACGAGTCCCGCAATTATACATTAAACG
GATAGAAACAAATATAGCGCGCAACTAGAGATAATTATCGCGCGGTGTCATCTGTTA
GATCTAGAGTCTCAAGCTTGGCGGCCACGTGACTAGTGGCACTGGCGTGTGTTAACACGTCG
TGACTGGGAAAACCTGGCGTACCCAACTTAACTGCGCTTGCAGCAGCATCCCCCTTGCAGCTGGC
GTAATCGGAAGAGGGCCGACCGATGCCCTCCAAACAGTGGCGCAGCCTGAATGGCAATGCTA
GAGCAGCTGAGCTGGATCAGATGTCGTTCCGCTTCACTGTTAAACTATCAGTGTGACAGGA
TATATGGCGGGTAAACCTAACAGAGAAAAGAGCGTTA
```

도면40a

도 40A, SEQ ID NO: 42

2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1735.

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의
PDISP/VP7(opt)[8]은 밀줄 그어져 있다.

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAATATCAAAGATACTCTCAGAAAGACC
AAAGGGCAATTGGAGACTTTCAACAAAGGGTAATATCCGGAAACCTCTCGGATTCCTGCCTGCAGC
TATCTGCACITATGTGAAGATACTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATGGCAG
AAAGGAAAGGCCATCGTGAAGATGCCCTGCCACAGTGGTCCAAAGATGGACCCCCACCCACG
AGGAGCATGTGGAAAAAGAACGTCACCCACGTCTCAAGCAAGTGATTGATGTGATAAC
ATGGTGGAGCACGACACACTTGTCTACTCCAAAATATCAAAGATACTCTCAGAAAGACCAAAAGG
GCAATTGAGACTTTCAACAAAGGGTAATATCCGAAACCTCTCGGATTCCTGCCTGCAGCTATCTG
TCACTTTATGTGAAGATACTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATGGCATAAGGA
AAGGCCATCGTGAAGATGCCCTGCCACAGTGGTCCAAAGATGGACCCCCACGGAGC
ATCGTGGAAAAGAACGACGTCCAACCCACGTCTCAAGCAAGTGATTGATGTGATACTCTCAG
ACGTAAGGGATGACGCCAACTATCCTCGCAAGACCCCTCCCTCATATAAGGAAGTTCATT
CATTGGAGGAGTTAAAATCTTAATGGTGGATAAAAGCGAACGTTGGGGAAACCCGAACCAAA
CCTCTCTAAACTCTCTCATCTCTTAAAGCAAACTCTCTGTCTTCTGGTGAACGATCT
TCAACGTTGAGCTGAGCTGCTGGCACCAGTACAACGTTTCTTACTGAAAGGAAATCAAAGATC
TCTTGTGGACACGTTGAGCTGCTGGAGGCTGCTGGAGCCACATTTAAGCTGTTACTTGTGCTGTT
TGGGAAAAGAACGAGCTGCTGGAGGCTGCTGGAGCCACATTTACTTGTGATCGATTCTGACT
TTGCGGGGTGCAATATCTACTCTGCTGACGAGGTTGCTGACTCTTCTCTCTTC
TTGCTGATTGTTCTAAGAAATCTAGTATTCTCTTGTGAAACAGAGTTTCCGTTGGTTGCAACTT
GGAGAAAGATGTTAAGCTCTGTATTTCTCCCAATTGTGCGGCCATGGCGAAAACGTC
GATTTCGGCTATGTTCTCTGTGTTGCTCAGATCTCGCCAGAACTATGGCT
GAACCTACCAATTCAGGCTTAATGGAATCCGTTACGCTAACTCCACTCAAGAGGGGATATTCTGA
CAAGTACCTGTGCTGTATTATCACAAGACGCTTACCCAGATACTGAGGGAGTGGAAAGGA
TAGTCTCACAGATGTTCTAACCAAGGGCTGGCCACCGGTTCTGACTTCAAGGAATACTCTA
GATGTCGACTCTCAGTGTGACCTGTTACGCTGACTACAACCTGGTACTTATGAAATACGAC
CAGAACCTGGAGCTGGATATGCGAGCTGCTGACCTCIAATGAGTGCCTGTCACCCCA
TGGACATCACATTATTTACTACCAAGCAGTCTGGAGAATCCAACAAGTGGATCAGTATGGCTCAAG
TIGCGTTGAGGCTGAAGGTGIGCCCTGAACACCCAAATGCTGGCATTGTTGTCAGACAACAAATGTT
GATTCTGTTGAGGCTATTTCTCTGTTGAAATTACTGTTACTCGTGTGCTGATTCTATGTTGTT
AGCGTTTAGAGGCTTATTTCTCTGTTGAAATTACTGTTACTCGTGTGCTGATTCTATGTTGTT
AGCGTTCTGCTGAGGCTGTTGTTGAAATTACTGTTACTCGTGTGCTGATTCTATGTTGTT
GGTCTGCTCTGAGGACACAAAAAGATTAAATTAAATTAAATTCTTGAGCTCTGTTAGCA
GGATCTGCTCTGAGGACACAAAAAGATTAAATTAAATTAAATTAAATTAAAGACCG
GGAATCTGATATCAAGCTGACCTGAGCTGCTGAAACATGTTGCAAAATGTTCTTAAGATT
GAATCTGTTGCCGGCTTGTGAGGATATTATCATAATTCTGTTGAAATTACGTTAAGCATGTAATAA
TTAACATGTAATGACGTTATTATGAGATGGTTTATGATTAGAGTCCCAGAATTACATT
TAATACGGATAGAAAACAAAATAGCGCGAAACTAGGATAATTATCGCGCGGTGTCATCTA
TGTACTAGAT

도면40b

도 40B

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의
PDISP/VP7(opt)를 암호화하는 뉴클레오티드 서열 (SEQ ID NO:62)

ATGGCGAAAACGTGCGATTTCCGTTATTGTTTCTCTTGTGTTGGTCTCTCAGATCTC
GCCAGAACTATGGCTGAACCTACCAATCACGGTCAATGGATACCGTTACGCTAATCCACTCA
AGAGGGGATATTCTGACAAGTACCCCTGTGCGTGTATTATCCAACAGAACGCTCTACCCAGATCAAT
GATGGGAGTGGAAAGGATAGCTCTCAGAGTTCTCTGTTACCAAGGGCTGGCCACCGGTTCCGCT
ACTTCAAGGAATACTCTAGTATTGTGCACTCTCAGTGTGACCCCCAGCTTATTGCGACTACAACCTG
GTGTTATGAAATGACCGAGAACCTGGAGCTGATATGTCCGAGCTGCTGACTCTGATCTCAATG
AGTGGCTGTGCAACCCATGGACATCACATTATTTACTACCAAGCAGTCTGGAGAATCCAACAAGTG
GATCAGTATGGCTCAAGTGTGACCGTGAAGGTGTGCTTGAACACCCAAATGCTGGCATTGGT
TGTGAGCACAACCTGGGATTCGTTGAAATGGTAGCGAAAACGAGAAGGCTGCTATAGTGGACG
TAGTCGAGCTGGATTAAACCAAGATCAATCTGACTACCAACCTTGTACCATCAGAAACTGAAAA
GCTCGGCCCCGGAGAACGCTGCGCTGATCCAGGTGGGGGGAGCAATGTGCTGACATTACTGCC
GACCCCTACCAACCAAGGAGATGAGCTGAGGATGAGCTGAGAAGAAATGCTGGCAGGTC
TTTATACATTGTGAGCTACATTAACAGGATTGTGCAAGTGAACGGTCCAGATCCCTGAA
CTCAGCAGCTCTATTATCGCTTGTAG

도면41

도 41, SEQ ID NO: 43

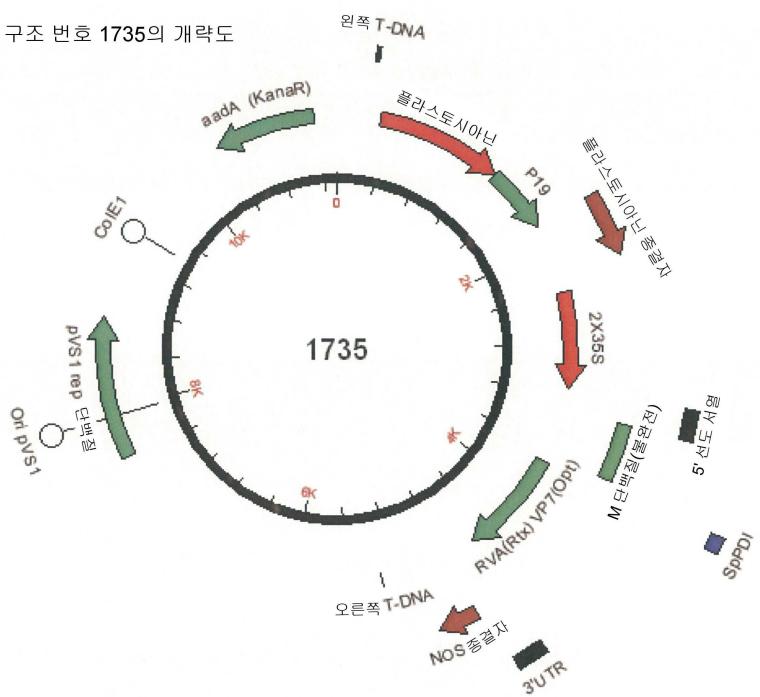
로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의
PDISP/VP7의 아미노산 서열

MAKNVAIFGLLFSLLVLVPSQIFAQNYGLNLPITGMDTVYANSTQEGLTSTLCLYYPTEASTQINDE
WKDSLSQLFLTKGWPVTSVYFKEYSSIVDFSVDPQLYCDYNLVLMKYDQNLELDMSSELADLILNEWLC
NPMDITLYQQSGESNEKWSMGSCTVVKVCPNTQMLGIGCQTNVDSFEMVAENEKLAIVDVVDGIN
HKINLTTTCTIRNCKKLGPRENVAVIQVGGSNVLDITADPTTPNQTERMMRNVWKKWWQVFYIVDYI
NQIVQVMSKRSRSLSNSAFFFYRV

도면42

도 42

구조 번호 1735의 개략도



도면43a**도 43A**군주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의
로타바이러스 A VP4의 암호화 서열 (SEQ ID NO: 50)

ATGGCTTCGCTCATTTAGACAATTGCTACGAATTCTTACAGTAGATTATCC
 GATGAGATACAAGAGATTGGATCAACTAAATCACAAATGTACAATTAACTCCTGGA
 CCATTGCGAACACAGGTTATGCTCCAGTTAACCTGGGACCTGGAGAAATTAATGAT
 TCTACGACAGTTGAACCATTGCTGGATGGCCTTATCAACCAACGACATTCAATCCA
 CCAGTCGATTATGGATGTTAGGCCAACACTGGCGTAATTGTTGAAGGT
 ACAAAATAACAGATAGTGGTAGCCACAATTAAATCGAGCCAATGTTCAGTCT
 GAAAATAGAACACTACACTATTTGGTATTCAGAACATTAACGGTATCCAATACT
 TCACAAGACCAGTGGAAATTATTGATGTCGTAACAAACTGCAAATGGAAGTATA
 GGACAATATGGACCATTACTATCCAGTCGAAATTATGCAAGTTGAAGCATAAT
 GAAAAATTATACATATGAAGGACAGACACCTAACGCTAGGACAGCACATTATTCA
 ACAACGAATTATGATTCTGTTAATATGACTGCTTTGTGACTTTTATATAATTCT
 AGATCTGAAAGACTCAAATGTACGGAAATACATTAATAATGGATTACCCAATACAA
 AATACTAGAAATGTTGACCATATTGCTGACTGCTAGAGATGTAACACTATAGA
 GCTCAAGCTAATGAAGATATTGTGATATCCAAGACATCATTGTTGAAAGAAATGCAA
 TATAATAGAGATATAACTATTAGATTTAAATTGCAAATACAATTATAAAATCAGGA
 GGGCTGGGATATAAGTGGTCAGAAATATCATTAAAGCAGCGAATTATCAATACACA
 TATACTCGTGATGGTGAAGAAGTTACCGCACATACTACTTGTTCAGTGAATGGCGTT
 AATGACTTCAGTTTAATGGAGGATATTACCAACTGATTGTTGTATCTAAATT
 GAAGTAATTAAAGAGAATTACAGTCTATATCGATTACTGGGATGATTCAAAGCA
 TTTCGTAACGTGGTGTATGTCGATCGTAGCAGCAAACCTTGAAATTCACTGTTATGTGT
 ACTGGAGGCAGCTATAATTAGTCTACAGCTGGACAATGGCCTGTTTAACTGGG
 GGAGCAGTTCTTACATTCACTGGTGTAAACACTACTTACAGTCAATTACAGATTC
 GTATCATTAAATCATTAAGATTAGATTAGACTAGCTGTCGAAGAACACACTTT
 AAACCTGACTAGAACTAGATTAGATAGATTGATGGCTGCTGCTGAGATCCAAT
 AATGGTAAAGAAATTATGAAATTGCTGACGATTTCACTTATATCATTAGTGCCA
 TCAAATGATGACTATCAGACTCTATAGCAAACACTCAGTTACTGTACGACAAGATTT
 GAAAGGCAGTTAGGAGAACTAAGAGAAGAGTTAACGCTTGTCTCAAAGAAATTGCA
 ATGTCGCAAGTTACGATTAGCGCTTCTACCAATTAGATATGTTCTCAATGTTTCT
 GGCATTAAGACTATTGATGCTGCAAATCAATGGCTACTAAATGTTATGAAAAAA
 TTCAAAAGTCAGGATTAGCGAATTCAGTTCAACACTGACAGATTCTTATCAGAC
 GCAGCATCATCAATATCAAGAGGTTCATCTACGTTGATGGATCTCAGCATCA
 GCATGGACGGATGTATCAACACAAACTGATATATCGTCATCAGTAAGTTCACT
 TCGACACAAACGCAACTATCAGTAGAGATTGAGACTAAAGGAATGGCAACACAA
 ACTGAGGGTATGAATTGATGATATACGCGGCTGTTGAAGACTAAGATAGAT
 AAATCGACTCAAATATCACCACAAACACAAACTGACATTGTTACTGAAGCATCGGAA
 AAATTCAACAAATAGGGCTTACCGTGTATAAACACAGTATGTTGAAGCT
 GGAATTGATGGAAAATTGCTTATAAAGTGGATACATTGAGGAAATACCATT
 GATGTACAAAATTCGCTGACTTAGTTACAGATTCTCAGTAATATCCGCTATAATT
 GATTTAAACACTTAAATTGAAACGATAATTACGCGATTACTAAGCAACAAGCA
 TTTAATCTTTAAGATCTGACCCAAGAGTTACGTGAATTCAATTACAGGACAAT
 CCTATAATTAGAAATAGAACACTGATTATGCAATGCAGGTTGTGA

도 43B군주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의
로타바이러스 A VP4의 최적화된 암호화 서열 (SEQ ID NO: 51)

도면43b

ATGGCTTCATTGATATATGCCAGTTGCTGACTAATAGCTATACTGTGGATTTGTCA
 GACGAAATCCAGGAAATAGGATCCACAAAGAGTCAGAACGTGACCATAAACCCCGGA
 CCGTTCGCCAGACTGGGTATGCCCGTAAACTGGGCCCGCGAGATTAACGAC
 AGCACCACCGTGAGGCCACTGCTGGATGGACCCCTACCAACCCACTACTTTAACCT
 CCAGTGGACTACTGGATGTTGGCTCCACGACACCTGGTGTAAATTGTAGAGGGC
 ACCAACAAATCCGATCGCTGGCACAATACTGATAGAACCCAACGTGCAGTCC
 GAGAACAGAACCTATACCATTCGGCATTCCAGGAACAGCTAACCGTGAGCAATACG
 AGCCAGGACCGAGTGAAGTTATCGATGTAGTAAAACCTACGCCATGGATCTATC
 GGGCAATACGGGCGCTGCTGCCTCACCTAACGCTACGCCGTGATGAAACATAAT
 GAGAAAACGTACACTTACGAGGGCAAACCCCAATGCCAGAACCTGCCACTACAGT
 ACAACCAACTATGACTCGGTGAACATGACAGCGTTGTGATTTTATATTATCCA
 AGATCAGAAGAACTCAAGTGTACTGAGTACATCAACAAATGGACTTCCACCCATCCAG
 AACACTCGAAATGTCGTCCTACTGTCCTAACGCTCGGGATGTGATCCACTATCGC
 GCCCAAGCTAATGAGGATATAGTCATTCAAAGACGAGCTTATGAAAGGAAATGCAG
 TATAACAGAGACATCACAACTCAGGTTCAAGTTCGCAAACTATTATAAGTCCGGG
 GGACTGGGTACAATGGAGTGGAGATCAGTTAACGGCGCTAACATCAGTACACC
 TATACTCGCACGGCGAACAGAGCTAACCGCCACACAACTTGCTCGGTTAATGGCGTG
 AACGATTTAGCTTCAACGGGGCTACCTGCTACTGATTTCGTGGTGAGCAAGTTT
 GAAGTCATCAAGGAAAATTCTACGTGTATATTGACTACTGGGATGATAGCCAGGCC
 TTCCGAAATGTTGTATGTTAGATCACTGGCGCAAACCTTAATTCAAGTCATGTG
 ACAGGAGGTCTACAATTTAGTCTTCCCGTGGCAGTGGCAGTGCTCACAGGG
 GGGCCTGTGAGCTTGCATTCCGCCAGTCACCTTGAGTACTCAGTCACAGACTTT
 GTGTCCTGAATAGCCTAACGGTTCAAGGTTAGACTTGAGTAAAGAGCCTCAGTT
 AAGCTCACTCGTACGAGGCTGGATCGGCTGTACGGCCTGCCGGCTGATCCCAAT
 AACGGCAAGGAATATTACGAGATAGCCGGAGATTTCGCTGATCAGTCTGGTCCG
 TCAAACGATGATTACCGACCCCCATTGCCAACAGTGTCACTGTCAGGCAAGATCTG
 GAGAGACAACTTGGCGAGCTGAGAGAGGTTCAACGCCCTGTCAGAGATCGCA
 ATGTCCTAGCTCATTGACCTGGCCCTGTTACCCCTGACATGTTCTCAATGTTCTCC
 GGCATAAAATCCACTATCGACGCTGCAAAGTCCATGCCACAAATGTGATGAAGAAG
 TTTAAGAAGAGCGGTCTGGCAAATAGCGTGTACGCTGACCGATAGTTGCGGAT
 GCCGCCAGTTCATTACCCGTGGATCCACCATAGTCCATTGGCTTCCGCCCT
 GCTTGGACTGACGTGAGTACACAGATAACTGACATTCTCTCTGCTCTCCAGTGTG
 AGCACACAAACTCCACGATATAAGACGACTGAGGCTCAAAGAGATGGCAACGCAA
 ACGGAAGGATGAAATTGATGACATCAGCGCCGAGTTGAAGACAAGAGATCGAT
 AAAAGCACTCAAATTAGCCCCAATACGATCCCTGACATTGTGACTGAGGCATCTGAA
 AAGTTATTCCCAACCGTCTTATCGGGTCTTAACAATGATGATGTTCTCGAGGCC
 GGCATCGATGGCAAGTTTGCTTAAAGTGGATACCTTCGAGGAGATTCTTTC
 GATGTCAGAAGTTGCTGACCTCGTAACGGATAGCCAGTGATAAGGCCATTATA
 GACTTAAACACATTGAAAAATTGAAACGATAATTATGGTATTACCAAGCAGCAGGCT
 TTTAACTTGTAAAGATCTGACCCCTCGCTGCTCAGAGAGTTATTAAACCAGGACAAC
 CCCATCATCAGAACAGGATCGAGCAGTGATTATGCACTGTCGCTGTAA

도 43c

군주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의
 로타바이러스 A VP7의 암호화 서열 (SEQ ID NO: 52)

ATGTATGGTATTGAATATACCAACAGTTCTAACCTTCTGATATCGATTATTCTACTA
 AATTACATACTTAAATCTTAACGAGATAATGGACTTTATAATTATAGATTCTT
 TTTATAATTGTGATATTGTCAACCTTCTCAGAGCACAAATTATGGTATTATCTT
 CCAATCACAGGCTCCATGGACACTGCATACGCTAACGCAAGAACATTC

도면43c

CTCACTTCTACACTTGCTATATTATCCGACTGAGGCTGCGACTGAAATAACGAT
AATTATGGAAAGACACACTGTCACAACATTCTACGAAAGGGGGCCAACCTGG
TCCGTATATTTAAAGAATATACTAACATGCATCGTTCTGATCCGCAGTTG
TATTGTGATTATAACGTAGTACTAATGAAATATGACCGACGTTGCAATTGGATATG
TCAGAATTGCGATCTAATATAACGAATGGTTGTAATCCAATGGATATTACT
CTGTATTATTATCAGCAAACGTACGAAGCGAATAAATGGATATCAATGGCTCATCA
TGTACAATTAAAGTATGTCCTAACACAAAACCTTGGAAATTGGATGCTTGACA
ACTGATCTACAACTTTGAAGAAGTTCGACAGCTGAAAAGTTGCTAATTACTGAC
GTGGTTGATGGCGTTAATCATAACTGCTGGATGTCACAAACAGAACGTGACTATTAGA
AACTGTAAGAAATTGGGACCAAGAGAAAACGTAGCCGTTACAAAGTTGGTGGTCT
GACATCCTCGATATAACTGCTGATCCAACACTGCACACAGACAGAACGGATGATG
CGAATTAACTGGAAAAATGGTGGCAAGTTTACTGTTACTGAGACTATGTAGAT
CAGATAATACAAGTTATGTCACAAAGATCAAGATCACTAAATTCAAGCAGCATTAT
TACAGAGTGTAG

도 43D

군주 RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]의
로타바이러스 A VP7의 최적화된 암호화 서열 (SEQ ID NO: 53)

ATGTACGGAATCGAGTATCACCCGTTCTGACATTCTATTAGTATTATCCTCTG
AACTATATTCTGAACTCACTACCCGATAATGGATTATTATATATAGGTTCTG
TTCATCATTGTAATTCTGAGGCCCTTCCTGAGGGCCAGAATTACGGCATAAACCTA
CCAATCACCGGTTCTATGGATACCGTTATGCTAACTCTACACAAGAGGAGACATT
CTCACATCAACCTATGCCCTACTATCCGACTGAAGCAGCCACAGAGATAAACGAT
AACTCTGGAAAGATACTTGAAGCCAGCTTCCTGACTAAGGGATGGCCACCGGA
TCGGCTACTTAAAGGAGTACACAAACATCGCAAGTTCAAGCGTGGATCCCCAGCTG
TATTGTGATTATAACGTTGCTGATGAAATACGACGCAACCTCCAGCTTGACATG
AGCGAGTTGGCAGACCTAATCCTCAATGAGTGGCTGTAACCCAATGGATATAACA
CTGTACTATTATCAGCAGACCGATGAAGCAGGAAACAAATGGATTCAATGGGAAGCAGC
TGTACCATCAAGTTGTCCTCTAACACACCAACTCTCGGATAGGGTGTCTGACC
ACAGACGCTACTACCTTTGAAGAAGTTCGACCGCGAAAAGCTGGTTATCACAGAT
GTGGTAGATGGCGTTAACACAAATTGGACGTAACCACAGCAACATGCACAATTAGG
AACTGCAAGAAGCTAGGACCCAGGGAAAACGTAGCCGTCATCCAAGTGGCGGAGT
GACATCCTAGACATCACCGCAGACCCACACAGCACCACAAACCGAGAGGATGATG
CGCATTAAATTGGAAGAAATGGTGGCAGGTGTTTACACTGTCGTTGACTATGTGGAC
CAGATCATTCAAGGTGATGAGCAAGCGGAGTCGCTCATTGAATAGTGTGCTGCCTTAT
TACAGAGTCTAA

도면44a

D-2X35S/CPMV-HT/RVA(Rtx) VP7(Opt)/NOS (구조 번호 1734)

도 44A (SEQ ID NO: 55)

IF-TrSP+Rtx_VP7(opt).s1+3c

AAATTGTGCGGGCCATGGATTATATTATCGTAGCCTCCTCATCTA

도 44B (SEQ ID NO: 56)

IF-Rtx_VP7(opt).s1-4r

도면44b

ACTAAAGAAAATAGGCCTCTAACACGCGATAATAGAAGGCTGCTGAGTCAGGGAA

도 44C (SEQ ID NO: 57)

균주 RVA/Vaccine/USA/Rotarix-A41CB052A/1988/G1P1A[8]의
로타바이러스 A VP7의 최적화된 암호화 서열

```
ATGTACGGCATCGAGTATAACAACAATTAACTTCAATTCCGATTTCCATCTGGTTA
AACTACATCCTTAAGTCCGTGACCAGAATTATGGATTATATTATCTATCGTAGCCTC
CTCATCTACGTGCCCTTTGCCCTGACCAGGGCCCAGAACTATGCCCTGAACCTTA
CCAATCACCGGTCAATGGATACCGTTACGCTAATCCACTCAAGAGGGATATT
CTGACAACTACCCCTGTGCCGTATTATCAAAGAAGCCTCACCCAGATCAATGAT
GGGGACTGGAAGGATAGTCTCTCACAGATGTTCTAACCAAGGGCTGGCCCACCGGT
TCCGCTACTTCAGGAATACTTAGTATTGTCGACTCTCAGTTGACCCCCAGCTT
TATTGCGACTACAACCTGGTACTTATGAAATACGACCAAGACCTGGAGCTGGATATG
TCCGAGCTGGCTGACCTGATCCTCAATGAGTGGCTGTGCAACCCATGGACATCACA
TTTATTACTACCAGCAGTCTGGAGAATCCAACAAGTGGATCAGTATGGGCTCAAGT
TGCACCGTGAACACCCAAATGCTGGGATTGGTTGTCAGACA
ACTAATGTGGATTCTGGTGAATGGTAGCCGAAACGAGAAGCTGGCTATAGTGGAC
GTAGTCATGGGATTAACCACAAGATCAATCTGACTACCACCTGTACCATCAGA
AACTGAAAAAGCTGGCCCCGGAGAACGTCGCCGTGATCCAGTGGGGGGAGC
AATGTGCTGACATTACTGCCGACCCCTACCACCAATCCACAGACGGAACGGATGATG
AGAGTCACATGGAAGAAATGGTGGCAGGTCTTATACCATTGTGGACTACATTAAC
CAGATTGTGCAAGTCATGAGTAAACGGTCCAGATCCCTGAACTCAGCAGCCTTCTAT
TATCGCGTTAG
```

도 44D (SEQ ID NO: 58)

2X35S 프로모터에서 NOS 종결자로의 발현 카세트 번호 1734.
RVA/Vaccine/USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의
VP7은 밀줄 그어져 있다.

```
GTCAACATGGTGGAGCAGCACACACTTGTCTACTCCAAAAATATCAAAGATACTGTC
TCAGAACGACAAAGGGCAATTGAGACTTTCAACAAAGGGTAATATCCGAAACCTC
CTCGGATTCCATTGCCAGCTATCTGTCACCTTATTGTGAAGATAGTGGAAAAGGAA
GGTGGCTCTACAAATGCCATCATTGCGATAAAAGGAAGGCCATCGTGAAGATGCC
TCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGACATCGTGGAAAAA
GAAGACGTTCCAACCACGCTTCAAACGCAAGTGGATTGATGTGATAACATGGTGGAG
CACGACACACTTGTCTACTCAAAATATCAAAGATACTGTCAGAAGACCAAAGG
GCAATTGAGACTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGC
CCAGCTATCTGTCACTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAA
TGCCATATTGCGATAAAAGGAAGGCCATCGTTGAAGATGCCCTGCCGACAGTGGT
CCCAAAGATGGACCCCCACCCACGAGGACATCGTGGAAAAGAAGACGTTCCAACC
ACGTCTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCA
CAATCCCACTATCCTCGCAAGACCCCTCCTATATAAGGAAGTTCATTTCATTTG
GAGAGGTATTAAATCTTAAAGGTTGATAAAAGCGAACGTGGGAAACCGAAC
CAAACCTTCTCTAAACTCTCTCATCTCTTAAAGCAAACCTCTCTTGTCTT
TCTTGCCTGAGCGATCTCAACGTTGTCAGATCGTGCCTCGGCACCAGTACAACGTT
```

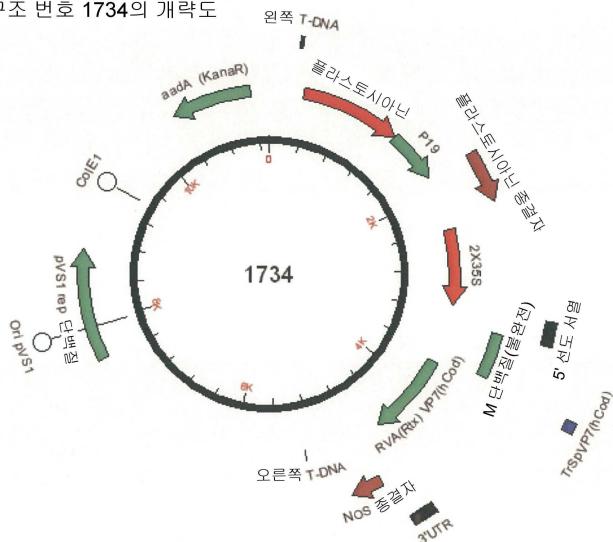
도면44c

TTCTTTCACTGAAGCGAAATCAAAGATCTTGTGGACACGTAGTGC GGCGCCATT
 AAATAACGTACTTGTCTATTCTGTCGGTGTGCTTGGGAAAAGAACGTTGC
 TGGAGGCTGCTGTTCAAGCCCCATACATTACTGTTACGATTCTGACTTCCGGCG
 GGTGCAATATCTACTTCTGCTGACGAGGTATTGTTGCTGTACTTCTTCTTCT
 TCTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTCTTCTTGAACAGAGTTTC
 CCGTGGTTTCGAACTTGGAGAAAGATTGTTAACGTTCTGTATATTCTGCCAAATT
 TGTCGGGCCCATGGATTATATTATCTATCGTAGCCTCCTCATCTACGTGGCCCTTT
 TGCCCTGACCAAGGGCCAGAACATGGCCTGAACCTAACCGTTCAATGGGA
 TACCGTTACGTAATTCCACTCAAGAGGGATATTCTGACAAGTACCCGTGCCT
 GTATTATCCAACAGAACGCTCACCCAGATCAATGATGGGAGTGAAGGAGATGCT
 CTACACAGATGTTCTAACCAAGGGCTGGCCACCGTCTCGTCAACTCAAGGAATA
 CTCTAGTATTGTCGACTTCTCAGTTGACCCCAAGCTTATTGCGACTACAACCTGGT
 ACTTATGAAATACGACCAAGAACCTGGAGCTGGATATGTCGAGCTGGCTGACCTGAT
 CCTCAATGAGTGGCTGTGCAACCCCAGAACATCACATTATATTACTACCAGCAGTC
 TGGAGAATCCAACAAAGTGGATCAGTATGGGCTCAAGTGTGACCGTGAAGGTGTG
 CTTGAACACCCAAATGCTGGCATTGGTGTGACAGAACTAATGTTGAGTGGCTTGA
 AATGGTAGCGAAAAGAGGAGCTGGCTATAGTGGACGTTAGTCGATGGGATTAACCA
 CAAGATCAATCTGACTACCACACTTGTACCATCAGAAACTGTAAGGCTCGGCCC
 CCAGGAGAACGTCGGTGTACCCAGGTGGGGGAGCAATGTGCTGACATTACTGC
 CGACCCCTACCAACCAATCCACAGACGGAACGGATGATGAGAGTCAACTGGAAGAAATG
 GTGGCAGGTCTTATACCATGTTGACTACATTAACCAAGATTGTCGAAAGTCATGAG
 TAAACGGTCCAGATCCCTGAACTCAGCAGCCTTCTATTATCGCCTTAGAGGCTAT
 TTTCTTAGTTGAATTACTGTTATTGGTGTGCATTCTATGTTGGTGAAGGGT
 TTTCTGTCGCTCAGTGTGTTATTGTAATTAAATTCTTGTGAGGCTCTGT
 TTAGCAGGTCGTCCTTCAGAACAGGACACAAAAAGATTAAATTATTAAAAAAA
 AAAAAAAAAAAGGGGAAATCGATATCAAGCTTATCGACACTGCAAGATCGTCAAAC
 ATTGCGAATAAGTTCTAAGATTGAATCCTGTTGCGGTCTTGCGATGATTATC
 ATATAATTCTGTTGAATTACGTTAAGCATGTAATAATTAAACATGTAATGATGACG
 TTATTATGAGATGGGTTTATGATTAGAGTCCCGCAATTATAACATTAAACCG
 ATAGAAAACAAATATAGCGCGAAACTAGGATAAAATTATCGCGCGGGTGT
 ATGTTACTAGAT

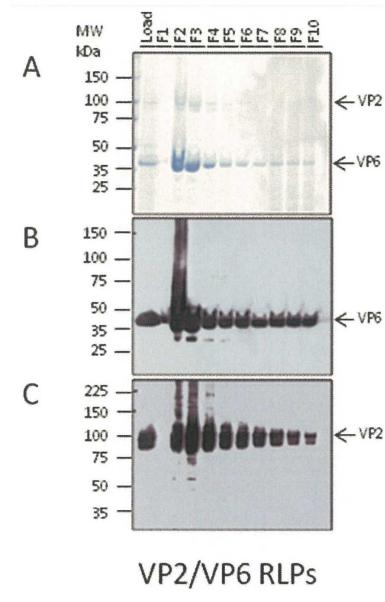
도 44E, SEQ ID NO: 59

로타바이러스 A 백신 USA/Rotarix-A41CB052A/1988/G1P1A[8] 균주의
 TrSp-VP7의 아미노산 서열
 MDYIIYRSLLIVVALFALTRAQNYGLNLPIIGSMDTVYANSTQEGIFLTSTLCLYY
 PTEASTQINDGEWKDSLQMFLTKWGPVSVYFKEYSSIVDFSVDPQLYCDYNLVL
 MKYDQNLELDMSLADLILNEWLCNPMDITLYQQSGESNKWISMSSCTVKVCLNT
 QMLGIGCQTTNVDSFEMVAEENEKLAIVDVVDGINHKINLTTTCTIRNCKKLGPREN
 VAVIQVGGSNVLDITADPTTNPQTERMMRVNWKKWWQVFYTIVDYINQIVQVMSKR
 RSLNSAAFYYRV

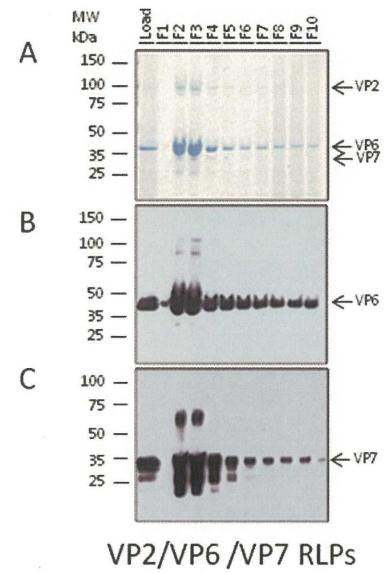
도면44d

도 44F
구조 번호 1734의 개략도

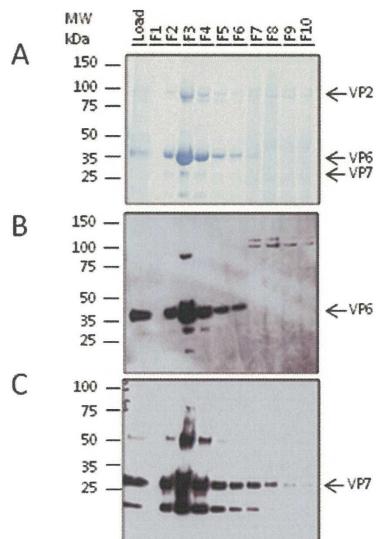
도면45



도면46

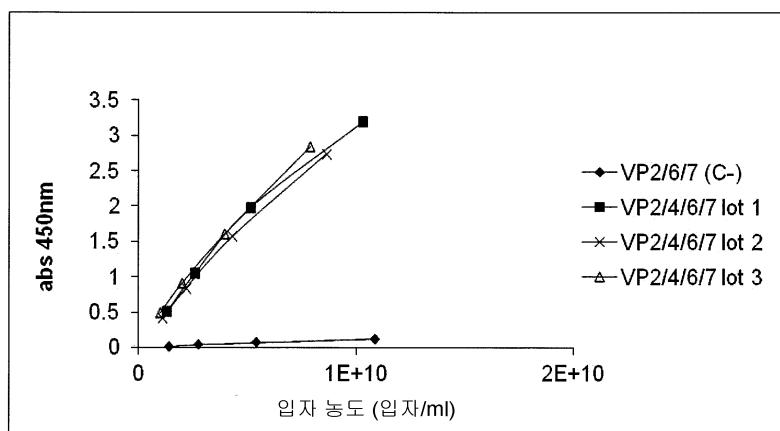


도면47

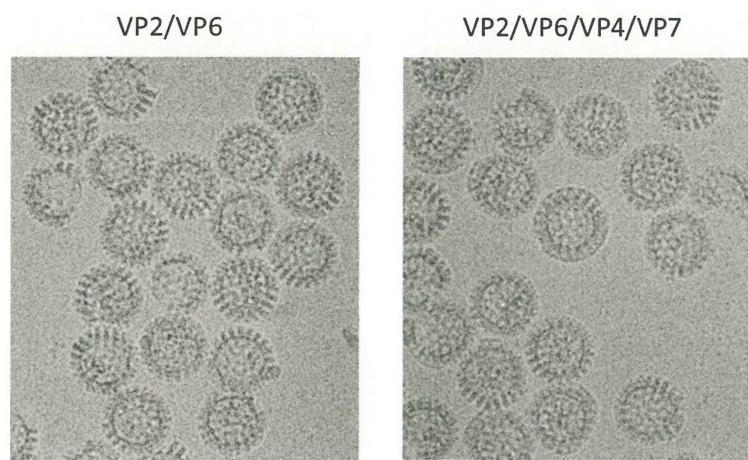


VP2/VP4/VP6 /VP7 RLPs

도면48



도면49



서 열 목록

SEQUENCE LISTING

<110> Medicago Inc.

Mitsubishi Tanabe Pharma Corporation

<120> Rotavirus-like Particle Production in Plants

<130> V84653WO

<150> US 61/646,058

<151> 2012-05-11

<160> 59

<170> PatentIn version 3.5

<210> 1

<211> 884

<212> PRT

<213> Nicotiana tabacum

<400> 1

Gly Arg Val Arg Ser Met Ala Tyr Arg Lys Arg Gly Ala Arg Arg Glu

1 5 10 15

Ala Asn Leu Asn Asn Asp Arg Met Gln Glu Lys Ile Asp Glu Lys

20 25 30

Gln Asp Ser Asn Lys Ile Gln Leu Ser Asp Lys Val Leu Ser Lys Lys

35 40 45

Glu Glu Ile Val Thr Asp Ser His Glu Glu Val Lys Val Thr Asp Glu

50 55 60

Leu Lys Lys Ser Thr Lys Glu Glu Ser Lys Gln Leu Leu Glu Val Leu

65 70 75 80

Lys Thr Lys Glu Glu His Gln Lys Glu Ile Gln Tyr Glu Ile Leu Gln

85 90 95

Lys Thr Ile Pro Thr Phe Glu Pro Lys Glu Thr Ile Leu Arg Lys Leu

100 105 110

Glu Asp Ile Gln Pro Glu Leu Ala Lys Lys Gln Thr Lys Leu Phe Arg

115 120 125

Ile Phe Glu Pro Lys Gln Leu Pro Ile Tyr Arg Ala Asn Gly Glu Arg

130 135 140

Glu Leu Arg Asn Arg Trp Tyr Trp Lys Leu Lys Lys Asp Thr Leu Pro

145 150 155 160

Asp Gly Asp Tyr Asp Val Arg Glu Tyr Phe Leu Asn Leu Tyr Asp Gln

165 170 175

Val Leu Thr Glu Met Pro Asp Tyr Leu Leu Leu Lys Asp Met Ala Val

180 185 190

Glu Asn Lys Asn Ser Arg Asp Ala Gly Lys Val Val Asp Ser Glu Thr

195 200 205

Ala Ser Ile Cys Asp Ala Ile Phe Gln Asp Glu Glu Thr Glu Gly Ala

210 215 220

Val Arg Arg Phe Ile Ala Glu Met Arg Gln Arg Val Gln Ala Asp Arg

225 230 235 240

Asn Val Val Asn Tyr Pro Ser Ile Leu His Pro Ile Asp Tyr Ala Phe

245 250 255

Asn Glu Tyr Phe Leu Gln His Gln Leu Val Glu Pro Leu Asn Asn Asp

260 265 270

Ile Ile Phe Asn Tyr Ile Pro Glu Arg Ile Arg Asn Asp Val Asn Tyr

275 280 285

Ile Leu Asn Met Asp Arg Asn Leu Pro Ser Thr Ala Arg Tyr Ile Arg

290 295 300

Pro Asn Leu Leu Gln Asp Arg Leu Asn Leu His Asp Asn Phe Glu Ser

305 310 315 320

Leu Trp Asp Thr Ile Thr Thr Ser Asn Tyr Ile Leu Ala Arg Ser Val

325 330 335

Val Pro Asp Leu Lys Glu Leu Val Ser Thr Glu Ala Gln Ile Gln Lys

340 345 350

Met Ser Gln Asp Leu Gln Leu Glu Ala Leu Thr Ile Gln Ser Glu Thr

355 360 365

Gln Phe Leu Thr Gly Ile Asn Ser Gln Ala Ala Asn Asp Cys Phe Lys

370 375 380

Thr Leu Ile Ala Ala Met Leu Ser Gln Arg Thr Met Ser Leu Asp Phe

385	390	395	400
Val Thr Thr Asn Tyr Met Ser Leu Ile Ser Gly Met Trp Leu Leu Thr			
405	410	415	
Val Val Pro Asn Asp Met Phe Ile Arg Glu Ser Leu Val Ala Cys Gln			
420	425	430	
Leu Ala Ile Val Asn Thr Ile Ile Tyr Pro Ala Phe Gly Met Gln Arg			
435	440	445	
Met His Tyr Arg Asn Gly Asp Pro Gln Thr Pro Phe Gln Ile Ala Glu			
450	455	460	
Gln Gln Ile Gln Asn Phe Gln Val Ala Asn Trp Leu His Phe Val Asn			
465	470	475	480
Asn Asn Gln Phe Arg Gln Ala Val Ile Asp Gly Val Leu Asn Gln Val			
485	490	495	
Leu Asn Asp Asn Ile Arg Asn Gly His Val Ile Asn Gln Leu Met Glu			
500	505	510	
Ala Leu Met Gln Leu Ser Arg Gln Gln Phe Pro Thr Met Pro Ile Asp			
515	520	525	
Tyr Lys Arg Ser Ile Gln Arg Gly Ile Leu Leu Leu Ser Asn Arg Leu			
530	535	540	
Gly Gln Leu Val Asp Leu Thr Arg Leu Leu Ala Tyr Asn Tyr Glu Thr			
545	550	555	560
Leu Met Ala Cys Ile Thr Met Asn Met Gln His Val Gln Thr Leu Thr			
565	570	575	
Thr Glu Lys Leu Gln Leu Thr Ser Val Thr Ser Leu Cys Met Leu Ile			
580	585	590	
Gly Asn Ala Thr Val Ile Pro Ser Pro Gln Thr Leu Phe His Tyr Tyr			
595	600	605	
Asn Val Asn Val Asn Phe His Ser Asn Tyr Asn Glu Arg Ile Asn Asp			
610	615	620	
Ala Val Ala Ile Ile Thr Ala Ala Asn Arg Leu Asn Leu Tyr Gln Lys			
625	630	635	640

Lys Met Lys Ala Ile Val Glu Asp Phe Leu Lys Arg Leu Tyr Ile Phe
 645 650 655
 Asp Val Ser Arg Val Pro Asp Asp Gln Met Tyr Arg Leu Arg Asp Arg
 660 665 670
 Leu Arg Leu Leu Pro Val Glu Ile Arg Arg Leu Asp Ile Phe Asn Leu
 675 680 685
 Ile Leu Met Asn Met Asp Gln Ile Glu Arg Ala Ser Asp Lys Ile Ala
 690 695 700
 Gln Gly Val Ile Ile Ala Tyr Arg Asp Met His Leu Glu Arg Asp Glu
 705 710 715 720
 Met Tyr Gly Tyr Val Asn Ile Ala Arg Asn Leu Glu Gly Phe Gln Gln
 725 730 735
 Ile Asn Leu Glu Glu Leu Met Arg Ser Gly Asp Tyr Ala Gln Ile Thr
 740 745 750
 Asn Met Leu Leu Asn Asn Gln Pro Val Ala Leu Val Gly Ala Leu Pro
 755 760 765
 Phe Ile Thr Asp Ser Ser Val Ile Ser Leu Ile Ala Lys Leu Asp Ala
 770 775 780
 Thr Val Phe Ala Gln Ile Val Lys Leu Arg Lys Val Asp Thr Leu Lys
 785 790 795 800
 Pro Ile Leu Tyr Lys Ile Asn Ser Asp Ser Asn Asp Phe Tyr Leu Val
 805 810 815
 Ala Asn Tyr Asp Trp Val Pro Thr Ser Thr Thr Lys Val Tyr Lys Gln
 820 825 830
 Val Pro Gln Gln Phe Asp Phe Arg Asn Ser Met His Met Leu Thr Ser
 835 840 845
 Asn Leu Thr Phe Thr Val Tyr Ser Asp Leu Leu Ala Phe Val Ser Ala
 850 855 860
 Asp Thr Val Glu Pro Ile Asn Ala Val Ala Phe Asp Asn Met Arg Ile
 865 870 875 880
 Met Asn Glu Leu

<210> 2

<211> 380

<212> PRT

<213> Nicotiana tabacum

<400> 2

Gly Arg Val Arg Ser Met Ala Ser Leu Ile Tyr Arg Gln Leu Leu Thr

1 5 10 15

Asn Ser Tyr Thr Val Glu Leu Ser Asp Glu Ile Asn Thr Ile Gly Ser

20 25 30

Glu Lys Ser Gln Asn Val Thr Ile Asn Pro Gly Pro Phe Ala Gln Thr

35 40 45

Asn Tyr Ala Pro Val Thr Trp Ser His Gly Glu Val Asn Asp Ser Thr

50 55 60

Thr Ile Glu Pro Val Leu Asp Gly Pro Tyr Gln Pro Thr Asn Phe Lys

65 70 75 80

Pro Pro Asn Asp Tyr Trp Ile Leu Leu Asn Pro Thr Asn Gln Gln Val

85 90 95

Val Leu Glu Gly Thr Asn Lys Thr Asp Ile Trp Val Ala Leu Leu Leu

100 105 110

Val Glu Pro Asn Val Thr Asn Gln Ser Arg Gln Tyr Thr Leu Phe Gly

115 120 125

Glu Thr Lys Gln Ile Thr Val Glu Leu Pro Thr Asp Phe Ser Val Ser

130 135 140

Arg Tyr Glu Val Ile Lys Glu Asn Ser Tyr Val Tyr Val Asp Tyr Trp

145 150 155 160

Asp Asp Ser Gln Ala Phe Arg Asn Met Val Tyr Val Arg Ser Leu Ala

165 170 175

Ala Asn Leu Asn Ser Val Lys Cys Ser Gly Gly Asn Tyr Asn Phe Gln

180 185 190

Ile Pro Val Gly Ala Trp Pro Val Met Ser Gly Gly Ala Val Ser Leu

195 200 205

His Phe Ala Gly Val Thr Leu Ser Thr Gln Phe Thr Asp Phe Val Ser
 210 215 220

Leu Asn Ser Leu Arg Phe Arg Phe Ser Leu Thr Val Glu Glu Pro Pro
 225 230 235 240
 Phe Ser Ile Leu Arg Thr Arg Val Ser Gly Leu Tyr Gly Leu Pro Ala
 245 250 255
 Phe Asn Pro Asn Asn Gly His Glu Tyr Tyr Glu Ile Ala Gly Arg Phe
 260 265 270
 Ser Leu Ile Ser Leu Val Pro Ser Asn Asp Asp Tyr Gln Thr Pro Ile
 275 280 285

Met Asn Ser Val Thr Val Arg Gln Asp Leu Glu Arg Gln Leu Gly Asp
 290 295 300
 Leu Arg Glu Glu Phe Asn Ser Leu Ser Gln Glu Ile Ala Met Thr Gln
 305 310 315 320
 Leu Ile Asp Leu Ala Leu Leu Pro Leu Asp Met Phe Ser Met Phe Ser
 325 330 335
 Asn Tyr Gly Ile Thr Arg Ser Gln Ala Leu Asp Leu Ile Arg Ser Asp
 340 345 350

Pro Arg Val Leu Arg Asp Phe Ile Asn Gln Asn Asn Pro Ile Ile Lys
 355 360 365

Asn Arg Ile Glu Gln Leu Ile Leu Gln Cys Arg Leu

370 375 380

<210> 3

<211> 402

<212> PRT

<213> Nicotiana tabacum

<400> 3

Gly Arg Val Arg Ser Met Asp Val Leu Tyr Ser Leu Ser Lys Thr Leu

1 5 10 15

Lys Asp Ala Arg Asp Lys Ile Val Glu Gly Thr Leu Tyr Ser Asn Val

20 25 30

Ser Asp Leu Ile Gln Gln Phe Asn Gln Met Ile Ile Thr Met Asn Gly
 35 40 45
 Asn Glu Phe Gln Thr Gly Gly Ile Gly Asn Leu Pro Ile Arg Asn Trp
 50 55 60
 Asn Phe Asp Phe Gly Leu Leu Gly Thr Thr Leu Leu Asn Leu Asp Ala
 65 70 75 80
 Asn Tyr Val Glu Thr Ala Arg Asn Thr Ile Asp Tyr Phe Val Asp Phe
 85 90 95

 Val Asp Asn Val Cys Met Asp Glu Met Val Arg Glu Ser Gln Arg Asn
 100 105 110
 Gly Ile Ala Pro Gln Ser Asp Ser Leu Arg Lys Leu Ser Gly Ile Lys
 115 120 125
 Phe Lys Arg Ile Asn Phe Asp Asn Ser Ser Glu Tyr Ile Glu Asn Trp
 130 135 140
 Asn Leu Gln Asn Arg Arg Gln Arg Thr Gly Phe Thr Phe His Lys Pro
 145 150 155 160

 Asn Ile Phe Pro Tyr Ser Ala Ser Phe Thr Leu Asn Arg Ser Gln Pro
 165 170 175
 Ala His Asp Asn Leu Met Gly Thr Met Trp Leu Asn Ala Gly Ser Glu
 180 185 190
 Ile Gln Val Ala Gly Phe Asp Tyr Ser Cys Ala Ile Asn Ala Pro Ala
 195 200 205
 Asn Thr Gln Gln Phe Glu His Ile Val Gln Leu Arg Arg Val Leu Thr
 210 215 220

 Thr Ala Thr Ile Thr Leu Leu Pro Asp Ala Glu Arg Phe Ser Phe Pro
 225 230 235 240
 Arg Val Ile Asn Ser Ala Asp Gly Ala Thr Thr Trp Tyr Phe Asn Pro
 245 250 255
 Val Ile Leu Arg Pro Asn Asn Val Glu Val Glu Phe Leu Leu Asn Gly
 260 265 270
 Gln Ile Ile Asn Thr Tyr Gln Ala Arg Phe Gly Thr Ile Val Ala Arg

275

280

285

Asn Phe Asp Thr Ile Arg Leu Ser Phe Gln Leu Met Arg Pro Pro Asn

290

295

300

Met Thr Pro Ser Val Ala Ala Leu Phe Pro Asn Ala Gln Pro Phe Glu

305 310 315 320

His His Ala Thr Val Gly Leu Thr Leu Lys Ile Glu Ser Ala Val Cys

325

330

335

Glu Ser Val Leu Ala Asp Ala Ser Glu Thr Met Leu Ala Asn Val Thr

340

345

350

Ser Val Arg Gln Glu Tyr Ala Ile Pro Val Gly Pro Val Phe Pro Pro

355

360

365

Gly Met Asn Trp Thr Asp Leu Ile Thr Asn Tyr Ser Pro Ser Arg Glu

370

375

380

Asp Asn Leu Gln Arg Val Phe Thr Val Ala Ser Ile Arg Ser Met Leu

385 390 395 400

Val Lys

<210> 4

<211> 331

<212> PRT

<213> Nicotiana tabacum

<400> 4

Gly Arg Val Arg Cys Met Tyr Gly Ile Glu Tyr Thr Thr Ile Leu Thr

1 5 10 15
Phe Leu Ile Ser Ile Val Leu Leu Asn Tyr Ile Leu Lys Ser Leu Thr

20

25

30

Ser Ala Met Asp Phe Ile Ile Tyr Arg Phe Leu Leu Ile Val Ile

35

40

45

Ala Ser Pro Phe Val Lys Thr Gln Asn Tyr Gly Ile Asn Leu Pro Ile

50

55

60

Thr Gly Ser Met Asp Thr Ala Tyr Ala Asn Ser Ser Gln Gln Glu Thr

65	70	75	80
Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ser Thr			
85	90	95	
Gln Ile Gly Asp Thr Glu Trp Lys Asp Thr Leu Ser Gln Leu Phe Leu			
100	105	110	
Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Thr Asp			
115	120	125	
Ile Ala Ser Phe Ser Ile Asp Pro Gln Leu Tyr Cys Asn Tyr Asn Val			
130	135	140	
Val Leu Met Lys Tyr Asp Ser Thr Leu Glu Leu Asp Met Ser Glu Leu			
145	150	155	160
Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr			
165	170	175	
Leu Tyr Tyr Tyr Gln Gln Thr Asp Glu Ala Asn Lys Trp Ile Ser Met			
180	185	190	
Gly Gln Ser Cys Thr Ile Lys Val Cys Pro Leu Asn Thr Gln Thr Leu			
195	200	205	
Gly Ile Gly Cys Ile Thr Thr Asn Thr Ala Thr Phe Glu Glu Val Ala			
210	215	220	
Thr Ser Glu Lys Leu Val Ile Thr Asp Val Val Asp Gly Val Asn His			
225	230	235	240
Lys Leu Asp Val Thr Thr Asn Thr Cys Thr Ile Arg Asn Cys Lys Lys			
245	250	255	
Leu Gly Pro Arg Glu Asn Val Ala Ile Ile Gln Val Gly Ser Asp			
260	265	270	
Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Ala Pro Gln Thr Glu Arg			
275	280	285	
Met Met Arg Val Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Val			
290	295	300	
Val Asp Tyr Ile Asn Gln Ile Val Gln Val Met Ser Lys Arg Ser Arg			
305	310	315	320
Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Val			

325

330

<210> 5		
<211> 22		
<212> DNA		
<213> Artificial sequence		
<220><223> Primer VP2F		
<400> 5		
ttccatggct taccgtaaaa gg	22	
<210> 6		
<211> 32		
<212> DNA		
<213> Artificial sequence		
<220><223> Primer VP2R		
<400> 6		
atgcggccgc aagctcggtc ataatcctca tg	32	
<210> 7		
<211> 22		
<212> DNA		
<213> Artificial Sequence		
<220><223> Primer VP4F		
<400> 7		
ttccatggct tccctcatct ac	22	
<210> 8		
<211> 31		
<212> DNA		
<213> Artificial sequence		
<220><223> Primer VP4R		
<400> 8		
atgcggccgc aagacggcac tggagaatga g	31	
<210> 9		
<211> 21		
<212> DNA		
<213> Artificial sequence		

<220><223> Primer VP6F
<400> 9
ttccatggat gtgctctact c 21
<210> 10
<211> 28
<212> DNA
<213> Artificial sequence
<220><223> Primer VP6R
<400> 10
atgcggccgc cttcacgagc atggaacg 28
<210> 11
<211> 22
<212> DNA
<213> Artificial sequence
<220><223> Primer VP7F
<400> 11
gtacatgtac ggaatcgagt ac 22
<210> 12
<211> 31
<212> DNA
<213> Artificial sequence
<220><223> Primer VP7R
<400> 12
atgcggccgc cacacggtag tagaaaggcag c 31
<210> 13
<211> 2700
<212> DNA
<213> Nicotiana tabacum
<400> 13
ggtaccgaat tcggacgcgt tcgttccatg gcttaccgta aaaggggtgc taggcgtcaa 60
gctaacctca acaaacaacgta taggatgcaa gagaagatcg atgagaagca ggattccaac 120
aagatccagc tctccgataa ggtgctctcc aagaaagaag agatcgttac tgattcccac 180
gaagagggtga aggtgacaga tgagcttaag aagtccacaa aagaagagtcc caagcagctc 240

ctttaggtgc tcaagacaaa agaggaacac cagaaagaga tccagtacga gatcctcaa	300
aagactatcc caacttcgaa gccaaaagag actatccta ggaagcttga ggatatccag	360
ccagagcttgc taagaagca gactaagctc ttcaaggatct tcgagccaaa gcagctcca	420
atctaccgtgc taacacgtga aaggaaacctt aggaacagggt ggtactggaa gctcaagaag	480
gataactctcc cagacgggtga ttacgatgtg agagagttact tcctcaacct ctacgatcag	540
gtgctcaactg agatgccaga ttacctcctc ctcaaggata tggctgtgga gaacaagaac	600
tccagggatgtcgaaaggt ggtggattcc gagactgctt ccatctgtga tgctatcttc	660
caggatgaag agactgaggg tgctgtgagg cgtttcattg ctgagatgag gcagagggtt	720
caggctgata ggaacgtggtaaactaccca tccatcctcc acccaatcga ttacgcttc	780
aacgagttact tccttcagca ccagcttgc gagccactca acaacgatata catctcaac	840
tacatcccag agaggattag gaacgacgtt aactacatcc tcaacatgga taggaacctc	900
ccatccactg ctcgttacat caggccaaac ctcctccagg ataggctcaa cctccacgt	960
aacttcgagt ccctctggaa tacaatcaact acattccaact acattctgc tcgttccgt	1020
gtgccagatctcaagaact cgtgtccact gaggctcaga tccagaagat gtcccaggat	1080
ctccagtttgc aggtctcac tatccagtcc gagactcagt tcctcaactgg tatcaactcc	1140
caggctgcta acgattgctt caagactctc attgtgtcta tgctctccca gaggactatg	1200
tccctcgatt tcgtgactac taactatatg tccctcatct ccgaaatgtg gctttgact	1260
gtggtgccaa acgatatgtt catccgttag tcccttggtt ctggccagct cgctatcgt	1320
aacactatca tctacccagc ttccgaaatg caaaggatgc actaccgtaa cggtgatcca	1380
cagactccat tccagatcgc agagcagcag atccagaact tccagggtggc aaactggctc	1440
cacttcgtga acaacaacca gttcaggcag gctgtgatcg atggtgtgtt gaaccagggt	1500
ctcaacgata acatccgtaa cggtcacgtt atcaaccagc tcatgaaagc tctcatgca	1560
ctctccagggc agcagttccc aactatgcct atcgattaca agcggtccat ccagagggg	1620
atccctctcc ttccaaacag gttggacag ctctggatc tcacttaggtt cctcgcttac	1680
aactacgaga ctctcatggc ttgcataact atgaacatgc agcacgttca gactctact	1740
actgagaagc tccagctcac ttccgtgact tccctctgca tgctcatcgaa aacgctact	1800
gtgatcccat cccccacagac actcttccac tactacaacg tgaacgtgaa cttccactcc	1860
aactacaacg agaggatcaa cgtatgttg gctatcatca ctgctgtaa caggcttaac	1920
ctctaccaaa agaagatgaa ggctatcggtt gaggattcc tcaagaggctt ctatcttc	1980
gatgtgtccaa ggggtccaga tgatcagatg taccgtctt gggataggct taggctctc	2040
ccagtggtggaa tcagaaggctt cgtatcttc aacctcatcc ttatgaacat ggatcagatc	2100


```
ttgaagctca gggagaccct atgttagtga tgaagggtga tgtaagagcg agcaaggcac 1020  
cacggcttag agtttcttga gcacagggtga ctccgagtct aggtcttcta cagggtccata 1080
```

gagggtcgaaac tccgagggatg atagggtcagg ctctgagtca aggagtgacc atagttgagg	1140
gtccgcacgt tgctaaca gttctgagag taacgcacgt acgagagggt ctccctgatac	1200
agggagctaa agcactgtatg attgtatatac agggagtaga ggccttacac cgagaactga	1260
caccacggtt tgctatacaa gtaggcactc agggAACACC gaacggtcga gcgtatgcac	1320
tttgtatagt agatgggtcg aaaggcattac gttcctacg tgatggcatt gccacttaggt	1380
gtctgaggta aggtctagcg tctcgctgc taggtcttga aggtccaccc ttggaccgag	1440
gtgaaggact tggtgttggt caagtccgtc cgacactagc taccacacaa ctgggtccac	1500

gagttgctat tgtaggcatt gccagtgcac tagttggtcg agtacctcg agagtagttt	1560
gagaggtccg tcgtcaaggg ttgatacgga tagctaattgt tcgcaaggta ggtctcccct	1620
taggaggagg aaaggttgtc cgaacctgtc gagcacctag agtgatccga ggagcgaatg	1680
ttgtatgtctt gagagtaccg aacgtgttga tacttgtacg tcgtgcgtt ctgagatgttt	1740
tgactcttcg aggtcgatgt aaggcactgtaa agggagacgt acgagtagcc tttgcgtatgttt	1800
cactagggtatgggtgtctgtg tgagaagggtatgttgc acttgcactt gaagggtgagg	1860
ttgtatgttgc tctccatgtt gctacgacac cgatgtgttgc gacgacgtt gtccgtatgttt	1920

gagatggttt tcttctactt ccgatagcaa ctccctaaagg agttctccga gatgtagaag	1980
ctacacaggc cccacggctt actagtctac atggcagaat ccctatccga atccgaggag	2040
ggtcacacctt agtcttccga gctatagaag ttggagtagg aataacttgta cctagtctag	2100
ctctcccgaa ggctattcta gcgcgtcccc caataatgc gaatggcact atacgtggaa	2160
ctctccctac tctacatgcc tatgcacttg taacgatccc ttggaaactccc taaggcgttc	2220
tagttggAAC ttctcgAAATA cgcaaggccA ctaatgcgAG tctagtgtt gtacgaggAG	2280
tttgttggTCG gtcaccggAGA acaaccacgA gagggtaagt agtgcataAG gaggcactAG	2340

agggagtaac gattcaacct acgatgacac aagcgagtct agcacttcga gtccttcac	2400
ctatgagagt tcggtagga gatttctag ttgaggctaa gggtgctaaa gatggagcac	2460
cgattgtatgc taacccacgg ttgaaggtga tgttccaca tgttcggtcca cggtgtcgtc	2520
aagctaaagg cattgaggtt cgttacgag tgaaggttgg agtgaaatgtt acacatgagg	2580
ctagaggaggc gaaagcacag gcgttatgtt cacctcgat agttgcgaca ccgaaagcta	2640
ttgtacttctt aatacttgcgat cgaaactact gagctccctt ggagatctctt taagctcgat	2700

<210> 15

<211> 2388

<212> DNA

<213> Rotavirus

<400> 15

ggtaccgaat tcggacgcgt tcgttccatg gcttcctca tctaccgtca gttgctact	60
aactcctaca ctgtggagct ttccgatgag atcaacacta tcggttccga gaagtcccag	120
aacgtgacta tcaacccagg accattcgct cagactaact acgctccagt gacttggta	180
cacggtaaga tgaacgattc cactactatc gagccagtg tcgatggacc ataccagcca	240
actaacttca agccaccaaa cgattactgg attctctca accaactaa ccagcaggtg	300
gtgcttggagg gaactaacaa gactgatatc tgggtggcac tccttcttgt ggagccaaac	360
gtgactaacc agtccaggca gtacactctc ttccggagaga ctaagcagat cacttggag	420
aacaacacta acaagtggaa gttttcgag atgtttagt ccaacgtgaa cgctgagttc	480
cagcacaaga ggactctcac ttccgataca aagctcgctg gtttatgaa gttctacaac	540
tctgtgtgga cttccacgg tgaaactcca cacgtacta ctgattactc ctccacttcc	600
aaccttccg aggtggagac tgtgatccac gtggagttct acatcatccc aaggtccaa	660
gagtctaagt gctccgagta catcaacact ggactccac caatgcaaaa cactaggaac	720
atcgtgccag tggcttgct ctctcggttcc gtgacttacc agagggtca ggtgaacgag	780
gatatcatca tctccaagac ttccctctgg aaagagatgc agtacaacag ggatattatc	840
atcaggttca agtcaacaa ctccatcgta aagctcgag gactcgata caagtggagt	900
gagatctctc tcaaggctgc taactaccag tactctacc tcagggatgg tgaacaggtg	960
acagctcaca ctacttgctc cgtgaacgggt gttaacaact tctcctacaa cgggtggttcc	1020
ctcccaactg atttctccgt gtcccggttac gaggtgatca aagagaactc ctacgtttac	1080
gtggattact gggatgattc ccaggcttc aggaacatgg tgtatgttag atccctcgct	1140
gctaacctca actccgtgaa gtgctccgtt ggaaactaca actccagat cccagtgaa	1200
gcttggccag ttagtctgg tggagctgtc tctctccact tcgctgggt tacactctcc	1260
actcagttca ctgatttcgt gtccctcaac tccctcaggat tcaggttctc cctcactgtg	1320
gaagagccac cattctccat cctcaggact aggggtgtccg gactttacgg actcccgatct	1380
ttcaacccaa acaacggaca cgagtaactac gagatcgctg gacgtttctc ctttatctcc	1440
ctcgccat ccaacgatga ttaccagact ccaattatga actccgtgac tgtgaggcag	1500
gatcttgaga ggcagctgg agatctcagg gaagagttca actccctctc ccaagagatc	1560
gctatgactc agtctcatgaa tctcgcttc ctccctactcg atatgttctc catgttctct	1620

ggtatcaagt ccactatcga tgtggctaag tctatggta ctaaggtat gaagaagttc	1680
aagaagtccg gactcgctac ttccatctcc gagcttactg gatctcttc caacgctgct	1740
tcttctgtgt ctaggtcctc ctccatcagg tccaacatct cctccatctc agtgtggact	1800
gatgtgtccg agcagatcgc tggatctcc gattccgtgc gtaacatctc cactcagact	1860
tccgctatct ccaagaggct taggctcaga gagatcacta ctcagactga ggaaatgaac	1920
ttcgatgata tctccgctgc tgtgctcaag actaagatcg ataggtccac tcacatctcc	1980
ccagatactc tcccagatat catcaactgag tcctccgaga agttcatccc aaagcgtgct	2040
taccgtgttc tcaaggatga tgaggtatg gaagctgatg tggatggaaa gttttcgct	2100
tacaaagtgg gaacttcga agaggtgcc a ttcgatgtgg ataagttcgt ggtatctcg	2160
actgattccc cagtgtatctc cgcttatcatc gattcaaga ctctcaagaa cctcaacat	2220
aactacggaa tcacttagtgc ccaggctctc gatctcatcc gtccgatcc aagggtgctc	2280
agggatttca tcaaccagaa caacccaatc atcaagaaca ggatcgagca gctcattctc	2340
cagtggcgcc tttagtact cgagggatcc tctagagaat tcgagctc	2388
<210> 16	
<211> 2388	
<212> DNA	
<213> Rotavirus	
<400> 16	
ccatggctta agcctgcgca agcaaggtac cgaagggagt agatggcagt caacgagtga	60
ttgaggatgt gacacctcga aaggctactc tagttgtat agccaaggct cttcagggtc	120
ttgcactgtat agttgggtcc tggtaagcga gtctgattga tgcgagggtca ctgaaccagt	180
gtgccacttc acttgctaag gtgatgatag ctggtcacg agtacctgg tatggtcgt	240
tgattgaagt tcgggtgtt gtaatgacc taagaggagt tgggttgatt ggtcgccac	300
cacgaactcc ctgtattttt ctgactatag acccaccgtg aggaagaaca cctcggttt	360
cactgattgg tcaggccgt catgtgagag aagcctctt gattcgatca gtgacaccc	420
ttgttgtat tttcacctt caagaagctc tacaagtcca gttgcactt gcgactcaag	480
gtcggttct cctgagagtg aaggctatgt ttgcagcgac caaaatactt caagatgtt	540
agacacaccc gaaagggtgcc acatttgggt gtgcgtatgat gactaatgag gaggtgaagg	600
ttggaaaggc tccacctctg acactagggt cacctcaaga tggatgttgg ttccagggtt	660
ctcagattca cgaggctcat gtagttgtga cctgagggtg gttacgtttt gtgatcttg	720
tagcacggtc accgaaacag gagagcaagg cactgaatgg tctcccgagt ccacttgctc	780

ctatagtagt agagggtctg aagggagacc tttctctacg tcatgttgtc cctataatag	840
tagtccaagt tcaagttgtt gaggtgcac ttcgagcctc ctgagccat gttcacctca	900
ctctagagga agttccgacg attgatggc atgaggatgg agtccctacc acttgtccac	960
tgtcgaggt gtgaacgag gcacttgcca caattgtga agaggatgtt gccaccaagg	1020
gagggttgac taaagaggca cagggcaatg ctccactagt ttctcttgag gatgcaaatg	1080
cacctaata gccctactaag ggtccgaaag tccttgatcc acataacaatc tagggagcga	1140
cgattggagt tgaggcactt cacgaggcca ccttgatgt tgaaggctta gggtcaccct	1200
cgaaccggc actacagacc acctcgac acagaggtga agcgaccaca atgtgagagg	1260
tgagtcaagt gactaaagca cagggagttt agggagtcca agtccaagag ggagtgcac	1320
cttctcggt gtaagaggtt ggagtctga tcccacaggc ctgaaatgcc tgagggtcga	1380
aagttgggtt tggctgt gctcatgtt ctctagcgac ctgcaaagag ggaatagagg	1440
gagcacggta ggttgctact aatggtctga ggttaataact tgaggcactg acactccgtc	1500
ctagaactct ccgtcgagcc tctagagtcc ctctcaagt tgagggagag ggttctctag	1560
cgatactgag tcgagtagct agagcgagag gagggtgagc tatacaagag gtacaagaga	1620
ccatagttca ggtgatagct acaccgattt agataccact gattccacta ctcttcaag	1680
ttcttcaggc ctgagcgatg aagtagagg ctgcatttgc ctagagagag gttgcac	1740
agaagacaca gatccaggag gaggtgtcc aggttgtaga ggaggttagag tcacacctga	1800
ctacacaggc tcgtctagcg acctagaagg ctaaggcacg cattgttagag gtgagtctga	1860
aggcgataga ggttctccga atccgatct ctctagtgat gagtctgact ccctacttg	1920
aagctactat agaggcgacg acacgatcc tgattctagc tatccaggtg agttagagg	1980
ggtctatgag agggtctata gtatgtactt aggaggctt tcaagtaggg ttgcac	2040
atggcacaag agttcctact actccactac ctgcactac acctacctt caagaagcga	2100
atgtttcacc ctgaaagct tctccacggta aagctacacc tattcaagca cctagagcac	2160
tgactaaggc gtcactagag gcgatagtag ctaaagttct gagagtctt ggagttgcta	2220
tttgatgcctt agtgatccag ggtccgagag cttagatgg caaggctagg ttccccacgag	2280
tccctaaagt agttggctt gttgggttag tagttctgt cctagctgt cgagtaagag	2340
gtcacggcag aaactactga gctccctagg agatcttta agctcgag	2388
<210> 17	
<211> 1254	
<212> DNA	
<213> Nicotiana tabacum	

<400> 17

ggtaccgaat tcggacgcgt tcgttccatg gatgtctct actccctctc caagactctc	60
aaggatgcta gggataagat cgtggaggga actctctact ccaacgttc cgatctcatc	120

cagcagttca accagatgat catcaactatg aacggaaacg agttccagac tggtggaatc	180
gaaaaacctcc caatcaggaa ctggaacttc gattcggac tcctcgaaac tactctctc	240
aacctcgatg ctaactacgt ggagactgct aggaacacta tcgattactt cgttgatttc	300
gtggataatg tgtgcatttga ttagatggtt cgtgagtccc agaggaacgg aattgctcca	360
cagtcgatt ccctcaggaa gctctccgtt atcaagttca agaggatcaa cttcgataac	420
tcctccgagt acatcgagaa ctggaaccc cagaacagaa ggcagaggac tggatticact	480
ttccacaaggc caaacatctt cccatactcc gcttcattca ctctcaacag gtcccaagcca	540

gctcacgata acctcatggg aactatgtgg ctcaacgctg gttctgagat ccaggtggca	600
ggattcgatt actcctgcgc tatcaacgct ccagctaaaca ctctcgatg ttctcgatc	660
gttcagctca gaaggggtgt cactactgtt actatcactc tcctccaga tgctgagagg	720
ttctccctcc caaggggtat caactccgtt gatggtgcta ctacttgta cttcaaccca	780
gtgatcctca ggcacaaacaa cgtggagggtg gagttcccttc tcaacggaca gatcatcaac	840
acttaccagg ctcgttccgg aactatcggtt gcttaggaact tcgataacaat caggctctcc	900
ttccagctta tgaggccacc aaacatgact ccattccgtt ctgcactctt cccaaacgca	960

cagccattcg agcaccacgc tactgtggg ctcaactctca agatcgagtc cgctgtgtgc	1020
gagtcgtgtgc tcgtgtatgc ttccgagact atgctcgcta acgtgacttc tgtgaggca	1080
gagtagctca tcccagtggg accagtgttt ccaccaggaa tgaactggac tgatctcatc	1140
actaactact cccatccag agaggataac ctccagagggtt tggtactgt ggctccatc	1200
cgttccatgc tcgtgaagtg atgactcgag ggatcctcta gagaattcga gctc	1254

<210> 18

<211> 1254

<212> DNA

<213> Nicotiana tabacum

<400> 18

ccatggctta agcctgcgca agcaaggtac ctacacgaga tgagggagag gttctgagag	60
---	----

tccctacgt ccctattcta gcacccct tgagagatga ggttgcaaaag gctagagtag	120
gtcgtaagt tggctacta gtagtgatac ttgccttgc tcaaggctg accaccttag	180
cctttggagg gttagtcattt gacccgttgaag ctaaaggctg aggaggctt atgagaggag	240

ttggagctac gattgatgca cctctgacga tccttgtat agctaataaag 300
 cacctattac acacgtacct actctaccaa gcactcaggg ttccttgcc ttaacgagg 360
 gtcaaggctaa gggagtcctt cgagaggcca tagtcaagt ttccttagtt gaagctattg 420
 aggaggctca ttagtgcctt gaccttggag gtcttgtt ccgtctccig acctaagtga 480

aagggttgcg gttttagaa gggtatgagg cgaaggaaat gagagtgtc cagggtcggt 540
 cgagtgcata tggagtagcc ttgatacacc gagttgcac caagactcta ggtccaccgt 600
 cctaagctaa tgaggacgca atagttgcga ggtcgattgt gagtcgtcaa gctcgtag 660
 caagtcgagt ctcccacga gtatgacga tgatagttag agagggtct acgactctcc 720
 aagaggaagg gtccacta gttgaggcga ctaccacat gaigaaccat gaagtgg 780
 cactaggagt ccggttgtt gcacccac ctcaaggaag agttgcctgt ctatgttg 840
 tgaatggtcc gagcaaagcc ttgatagec acgtatgtt gtccgagg 900

aagggtcaat actccgtgg ttgtactga ggtggcacc gacgtgagaa gggtttcg 960
 gtcgtaagc tcgtggcgc atgacaccct gagtgagat tctagctcag ggcacacacg 1020
 ctcagacacg agcactacg aaggctctga tacgagcgt tgactgaag acactccgtt 1080
 ctcatgcgtt agggtcaccc tggcacaaa ggtggcctt acttgacctg actagatgt 1140
 tgattgtga gggtagtgc ttcctattt gagggtccc acaagtgaca ccgaaggtag 1200
 gcaaggtagc agcacttcac tactgagctc cctaggat ctcttaagct cgag 1254

<210> 19

<211> 1041

<212> DNA

<213> Nicotiana tabacum

<400> 19

ggtaccgaat tcggacgcgt tcgttgcatg tacggaatcg agtacactac tatccctact 60
 ttccatctt ccatcgtgt cctcaactac atcccaatg ccctcactc cgctatggat 120
 ttccatctt accgtttctt ctcctcata gtatcggtt cccattgtt taagactcag 180
 aactacggta tcaacccccc aatcaactggta tctatggata ctgcttacgc taactcc 240
 cagcaagaga ctccctcac ttccacactc tgcctctact acccaactga ggcacccaca 300
 cagatcgag atacagatgtt gaaggatact ctctccacg ttcctcac taaggatgg 360
 ccaactggtt ccgtgtactt caaagatgtc actgatatcg ctccctctc catcgatcca 420

cagctctact gcgattacaa cgtggtgctt atgaagtacg attccactt tgagttgtat 480
 atgtccgagc ttgtgtatct catcctcaac gagtggctt gcaacccaaat ggtatcact 540

ctctactact accagcagac tcatgaggct aacaagtgg a tctctatgg acagtccgc	600
actatcaaag tgtccccact caacactcg actctcgaa tcggatgcat cactactaac	660
actgtactt tcgaggaagt ggctactcc gagaagctcg tcatgactga tgtggtgat	720
gttgttaacc acaagctcg a tcatgactt aacacatgca caatcaggaa ctgcaagaag	780
ctcgacca gggaaaacgt ggctatcatc caagtggag gtccgatgt gctcgatata	840
actgctgatc caactactgc tccacagact gagaggatga tgagggtgaa ctggaagaag	900
tggggcagg tttctacac tcatgatggat tacatcaacc agatcggtca ggtgatgtcc	960
aagaggtccc gttctctaa ctccgctgt ttctactacc gtgtgtgatg actcgaggaa	1020
tcctcttagag aatcgagct c	1041
<210> 20	
<211> 1041	
<212> DNA	
<213> Nicotiana tabacum	
<400> 20	
ccatggctta agcctcgca agcaacgtac atgccttagc tcatgtgatg ataggagtga	60
aaggagtaga ggttagcacga ggagttgatg taggagttca gggagtgaag gcgataccta	120
aagtagtaga tggcaaagga ggaggagtag cactagcgaa gggtaagca attctgagtc	180
ttgtatgcat agttggaggg ttagtacact agataccat gacgaatgctt attgaggagg	240
gtcggtctct gaaaggagtg aaggtgtgag acggagatga tgggttact ccgttaggt	300
gtcttagcctc tatgtctcac cttcttatga gagagggtcg agaaggagtg attccctacc	360
ggttgaccaa ggcacatgaa gtttctcatg tgactatagc gaaggaagag gtagcttagt	420
gtcgagatga cgctaatgtt gcaccacgaa tacttcatgc taaggtgaga actcgaacta	480
tacaggctcg aacgactaga gtaggagttt ctcaccgaga cttgggtta cctatagtga	540
gagatgatga tggcgtctg actactccga ttgttacact agagataccc tgcaggacg	600
tgtatgttca acacgggtga gttgtgatc tgagagccctt agcctacgtt gtgtatgtt	660
tgacgatgaa agtccttca ccgtatgaaagg ctctcgagc actagtactt acaccaccta	720
ccacaattgg tttcgagctt acactgatga ttgtgtacgt gttgttactt gacgttcttc	780
gagcctgggtt ccctttgca ccgtatgatg gttcacccctc caaggctaca cgagctatag	840
tgacgactag gtatgtacg aggtgtctga ctctctactt actccactt gacgttcttc	900
accaccgtcc aaaagatgtg acaccaccta atgtatgtt gtttagcaagt ccactacagg	960
ttctccaggcaagagatg gaggcgacga aagatgatgg cacacactac tgacgtccct	1020

aggagatctc ttaagctcg a	1041
<210> 21	
<211> 48	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer IF-WA_VP2(opt).s1+3c	
<400> 21	
aaatttgcg gccccatggc ataccggaag agaggagcaa agcgcgaa	48
<210> 22	
<211> 50	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer IF-WA_VP2(opt).s1-4r	
<400> 22	
actaaagaaa ataggccttt aaagctcggt cattattcgc atattgtcg a	50
<210> 23	
<211> 4903	
<212> DNA	
<213> Artificial sequence	
<220><223> Construct 1191	
<400> 23	
tggcaggata tatttggtg taaacaatt gacgcttaga caacttaata acacattgcg	60
gacgtttta atgtactgaa ttaacgccga atccgggct ggtatattta tatgttgtca	120
aataactcaa aaaccataaa agttaagtt agcaagtgt tacattttt cttgaacaaa	180
aatattcacc tactactgtt ataaatcatt attaacatt agagtaaaga aatatggatg	240
ataagaacaa gagtagtgat atttgacaa caattttgtt gcaacatttg agaaaatttt	300
gttggctct ctttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata	360
aaaacataat gtgagtagtga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac	420
aaatatcatt gaggaatttg acaaaagcta cacaataag ggttaattgc tgtaataaa	480
taaggatgac gcatttagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga	540
aagaataat tattttaaa attaaaagtt gagtcatttg attaacatg tgattattta	600
atgaattgtat gaaagagttt gattaaagtt gtatttagtaa ttggcaattt gttcaattt	660
taatttgcata ttgtatctt tccttatata tgccccatag agtcaggtaa ctcattttta	720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg	780
gtatattac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata	840
acatccaatc caaccaatca caacaatcct gatgagataa cccacttaa gcccacgcat	900
ctgtggcaca tctacattat ctaatcaca cattttcca cacatctgag ccacacaaaa	960
accaatccac atcttatac cccattctat aaaaatcac acttttgag tctacactt	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggaac gagctataca aggaaacgac gctaggaaac aagctaacag tgaacgttgg	1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg	1200
actgagtgcc ggctacataa cgatgagacg aattcgaatc aagataatcc cttgggttc	1260
aaggaaagct ggggttcgg gaaagtgtt ttaagagat atctcagata cgacaggacg	1320
gaagcttcac tgacagagt cttggatct tggacggag attcggttaa ctatgcgca	1380
tctcgattt tcggttcga ccagatcggta tttttttttt gtattcggtt tcgaggagtt	1440
agtatcaccc ttctggagg gtcgcgaact cttagatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gttggccca atcgaagttgg aaagtaatgt atcaagagga	1560
tgcctgaag gtactcaaacc cttcgaaaaa gaaagcgagt aagttaaat gttttcggt	1620
ctcctattn taatatggtt tttttttttt aattttgttc ttgtagaaga gcttaattaa	1680
tcgttgttgt taatggaaatc tttttgtatg agatgaactg gtgtatgtt attcattac	1740
ataagtggag tcagaatcag aatgtttctt ccataactaa ctagacatga agacctgccc	1800
cgtacaattt tttttttttt gaacaactaa aattgaacat cttttgccac aactttataa	1860
gtggtaata tagctcaat atatggtcaa gttcaataga ttaataatgg aaatatcgt	1920
tatcgaaatt cattaacaat caacttaacg ttattaaacta ctaattttt atcatcccct	1980
ttgataatg atagtagacacc aatttaggaag gagcatgctc gcctaggaga ttgtcggttcc	2040
ccgccttcag ttgtcaagct gctctagccg tttttttttt acgcaaaacgg cctctccccg	2100
cgcgttggga attactagcg cgtgtcgaca agcttgcattt ccggtaaca tggggagca	2160
cgacacactt gtctactcca aaaatataa agatacagtc tcagaagacc aaaggcaat	2220
tgagactttt caacaaaggta aatatccgg aaaccccttc ggatttcattt gcccagctat	2280
ctgtcactttt atttgtcaaga tagtggaaaa ggaagggtggc ttctacaat gccatcattt	2340
cgataaagga aaggccatcg ttgtttttttt gttttttttt gttttttttt gttttttttt	2400
cccaaaaaacgg aggagcatcg tggaaaaaaa gttttttttt gttttttttt gttttttttt	2460

ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga	2520
tacagtctca gaagacaaa gggcaattga gactttcaa caaaggtaa tatccggaaa	2580
cctccctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tgaaaagga	2640
aggtggctcc tacaatgcc atcatgcga taaaggaag gccatcgting aagatgcctc	2700
tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtag aaaaagaaga	2760
cgttccaacc acgtttcaa agcaagtgga ttgatgtat atctccactg acgttaaggga	2820
tgacgcacaa tcccactatc ctgcgaaga cccttcctct atataaggaa gttcattca	2880
tttggagagg tattaaaatc ttaatagtt ttgataaaag cgaacgtggg gaaacccgaa	2940
ccaaacccitc ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtcttc	3000
ttgcgtgagc gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgtttctt	3060
tcactgaagc gaaatcaaag atctcttgtt ggacacgttag tgccgcgcc ttaaataaag	3120
tgtacttgc ctattcttgtt cggtgtggc ttggaaaag aaagcttgct ggaggctgct	3180
gttcagcccc atacattact ttttacgatt ctgctgactt tcggcgggtt caaatctct	3240
acttctgctt gacgaggtat tttgcctgt acttcttct tttcttctt gctgatttgt	3300
tctataagaa atctagtatt ttcttgaaa cagagtttc ccgtgggtt cgaacttgga	3360
gaaagattgt taagcttctg tatattctgc ccaaatttg cgggcccgcg gatggcgaaa	3420
aacgttgcga tttcggctt attgtttctt cttcttgtt tggttccttc tcagatcttc	3480
gcctgcaggc tcctcagcca aaacgacacc cccatctgtc tatccactgg cccctggatc	3540
tgctgcccua actaactcca tggtgaccct gggatgcctg gtcaagggtt atttccctga	3600
gccagtgaca gtgacctgga actctggatc cctgtccagc ggtgtgcaca cttcccaagc	3660
tgtcctgcag tctgacctct acactctgag cagctcagtg actgtccctt ccagcacctg	3720
gcccagcgag accgtcacct gcaacgttgc ccacccggcc agcagcacca aggtggacaa	3780
gaaaattgtg cccaggatt gtgggtgtaa gccttgata tttacagtcc cagaagtatc	3840
atctgtcttc atttccccca caaagccaa ggatgtgcacc accattactc tgactctaa	3900
ggtcacgtgtt gttgtggtag acatcagca ggtatgtccc gaggtccagt tcagctggtt	3960
tgttagatgtatgttgc acacagctca gacgcaaccc cgggaggagc agttcaacag	4020
cacttccgc tcagtcagtg aacttcccat catgcaccag gactggctca atggcaagga	4080
gcgatcgctc accatcacca taccatcac catcaccatt aaaggcttat tttctttagt	4140
ttgaatttac ttttattcgg tttgcatttc tatgttttgtt gagcgggttt ctgtgctcag	4200
agtgtgttta tttatgtaa ttaatttctt ttgtgagctc ctgttttagca ggtcgccct	4260
tcagcaagga cacaaaaaga ttttaattttt attaaaaaaa aaaaaaaaaa agaccggaa	4320

tgcatatac agtttatcga cctgcagatc gttcaaacat ttggcaataa agtttcttaa	4380
gattgaatcc tggcccggt ctgcgatga ttatcatata atttctgtt aattacgtta	4440
agcatgtaat aattaacatg taatgcata cgttattt gagatgggtt tttatgatta	4500
gagtcggca atiatacatt taatacgcga tagaaaacaa aatatacgca gcaaactagg	4560

ataaattatc gcgcgcggtg tcatctatgt tactagatct ctagagtctc aagcttggcg	4620
cgcacgtg actagtggca ctggccgtcg tttacaacg tcgtgactgg gaaaaccctg	4680
cgcttaccca acttaatcgc ctgcagcac atccccctt cgccagctgg cgtaatagcg	4740
aagaggcccg caccgatcgc cttcccaac agttgcgcag cctgaatggc gaatgctaga	4800
gcagcttgcag ctggatcag attgtcggtt cccgccttca gttaaacta tcagtgttg	4860
acaggatata ttggcgggta aacctaagag aaaagagcgt tta	4903

<210> 24

<211> 2721

<212> DNA

<213> Artificial sequence

<220><223> Expression cassette number 1733

<400> 24

gtcaacatgg tggagcacga cacacttgc tactccaaaa atatcaaaga tacagtctca	60
gaagacaaaa gggcaattga gactttcaa caaaggtaa tatccgaaa cctccctcgga	120
ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc	180
tacaaatgcc atcattgcga taaaggaaag gccatcggt aagatgcctc tgccgacagt	240
ggtcccaag atggaccccc acccacgagg agcatcggtt aaaaagaaga cgttccaacc	300
acgtcttcaa agcaagtggta ttgatgtgat aacatggtgg agcacgacac acttgtctac	360

tccaaaaata tcaaagatac agtctcagaa gaccaaggaa caattgagac ttttcaacaa	420
aggtaatat ccggaaacct cctcgattt cattgcccag ctatctgtca ctttattgt	480
aagatagtgg aaaaggaagg tggctctac aaatgccatc attgcataa aggaaaggcc	540
atcggttgaag atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagc	600
atcggtgaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattt atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttccctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggtttt ataaaagcga	780

acgtgggaa acccgaacca aactttctt taaactctt ctcattctc ttaaagcaaa	840
cttctctttt gttttctt cgtgagcgtt cttcaacgtt gtcagatcgt gttcggcac	900

cagtacaacg tttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtatgc	960
ggccccaattt aataacgtgt acttgtccta ttcttgctgg tgggtcttggaaaagaaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgc gaggtattgt tgccgtact tctttctct	1140
tcttcttgct gattggttct ataagaatc tagtatttc tttgaaacag agtttcccg	1200
tggtttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg	1260
gccccatgtac ggcacatcgagt atacaacaat tttaatttc ctgatttcca tcattctgtt	1320
aaactacatc ctaagtcgg tgaccagaat tatggattat attatctatc gtggcctct	1380
catctacgtg gcccttttg ccctgaccag ggcccagaac tatggcctga acttaccaat	1440
caccgggtca atggataccg ttacgctaa ttccactcaa gaggggatat ttctgacaag	1500
tacctgtgc ctgtattatc caacagaagc ctctacccag atcaatgtg gggagtggaa	1560
ggatagtctc tcacagatgt tcctaaccaa gggctggccc accgggtccg tctacttcaa	1620
ggaataactct agtattgtcg acttctcagt tgaccccccag ctttatttgcg actacaacct	1680
gtgtacttatg aaatacgacc agaacctgga gctggatatg tccgagctgg ctgacacctat	1740
cctcaatgag tggctgtgca accccatgga catcacatta tattactacc agcagtctgg	1800
agaatccaac aagtggatca gtatgggctc aagttgcacc gtgaagggtgt gtccttcaa	1860
cacccaaatg ctggcattt gttgtcagac aactaatgtg gattcggtt aaatggtagc	1920
cgaaaacgag aagctggcta tagtggacgt agtcgtatggg attaaccaca agatcaatct	1980
gactaccacc acttgtacca tcagaaactg taaaaagctc ggccccggg agaacgtcgc	2040
cgtgatccag gtggggggga gcaatgtgct cgacattact gccgacccta ccaccaatcc	2100
acagacggaa cggatgatga gaggtaactg gaagaaatgg tggcaggctt tttataccat	2160
tgtggactac attaaccaga ttgtgcaagt catgagtaaa cggccatgtt ccctgaactc	2220
agcagcccttc tattatcgctt ttttagggcc tattttcttt agttgttatt tactgttatt	2280
cgggtgtcat ttctatgttt ggtgagcggt ttctgtgct cagagtgtgt ttatattatg	2340
taatttaatt tcttgtgag ctccgttta gcaggcgat ccttcagcaa ggacacaaaa	2400
agatttaat ttataaaaaaa aaaaaaaaaaaa aaaagacggg gaattcgata tcaagcttat	2460
cgaccctgcag atcgttcaaa catttggcaa taaagttct taagatgaa tcctgttgc	2520
ggcttgcga tgattatcat ataatttctg ttgaattacg ttaagcatgt aataattaac	2580
atgtaatgca tgacgttatt tatgagatgg gttttatga ttagatccc gcaattatac	2640
atttaaatcg cgatagaaaaaaa caaaatatacg cgccaaact aggataaatt atcgcgcgcg	2700
gtgtcatcta ttttactaga t	2721

<210> 25

<211> 890

<212> PRT

<213> Rotavirus

<400> 25

Met Ala Tyr Arg Lys Arg Gly Ala Lys Arg Glu Asn Leu Pro Gln Gln

1 5 10 15

Asn Glu Arg Leu Gln Glu Lys Glu Ile Glu Lys Asp Val Asp Val Thr

20 25 30

Met Glu Asn Lys Asn Asn Arg Lys Gln Gln Leu Ser Asp Lys Val

35 40 45

Leu Ser Gln Lys Glu Glu Ile Ile Thr Asp Ala Gln Asp Asp Ile Lys

50 55 60

Ile Ala Gly Glu Ile Lys Ser Ser Lys Glu Glu Ser Lys Gln Leu

65 70 75 80

Leu Glu Ile Leu Lys Thr Lys Glu Asp His Gln Lys Glu Ile Gln Tyr

85 90 95

Glu Ile Leu Gln Lys Thr Ile Pro Thr Phe Glu Ser Lys Glu Ser Ile

100 105 110

Leu Lys Lys Leu Glu Asp Ile Arg Pro Glu Gln Ala Lys Lys Gln Met

115 120 125

Lys Leu Phe Arg Ile Phe Glu Pro Lys Gln Leu Pro Ile Tyr Arg Ala

130 135 140

Asn Gly Glu Lys Glu Leu Arg Asn Arg Trp Tyr Trp Lys Leu Lys Lys

145 150 155 160

Asp Thr Leu Pro Asp Gly Asp Tyr Asp Val Arg Glu Tyr Phe Leu Asn

165 170 175

Leu Tyr Asp Gln Ile Leu Ile Glu Met Pro Asp Tyr Leu Leu Leu Lys

180 185 190

Asp Met Ala Val Glu Asn Lys Asn Ser Arg Asp Ala Gly Lys Val Val

195 200 205

Asp Ser Glu Thr Ala Asn Ile Cys Asp Ala Ile Phe Gln Asp Glu Glu
 210 215 220
 Thr Glu Gly Val Val Arg Arg Phe Ile Ala Asp Met Arg Gln Gln Val
 225 230 235 240
 Gln Ala Asp Arg Asn Ile Val Asn Tyr Pro Ser Ile Leu His Pro Ile
 245 250 255
 Asp His Ala Phe Asn Glu Tyr Phe Leu Asn His Gln Leu Val Glu Pro
 260 265 270
 Leu Asn Asn Glu Ile Ile Phe Asn Tyr Ile Pro Glu Arg Ile Arg Asn
 275 280 285
 Asp Val Asn Tyr Ile Leu Asn Met Asp Met Asn Leu Pro Ser Thr Ala
 290 295 300
 Arg Tyr Ile Arg Pro Asn Leu Leu Gln Asp Arg Leu Asn Leu His Asp
 305 310 315 320
 Asn Phe Glu Ser Leu Trp Asp Thr Ile Thr Thr Ser Asn Tyr Ile Leu
 325 330 335
 Ala Arg Ser Val Val Pro Asp Leu Lys Glu Lys Glu Leu Val Ser Thr
 340 345 350
 Glu Ala Gln Ile Gln Lys Met Ser Gln Asp Leu Gln Leu Glu Ala Leu
 355 360 365
 Thr Ile Gln Ser Glu Thr Gln Phe Leu Ala Gly Ile Asn Ser Gln Ala
 370 375 380
 Ala Asn Asp Cys Phe Lys Thr Leu Ile Ala Ala Met Leu Ser Gln Arg
 385 390 395 400
 Thr Met Ser Leu Asp Phe Val Thr Thr Asn Tyr Met Ser Leu Ile Ser
 405 410 415
 Gly Met Trp Leu Leu Thr Val Ile Pro Asn Asp Met Phe Leu Arg Glu
 420 425 430
 Ser Leu Val Ala Cys Glu Leu Ala Ile Ile Asn Thr Ile Val Tyr Pro
 435 440 445
 Ala Phe Gly Met Gln Arg Met His Tyr Arg Asn Gly Asp Pro Gln Thr

450	455	460
Pro Phe Gln Ile Ala Glu Gln Gln Ile Gln Asn Phe Gln Val Ala Asn		
465	470	475
Trp Leu His Phe Ile Asn Asn Asn Arg Phe Arg Gln Val Val Ile Asp		
485	490	495
Gly Val Leu Asn Gln Thr Leu Asn Asp Asn Ile Arg Asn Gly Gln Val		
500	505	510
Ile Asn Gln Leu Met Glu Ala Leu Met Gln Leu Ser Arg Gln Gln Phe		
515	520	525
Pro Thr Met Pro Val Asp Tyr Lys Arg Ser Ile Gln Arg Gly Ile Leu		
530	535	540
Leu Leu Ser Asn Arg Leu Gly Gln Leu Val Asp Leu Thr Arg Leu Val		
545	550	555
Ser Tyr Asn Tyr Glu Thr Leu Met Ala Cys Val Thr Met Asn Met Gln		
565	570	575
His Val Gln Thr Leu Thr Thr Glu Lys Leu Gln Leu Thr Ser Val Thr		
580	585	590
Ser Leu Cys Met Leu Ile Gly Asn Thr Thr Val Ile Pro Ser Pro Gln		
595	600	605
Thr Leu Phe His Tyr Tyr Asn Ile Asn Val Asn Phe His Ser Asn Tyr		
610	615	620
Asn Glu Arg Ile Asn Asp Ala Val Ala Ile Ile Thr Ala Ala Asn Arg		
625	630	635
Leu Asn Leu Tyr Gln Lys Lys Met Lys Ser Ile Val Glu Asp Phe Leu		
645	650	655
Lys Arg Leu Gln Ile Phe Asp Val Pro Arg Val Pro Asp Asp Gln Met		
660	665	670
Tyr Arg Leu Arg Asp Arg Leu Arg Leu Leu Pro Val Glu Arg Arg Arg		
675	680	685
Leu Asp Ile Phe Asn Leu Ile Leu Met Asn Met Glu Gln Ile Glu Arg		
690	695	700

Ala Ser Asp Lys Ile Ala Gln Gly Val Ile Ile Ala Tyr Arg Asp Met

705 710 715 720

Gln Leu Glu Arg Asp Glu Met Tyr Gly Tyr Val Asn Ile Ala Arg Asn

725 730 735

Leu Asp Gly Tyr Gln Gln Ile Asn Leu Glu Glu Leu Met Arg Thr Gly

740 745 750

Asp Tyr Gly Gln Ile Thr Asn Met Leu Leu Asn Asn Gln Pro Val Ala

755 760 765

Leu Val Gly Ala Leu Pro Phe Val Thr Asp Ser Ser Val Ile Ser Leu

770 775 780

Ile Ala Lys Leu Asp Ala Thr Val Phe Ala Gln Ile Val Lys Leu Arg

785 790 795 800

Lys Val Asp Thr Leu Lys Pro Ile Leu Tyr Lys Ile Asn Ser Asp Ser

805 810 815

Asn Asp Phe Tyr Leu Val Ala Asn Tyr Asp Trp Ile Pro Thr Ser Thr

820 825 830

Thr Lys Val Tyr Lys Gln Val Pro Gln Pro Phe Asp Phe Arg Ala Ser

835 840 845

Met His Met Leu Thr Ser Asn Leu Thr Phe Thr Val Tyr Ser Asp Leu

850 855 860

Leu Ser Phe Val Ser Ala Asp Thr Val Glu Pro Ile Asn Ala Val Ala

865 870 875 880

Phe Asp Asn Met Arg Ile Met Asn Glu Leu

885 890

<210> 26

<211> 6745

<212> DNA

<213> Artificial sequence

<220><223> Construct 193

<

400> 26

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgtttta atgtactgaa ttaacgccga atccggct ggtatattta tatgttgtca	120
aataactcaa aaaccataaa agttaagt agcaagtgt tacattttt cttgaacaaa	180
aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg	240
ataagaacaa gagtagtgat atttgacaa caatttgtt gcaacattig agaaaatttt	300
gttgcctct ctttcattt gtcaaaaaca atagagagag aaaaaggaag agggagaata	360
aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtacaaa atagttgtac	420
aaatatcatt gaggatttg acaaaagcta cacaataag ggttaattgc tgtaaataaa	480
taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga	540
aagaataaat tattttaaa attaaaagt gatcatttg attaaacatg tgattattta	600
atgaattgtat gaaagagttt gattaaagt gtattagtaa ttggaaattt gtgtcaaatt	660
taatttgcata ttgtatctt tcctatataat tgccccatag agtcagttaa ctcattttt	720
tatttcata agcaataaag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg	780
gtatatttac taaaaatct aagccacgtt ggaggataac aggatccccg taggaggata	840
acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat	900
ctgtggcaca tctacattat ctaaatcaca cattttcca cacatctgag ccacacaaaa	960
accaatccac attttatca cccattctat aaaaaatcac actttgtgag tctacacttt	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggAAC gagctataca agggaaacgac gctaggAAC aagctaacag tgaacgttgg	1140
gatggaggat caggaggatc cacttctccc ttcaaacttc ctgacgaaag tccgagttgg	1200
actgagtggc ggctacataa cgatgagacg aattcaatc aagataatcc cttgggttc	1260
aaggaaagct ggggttcgg gaaagttgtt ttaagagat atctcagata cgacaggacg	1320
gaagcttcac tgacagagt cttggatct tggacggag attcggttaa ctatgcagca	1380
tctcgatttt tcggttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt	1440
agtatcaccg ttctggagg gtgcgaact cttcagcatc tctgtgagat ggcaattcg	1500
tctaagcaag aactgctaca gttggccca atcgaagtgg aaagtaatgt atcaagagga	1560
tgcctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttttttt gcttcttcgt	1620
ctcctattta taatatggtt ttgttattttt aattttgttc ttgtagaaga gcttaattaa	1680
tcgttgttgc tattttgtatg agatgaactg gtgtatgttta attcatttac	1740
ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg	1800
cgtacaatttgc ttatatttta gaacaactaa aattgaacat ctttgccac aactttataa	1860
gtggtaata tagtcaa atatggtcaa gttcaataga ttaataatgg aaatatcgt	1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaatttat atcatccct	1980
ttgataaatg atagtagcacc aatttaggaag gagcatgctc gcctaggaga ttgtcgttc	2040
ccgccttcag tttgcaagct gctctagccg tgttagccat acgcaaaccg cctctccccg	2100
cgcgttggga attactagcg cgtgtcgaga cgcggttg ttgtgactcc gaggggttgc	2160
ctcaactct atcttataac cggcgtggag gcatggagc aggggtattt tggtcatttt	2220
aatagatgt ggaaaatgac gtgaaatata cttaaagacg aagtcttgc gacaaggggg	2280
ggcccacgcc gaatttaata ttaccggcgt ggcccccct tatcgagtg ctgttagcac	2340
gagcggtcca gatttaagt agaaaattc ccgcccacta gggtaaagg tgttcacact	2400
ataaaagcat atacgatgtg atggtatttg gtcgacaagc ttgcattccg gtcaacatgg	2460
tggagcacga cacacttgc tactccaaaa atatcaaaga tacagtcga gaagacccaa	2520
ggcaattga gactttcaa caaaggtaa tatccggaaa ctcctcgga ttccattgcc	2580
cagctatctg tcactttattt gtgaagatag tggaaaagga aggtggctcc tacaatgcc	2640
atcattgcga taaaggaaag gccatcggt aagatgcctc tgccgacagt ggtccaaag	2700
atggaccccc acccacgagg agcatcggtt aaaaagaaga cgttccaaacc acgtctcaa	2760
agcaagtgga ttgatgtgat aacatgggtt agcacgacac acttgtctac tccaaaata	2820
tcaaagatac agtctcagaa gaccaaaggg caattgagac tttcaacaa aggtaatat	2880
ccggaaacct cctcgattt cattgcccag ctatctgtca ctttatgtt aagatagtgg	2940
aaaaggaagg tggctctac aaatgccatc attgcataa aggaaggcc atcggttgaag	3000
atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagc atcggtggaaa	3060
aagaagacgt tccaaccacg tcttcaaagc aagtggattt atgtgatatac tccactgacg	3120
taagggatga cgacaaatcc cactatcctt cgcaagaccc ttccctata taaggaagtt	3180
catttcattt ggagaggtat taaaatctt ataggttttg ataaaagcga acgtggggaa	3240
acccgaacca aaccccttc taaactctct ctcattctc tttaagccaa cttctcttt	3300
gtcttcttg cgtgagcgat cttcaacgtt gtcagatgtt gttcgac cagtacaacg	3360
ttttcttca ctgaagcgaa atcaaagatc tctttgttga cacgtgtgc ggcgccatta	3420
aataacgtgt acttgtccta ttcttgcgg tgggtcttgg ggaaaagaaaa gcttgcttgg	3480
ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg gcgggtgcaa	3540
tatctctact tctgtttgac gaggtattgt tgcgtgtact tctttcttct tcttcttgc	3600
gattggttct ataagaaatc tagtattttc ttgaaacag agttttcccg tggttttcgaa	3660
acttggagaa agattgttaa gttctgtat atttgcctt aatttgcgg gcccgcggat	3720

ggcgaaaaac gttgcgattt tcggcttatt gtttctttt cttgtgttgg ttccttctca	3780
gatcttcgcc tcaggcgtcc tcagccaaa cgacacccc atctgtctat ccactggccc	3840
ctggatctgc tgcccaaact aactccatgg tgaccctggg atgcctggc aagggttatt	3900
tccctgagcc agtgacagtgc acctggaaact ctggatccct gtccagcgggt gtgcacacct	3960
tcccagctgt cctgcagtct gacctctaca ctctgagcagc ctcatgtact gtcccttcca	4020
gcacctggcc cagcgagacc gtcacctgca acgttgcaccc cccggccagc agcaccaagg	4080
tggacaagaa aatttgtgccc agggatttg gttgtaaagcc ttgcataatgt acagtcccag	4140
aagtatcatc tgtcttcatac ttccccccaa agcccaagga tgtgctcacc attactctga	4200
ctcctaaggcacgtgtt gttgttagaca tcagcaagga tgatcccgag gtccagttca	4260
gctggttgtt agatgtatgtg gaggtgcaca cagctcagac gcaaccccg gaggagcgt	4320
tcaacagcac ttccgcgtca gtcagtgaac ttcccatcat gcaccaggac tggctcaatg	4380
gcaaggagcg atcgcctacc atcaccatca ccatcaccat caccataaa ggcctat	4440
cttttagttt aatttactgt tattcggtt gcattctat gttggtgag cggtttctg	4500
tgctcagagt gtgttattt tatgttaattt aatttcttg ttagctcctg ttttagcaggt	4560
cgtcccttca gcaaggacac aaaaagattt taatttatt aaaaaaaaaa aaaaaaaaga	4620
ccggaaattt gatatcaagc ttatcgacct gcagatcgaa caaacattt gcaataaagt	4680
ttcttaagat tgaatcctgt tgccggctt gcgatgatta tcatataatt tctgttgaat	4740
tacgttaagc atgtataat taacatgtaa tgcgtacgt tatttatgat atgggtttt	4800
atgatttagag tcccgcaatt atacatttaa tacgcgtatg aaaacaaaat atagcgcga	4860
aactaggata aattatcgcg cgccgtgtca tctatgttac tagatctcta gagtctcaag	4920
cttggcgcgc cataaaatga ttatatttgc aatatatttgc attgtgcaag tagatagaaa	4980
ttacatatgt tacataaacac acgaaataaa caaaaaaaga caatccaaaa acaaacaccc	5040
aaaaaaaaat aatcacttta gataaactcg tatgaggaga ggcacgttca gtgactcgac	5100
gattcccgag caaaaaaagt ctcccggtca cacatatagt gggtgacgca attatctta	5160
aagtaatcct tctgttgcgt tgcattgtat aacatccagt ctgcgtcagg attgcaaga	5220
attatagaag ggatcccacc ttatatttgc ttctttttc catatttagg gttgacagtgc	5280
aaatcagact ggcaacctat taattgcgtt cacaatggaa cgaacttgc gggatgtcg	5340
tcgtatgtat tataggtggc gtgttcatgc tagttggta aatcgatgtt accgttccaa	5400
tagttgtgtc gtccgagact tctagccag gtggcttcccgatcgact tggtccgcag	5460

atgttagaggc tgggttgtcg gattccattc cttccattgt cctggtaaa tcggccatcc	5520
atccaagggtc agattgagct tgggtatg agacaggatg tatgtaagta taaggctcta	5580
tgcattacatg gtatagatgg gttccctcc aggagtgttag atttcgtag cagcgaagat	5640
ctgattctgt gaagggcgac acatacggtt caggttgtgg agggataaat ttgttgctg	5700
aatatccag ccattgaagt tttgtgcc attcatgagg gaattctcc ttgtatcatgt	5760
caagatattc ctcccttagac gttgcagtct ggataatagt tctccatgt gcgtcagatt	5820
tgcgaggaga gaccttatga tctcgaaat ctccctgtt ttaataatct ccgtcccttg	5880

atatgtaatc aaggacttgt ttagagtttc tagctggctg gatattaggg tgatttcctt	5940
caaaatcgaa aaaagaagga tccctaatac aaggtttttt atcaagctgg agaagagcat	6000
gatagtgggt agtgccatct tcatgtaaac cagaagcaac accaaggaaag aaaataagaa	6060
aaggtgtgag tttctcccag agaaaactgga ataaatcatc tctttgagat ggcacttgg	6120
gataggttaag gaaaaacatat tttagattgga gtctgaagtt cttaactagca gaaggcatgt	6180
tgttgtgact ccgaggggtt gcctcaaact ctatcttata accggcgtgg aggcatggag	6240
gcaggggtat tttgggtcatt ttaatagata gtggaaaatg acgttggatt tacttaaaga	6300

cgaagtctt	gcgacaaggg	ggggcccacg	ccgaatttaa	tattaccggc	gtggccccc	6360
cttatcgca	gtgcttagc	acgagcggtc	cagatttaaa	gtagaaaatt	tcccgcccc	6420
tagggttaaa	ggtgttcaca	ctataaaagc	atatacgatg	tgatggtatt	tgactagtg	6480
cactggccgt	cgtttacaa	cgtcgact	ggaaaaaccc	tggcgtaacc	caacttaatc	6540
gccttgcagc	acatccccct	ttcgccagct	ggcgtaatag	cgaaggaggcc	cgcaccgatc	6600
gcccttccca	acagttgcgc	agcctgaatg	gcgaatgcta	gagcagcttgc	agcttggatc	6660
agattgtcggt	ttcccgccct	cagtttaaac	tatcagtggtt	tgacaggata	tattggcggg	6720

taaacctaag agaaaagagc gttta 6745

<210> 27

<211> 4413

<212> DNA

<213> Artificial sequence

<220><223> Expression cassette number 1710

<400> 27

```
gtcaacatgg tggagcacga cacacttgc tacccaaaa atatcaaaga tacagtctca 60  
gaagacccaaa gggcaattga gactttcaa caaagggtaa tatccggaaa cctcctcgga 120  
ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180  
tacaatgcc atcattgcga taaaggaaag gccatcggt aagatgcctc tgccgacagt 240
```

ggtcccaaag atggaccccc acccacgagg agcatcggtt aaaaagaaga cgttccaacc	300
acgtcttcaa agcaagtggg ttgatgttat aacatgggtt agcacgacac acttgtctac	360
tccaaaata tcaaagatac agtctcgaa gaccaaggaa caattgagac tttcaacaa	420
aggtaatat ccggaaacct cctcgattt cattgccag ctatctgtca ctttatttg	480
aagatagtgg aaaaggagg tggctctac aaatgccatc attgcataa aggaaaggcc	540
atcggtgaag atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagc	600
atcggtgaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattt atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttccctata	720
taaggaagtt catttcattt ggagaggtt taataatctta ataggtttg ataaaagcga	780
acgtgggaa acccgAACCA aaccccttc taaactctt ctcattctc ttaaagcaaa	840
catttcattt gtcttcttgc cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac	900
cagtacaacg tttctttca ctgaagcgaa atcaaagatc tctttgttgg cacgtatgc	960
ggccatatta aataacgtgt acttgccta ttcttgcgg tgggttttgg ggaaaagaaa	1020
gtttgttggg ggctgctgtt cagccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgc gaggtattgt tgccctgtact tcttttttct	1140
tcttcttgct gattggttct ataagaaatc tagtattttc ttgaaacag agtttcccg	1200
tggtttgcg acitggagaa agattgttaa gcttctgtat attctgcacca aatttgcgg	1260
gcccatggca taccggaga gaggagcaaa gcgcgaaac ctgcgcac agaacgagag	1320
actgcaagaa aaagagatag agaaagatgt cgacgtacca atggaaaaca agaataacaa	1380
taggaaacaa cagctgtccg acaaagtctt gtcggcggaaag gagaaattt tcactgacgc	1440
ccagggacat attaaaattt ccggagaaat aaagaagacg tcgaaagaag aatctaaaca	1500
gctgctgaa attctgaaaa caaaagaaga ccatcagaaa gagattcaat atgaaatttt	1560
gcaaaaaaca atacctacat ttgagtccaa agaaagtatc ctcaagaagc ttgaagacat	1620
aagacccggag caggcaaaaa aacagatgaa actctttcgat atttcgatc caaaacagct	1680
ccctatatacg cgcccaatg gcgagaaggaa gctacgcaac cgggtgtact ggaagttgaa	1740
aaaagacacc ctgcccggatc gagattatgaa cgtccggag tatttcctca atctctatgaa	1800
tcagatcctc atcgaaatgc cgactatct gtcctcaag gacatggccg tggagaacaa	1860
aaatagcaga gacgcggca aagttgtcgat ctctgagact gccaatattt gtgtgtccat	1920
cttccaggat gaggagacccg agggagtcgt ccgttagattc atcgctgata tgccgcacaa	1980

ggtccaggct gatcgtaaca ttgtcaatta cccttcac cttcaccta ttgatcatgc	2040
attcaat gag tatttctta accaccagt ggtggagccg ctgaacaatg agataat ctt	2100
caattacata ccagagagga taaggaatga cgtgaattac atcctgaaca tggatatgaa	2160
tctgccatct acagccaggt atatcaggcc aaacttgtg caggatagac tgaatctca	2220
cgataat ttt gagtccctgt gggataccat cacaacatcc aactacattc tggccaggc	2280
cgtcg tccc gatttgaagg agaaggagct ggtctccacc gaagcacaga tccagaaaat	2340
gagccaggac ctgcagctgg aggcctcac tattcagagc gagacacagt ttttagccgg	2400
gattaacagt caggctgcca atgattgttt caagaccctc atagccgcca tgctgtctca	2460
aagaaccatg tcitggact ttgtgaccac gaactatatg agccta atct ccggaa atgt	2520
gctacttaca gtgattccca acgatatgtt cctccggag tcactagtgg cctgtgagct	2580
ggcgatcatc aacaccatcg tgtatccagc attcggaa atg cagaatgc attacggaa	2640
tggcgaccct cagacaccct tccagatcgc agaacagcag atccagaatt tccaggtggc	2700
gaactggctc cattttatta acaataacag attcaggcaa gttgtgattt atggagttct	2760
gaatcagact ctgaacgaca atatacggaa tggacaggc atcaaccagc tcatggaa	2820
attgatgcaa ctacgcgac agcagttccc cacgatgcct gtggattaca aacggagcat	2880
ccaaacggggc attctgttc tctccaaatag gctggggcag cttgtcgact taacccgact	2940
ggtctcctat aactacgaga cgcta atggc ttgtgtgacc atgaacatgc agcacgtgca	3000
aaccctgaca actgagaagt tgcagctac ttctgtgact tcgc ttgtgta tgttaattgg	3060
taacacaacc gtgattccgt ccccacagac actgttccac tactacaaca tcaacgtgaa	3120
tttccactcc aattataatg agcggatcaa cgacgcgctc gccataatta ccgcagcaaa	3180
taggctgaat ctatcaga aaaaaatgaa gtccatagtg gaagacttcc taaaacggct	3240
ccagat ttc gacgtaccac gagtgcctga cgaccaatg tacaggctga gggatgcct	3300
tcggcttta cccgttgaac ggagacggct tgacatatc aacttgcattt tgatgaaat	3360
ggagcagatc gaacgcgtt ctgataagat tgctcagggg gttatcatcg cataccgaga	3420
tatgcagctg gaacgcgacg agatgtacgg atatgttaat attgcacggaa atcttgcattt	3480
ctaccagcaa attaacttgg aggaactcat ggcgcacccgt gattacggac aaattacgaa	3540
catgcttc aacaatcaac ccgttgcct tgggggtgca ttgccttcg ttacggactc	3600
atccgtgatc agtcta atcg ccaagctcgca cgcaaccgtc ttgcctcaga tagtgaagct	3660
cagggaaatggt gacacactga agccatact gtacaaaata aactcggtt ccaatgactt	3720
ttaccttgc gccaactacg actggatccc cacaagtaca actaaggct acaaacaggt	3780
gccacaacca ttgcacttta gagccagcat gcacatgcgt acttctaacc ttacgttac	3840

cgtctactct gacctactgt catttggttc agcggacacg gtagagccca ttaacgcagt	3900
cgcattcgac aatatgcgaa taatgaacga gctttaaagg cctatttct ttagttgaa	3960
tttactgtta ttccgggtgtgc atttctatgt ttggtgagcg gtttctgtg ctcagagtgt	4020
gtttatTTTA tGtaatttaa tttcttttgtg agctccgtt tagcaggctg tcccttcagc	4080

aaggacacaa aaagatttta attttattaa aaaaaaaaaaaa aaaaaagacc ggaaattcga	4140
tatcaagctt atcgacctgc agatcggtca aacatttggc aataaagttt cttaagattt	4200
aatccgttg ccggcttgc gatgattatc atataatttc ttttgaatta cgttaagcat	4260
gtaataatttta acatgtaatg catgacgtt tttatgagat gggtttttat gattagatc	4320
ccgcaatttat acatTTATAA cgcgatagaa aacaaaatat agcgccaaa ctaggataaa	4380
ttatcgccgcg cgggtgtcatc tatgttacta gat	4413

<210> 28

<211> 49

<212> DNA

<213> Artificial sequence

<220><223> Primer IF-WA_VP6(opt).s1+3c

<400> 28

aaatttgcg gccccatgga ggtcctttat agtctctcca aaacgctga	49
--	----

<210> 29

<211> 52

<212> DNA

<213> Artificial sequence

<220><223> Primer IF-WA_VP6(opt).s1-4r

<400> 29

actaaagaaa ataggcctct acttgatcaa catactccgg atagaggccca ca	52
--	----

<210> 30

<211> 2934

<212> DNA

<213> Artificial sequence

<220><223> Expression cassette number 1713

<400> 30

gtcaacatgg tggagcacga cacacttgc tactccaaaa atatcaaaga tacagtctca	60
--	----

gaagacacaaa gggcaattga gactttcaa caaaggtaa tatccggaaa cctccctcgga	120
---	-----

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc	180
tacaatgcc atcattgcga taaaggaaag gccatcgta aagatgcctc tgccgacagt	240
ggtcccaaag atggaccccc acccacgagg agcatcgta aaaaagaaga cgttccaacc	300
acgtctcaa agcaagtgga ttgtatgtat aacatggtg agcacgacac acttgctac	360
tccaaaata tcaaagatac agtctcagaa gaccaaggaa caattgagac tttcaacaa	420
aggtaatat ccggaaacct cctcgattc cattgcccag ctatctgtca ctttattgtg	480
aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcataa aggaaaggcc	540
atcggttaag atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagc	600
atcggtgaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattt atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttccctata	720
taaggaagtt catttcattt ggagaggtat taaaatctt ataggtttg ataaaagcga	780
acgtgggaa acccgaacca aaccccttc taaactctct ctcatctctc ttaaagcaaa	840
cttctctttt gctttcttg cgtgagcgat cttcaacgtt gtcagatgt gcttcggcac	900
cagtacaacg tttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtatgc	960
ggccgcatta aataacgtgt acttgtccta ttcttgcgg tgtggcttg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgcg gaggtattgt tgccctgtact tctttttct	1140
tcttcttgct gattgggtct ataagaaatc tagtattttc tttgaaacag agtttcccg	1200
tgggtttcga acttggagaa agattgttaa gcttctgtat attctgccc aatttgtcg	1260
gccccatggag gtccttata gtctctccaa aacgctgaag gacgctaggg acaagatcgt	1320
ggagggtaca ctttatagca atgtcagcga cctaatacag cagtttaatc aaatgatcgt	1380
tacaatgaat gggaaatgatt tccaaactgg cggattttgtt aatctgccc tgaggaactg	1440
gacatttcgtt ttccgcctgc tggcacgac tctccttaat ctgcgtgaa attatgtaga	1500
aaacgccaga acgattatcg agtactttt cgttccatt gataacgtt gtatggatga	1560
gatggcccgcc gagtccaaac ggaacggagt tgctccacag tccgaggccc ttccgaaact	1620
cgccggcatt aagtcaagc gtattaaattt cgacaactcc tccgaatata tagagaactg	1680
gaacttgcag aatcgctcgcg agagaaccgg cttcggttc cataaaccta atatcttcc	1740
gtatagcgcc tcattcaccc tgaataggag tcagccatg cacgacaacc tcatgggtac	1800
aatgtggctg aatgcgggaa gtgaaataca ggtcgccggg ttgcattact cctgtgccat	1860
taatgcaccc gcaaacatcc agcagttcga acatatcgta caactaagac gggctctcac	1920
gaccgcgaca attacactcc tgcccgacgc cgagcgcttc tccttcccc gctgtatcaa	1980

ctcagctgat ggcgccacca cttgggttctt caaccctgtt atattgcgcc ctaacaacgt	2040
agaggtggag tttctctaa acggacagat catcaatacc taccaaggcca ggttcggcac	2100
gattattgca agaaaattcg acgctatcg gctgctcttc caactgatga ggccccccaa	2160

tatgactccc gctgtgaacg ctttgttcc gcaggcttag cctttccagc accacgccac	2220
cgtcggtttg actcttcgaa tagagagcgc ggtctgcgaa tcagtgcgttgg cagacgccaat	2280
cgagacgctg ctggcaaacg ttaccggcgt gcggcaagag tatgccatcc cagtagggcc	2340
tgtgtttcca cccggcatga actggactga actaattact aactatagcc catccagaga	2400
agacaacttg cagcgggtct tcactgtggc ctctatccgg agtatgttga tcaagttagag	2460
gcctattttc tttagtttga atttactgtt attcggtgtt catttctatg tttggtgagc	2520
ggttttctgt gctcagatgt tgtttatttt atgttaatttatttcttgc gagctcgtt	2580

ttagcaggcgtcccttcag caaggacaca aaaagatttt aattttatta aaaaaaaaaaa	2640
aaaaaaaaagac cgggaattcg atatcaagct tatcgacctg cagatcggtt aaacatttgg	2700
caataaaagt tcttaagatt gaatcctgtt gccggcttg cgatgattat catataattt	2760
ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga	2820
tgggttttta tgatttagagt cccgcaatta tacatttaat acgcgataga aaacaaaata	2880
tagcgcgcaaa actaggataaa attatcgccgc gcgggtgtcat ctatgttact agat	2934

<210> 31

397

<212> PRT

<213> Rotavirus

<400> 31

Met Glu Val Leu Tyr Ser Leu Ser Lys Thr Leu Lys Asp Ala Arg Asp

1 5 10 15

Lys Ile Val Glu Gly Thr Leu Tyr Ser Asn Val Ser Asp Leu Ile Gln

20 25 30

Gln Phe Asn Gln Met Ile Val Thr Met Asn Gly Asn Asp Phe Gln Thr

35 40 45

Gly Gly Ile Gly Asn Leu Pro Val Arg Asn Trp Thr Phe Asp Phe Gly

50 55 60

Leu Leu Gly Thr Thr Leu Leu Asn Leu Asp Ala Asn Tyr Val Glu Asn

65 70 75 80

Ala Arg Thr Ile Ile Glu Tyr Phe Ile Asp Phe Ile Asp Asn Val Cys

85 90 95

Met Asp Glu Met Ala Arg Glu Ser Gln Arg Asn Gly Val Ala Pro Gln

100 105 110

Ser Glu Ala Leu Arg Lys Leu Ala Gly Ile Lys Phe Lys Arg Ile Asn

115 120 125

Phe Asp Asn Ser Ser Glu Tyr Ile Glu Asn Trp Asn Leu Gln Asn Arg

130 135 140

Arg Gln Arg Thr Gly Phe Val Phe His Lys Pro Asn Ile Phe Pro Tyr

145 150 155 160

Ser Ala Ser Phe Thr Leu Asn Arg Ser Gln Pro Met His Asp Asn Leu

165 170 175

Met Gly Thr Met Trp Leu Asn Ala Gly Ser Glu Ile Gln Val Ala Gly

180 185 190

Phe Asp Tyr Ser Cys Ala Ile Asn Ala Pro Ala Asn Ile Gln Gln Phe

195 200 205

Glu His Ile Val Gln Leu Arg Arg Ala Leu Thr Thr Ala Thr Ile Thr

210 215 220

Leu Leu Pro Asp Ala Glu Arg Phe Ser Phe Pro Arg Val Ile Asn Ser

225 230 235 240

Ala Asp Gly Ala Thr Thr Trp Phe Phe Asn Pro Val Ile Leu Arg Pro

245 250 255

Asn Asn Val Glu Val Glu Phe Leu Leu Asn Gly Gln Ile Ile Asn Thr

260 265 270

Tyr Gln Ala Arg Phe Gly Thr Ile Ile Ala Arg Asn Phe Asp Ala Ile

275 280 285

Arg Leu Leu Phe Gln Leu Met Arg Pro Pro Asn Met Thr Pro Ala Val

290 295 300

Asn Ala Leu Phe Pro Gln Ala Gln Pro Phe Gln His His Ala Thr Val

305 310 315 320

Gly Leu Thr Leu Arg Ile Glu Ser Ala Val Cys Glu Ser Val Leu Ala

325	330	335
Asp Ala Asn Glu Thr Leu Leu Ala Asn Val Thr Ala Val Arg Gln Glu		
340	345	350
Tyr Ala Ile Pro Val Gly Pro Val Phe Pro Pro Gly Met Asn Trp Thr		
355	360	365
Glu Leu Ile Thr Asn Tyr Ser Pro Ser Arg Glu Asp Asn Leu Gln Arg		
370	375	380

Val Phe Thr Val Ala Ser Ile Arg Ser Met Leu Ile Lys

385	390	395
-----	-----	-----

<210> 32

<211> 2934

<212> DNA

<213> Artificial sequence

<220><223> Expression cassette number 1714

<400> 32

gtcaacatgg tggagcacga cacacttgc tactccaaaa atatcaaaga tacagtctca	60
gaagacccaaa gggcaattga gactttcaa caaaggtaa tatccggaaa ctcctcgga	120
ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc	180
tacaaatgcc atcattgcga taaaggaaag gccatcggt aagatgcctc tgccgacagt	240

ggtccccaaag atggaccccc acccacgagg agcatcggtt aaaaagaaga cgttccaacc	300
acgtttcaa agcaagtggg ttgatgtat aacatgggtt agcacgacac acttgtctac	360
tccaaaata tcaaagatac agtctcagaa gaccaaggaa caattgagac tttcaacaa	420
aggtaatat ccggaaacct ctcggattt cattggccag ctatctgtca ctttatttg	480
aagatagtgg aaaaaggaaagg tggctctac aaatgccatc attgcgataa aggaaaggcc	540
atcggtttagt atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagc	600
atcggtggaa aagaagacgt tccaaaccacg tcttcaaagc aagtggattt atgtgatatac	660

tccactgacg taaggatga cgcacaatcc cactatcctt cgcaagaccc ttccctata	720
taaggaagtt catttcattt ggagaggat taaaatctta ataggtttt ataaaagcga	780
acgtggggaa acccgaaacca aaccttctt taaactctt ctcatcttc ttaaagcaaa	840
cttctctttt gtctttctt cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac	900
cagtacaacg tttctttca ctgaagcgaa atcaaagatc tctttgttga cacgtatgc	960
ggccatca aataacgtgt acttgccta ttcttgcgg tgtggcttg ggaaaagaaa	1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgccctgtact tctttttct	1140
tcttcttgc gattggttct ataagaaatc tagtattttc tttgaaacag agtttcccg	1200
tggtttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcg	1260
gccccatggag gtccttata gtctctccaa aacgctgaag gacgcttaggg acaagatcgt	1320
ggagggtaca ctttatagca atgtcagcga cctaatacag cagtttaatc aaatgatcgt	1380
tacaatgaat gggaatgatt tccaaactgg cggtatttgtt aatctgccc tgaggaactg	1440
gacattcgat ttccggctgc tggcacgac tctccttaat ctgcgtgaa attatgtaga	1500
aaacgccaga acgattatcg agtactttat cgatttcatt gataacgttt gtatggatga	1560
gatggccgc gagtccacaac ggaacggagt tgctccacag tccgaggccc ttccgaaact	1620
cgcggcatt aagtcaagc gtattaattt cgacaactcc tccgaatata tagagaactg	1680
gaacttgcag aatcgctcagc agagaaccgg cttcgtgtc cataaaccta atatcttcc	1740
gtatagcgc tcattcaccc tgaataggag tcagcccatg cacgacaacc tcatgggtac	1800
aatgtggctg aatgcgggaa gtgaaataca ggtcgccggg ttcgattact cctgtgccat	1860
taatgcaccc gcaaacaatcc agcagttcga acatatcgta caactaagac gggctctcac	1920
gaccgcgaca attacactcc tgcccgacgc cgagcgttc tcctttcccc gcgtaatcaa	1980
ctcagctgat ggcgcacca ctgggttctt caaccctgtt atattgcgc ctaacaacgt	2040
agaggtggag ttctcttaa acggacagat catcaatacc tbeccaagcca gtttcggcac	2100
gattattgca agaaatttcg acgctatcag gctgcttcc caactgtatga ggcggccaa	2160
tatgactccc gctgtgaacg cttgtttcc gcaggctcag ctttccagc accacgcac	2220
cgtcggttg actcttcgaa tagagagcgc ggtctgcgaa tcagtgtgg cagacgcca	2280
cgagacgctg ctggcaaacg ttaccggcgt gcccgaagag tatgccatcc cagtagggcc	2340
tgtgtttcca cccggcatga actggactga actaattact aactatagcc catccagaga	2400
agacaacttg caggggtct tcaactgtggc ctctatccgg agtatgtga tcaagtagag	2460
gcctatttc tttagttga atttactgtt attcgggtgt catttctatg tttggtgagc	2520
ggtttctgt gtcagagtg tgtttatttt atgttaattt atttcttgt gagctcctgt	2580
tttagcaggc tcgtccatcc caaggacaca aaaagattt aattttatta aaaaaaaaaa	2640
aaaaaaaagac cggaaattcg atatcaagct tatcgacccg cagatcgcc aaacattgg	2700
caataaaagtt tcttaagatt gaatcctgtt gccggcttg cgatgattat catataattt	2760

ctgttgaatt acgttaagca tctaataatt aacatgtaat gcatgacgtt atttatgaga	2820
tgggtttta tgatttagt cccgcaatta tacatthaat acgcgataga aaacaaaata	2880
tagcgcgcaa actaggataa attatcgccc gcgggtgtcat ctatgttact agat	2934
<210> 33	
<211> 53	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer IF-Rtx_VP4(opt).s1+3c	
<400> 33	
aaatttgtcg ggccatggc tagcctgatc tacagacaac tcttgaccaa ttc	53
<210> 34	
<211> 55	
<212> DNA	
<213> Artificial sequence	
<220><223> Primer IF-Rtx_VP4(opt).s1-4r	
<400> 34	
actaaagaaa ataggccttc agagttaca ttgcaggatt aattgctcaa tccta	55
<210> 35	
<211> 4068	
<212> DNA	
<213> Artificial sequence	
<220><223> Expression cassette number 1731	
<400> 35	
gtcaacatgg tggagcacga cacacttgc tactccaaa atatcaaaga tacagtctca	60
gaagaccaaa gggcaattga gactttcaa caaaggtaa tatccgaaa ctcctcgga	120
ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc	180
tacaaatgcc atcattgcga taaaggaaag gccatcggt aagatgcctc tgccgacagt	240
ggtcccaaag atggaccccc acccacgagg agcatcggtt aaaaagaaga cgttccaacc	300
acgtcttcaa agcaagtgga ttgatgtat aacatggtg agcacgacac acttgtctac	360
tccaaaaata tcaaagatac agtctcagaa gaccaaggaa caattgagac tttcaacaa	420
aggtaatat ccggaaacct cctcggtt cattgcccag ctatctgtca ctattttgt	480
aagatagtgg aaaaggaaagg tggctctac aaatgcctc attgcataa aggaaaggcc	540
atcggttgaag atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagc	600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatata	660
tccactgacg taaggatga cgcacaatcc cactatcctt cgcaagaccc ttccctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta atagtttg ataaaagcga	780
acgtgggaa acccgAACCA aaccccttc taaactctc ctcatctc ttaaagcaaa	840
cttctctttt gtctttctt cgtgagcgtt cttaacgtt gtcagatcgt gcttcggcac	900
cagtacaacg tttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtatgc	960
ggccattt aataacgtgt acttgtccta ttcttgtcg tggtgttttgggaaagaaaa	1020
gtttgtgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgc gaggtattgt tgccctgtact tctttttct	1140
tcttcttgc gattgggtct ataagaaatc tagtattttc tttgaaacag agttttcccg	1200
tggtttgcgaa acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcg	1260
gcccattggct accctgatct acagacaact ctgaccaat tcatttctg tggatctca	1320
tgacgaaatc gaggcattt ggtccgagaa gacccagaac gtgaccatca accctggacc	1380
tttgctcag acccgctatg cccctgtgaa ttggatcac ggagaaatca acgacagtac	1440
gaccgtcgaa cccattctgg acggccata ccaacccacc accttcaccc cacctaata	1500
ttatggatt ttaatcaact ccaacacaaa cggagtggtc tacgagtcca ctaataactc	1560
cgatTTTgg accggcgttg tagccatcgaa gccacacgtc aatcctgtcg atcggcagta	1620
tatgatattc ggcgagtcca aacagttaa cgtttcaat gacagcaaca aatgaaagtt	1680
tctggagatg ttcccgact cctctcagaa cgaattctat aatagacgga cccttacctc	1740
cgtacacga ctcgtggta ttttaagta cggcggcagg gtgtggacat ttcacggta	1800
aaccctcga gcaaccactg actccagtag cactgaaac ctgaacaata tatctattac	1860
catccacagc gaattctaca taatccaaat atctcaggaa agtaagtgtacgaaatata	1920
caacaacgga ctcccccaa ttcagaatac acggAACGTG gtgcctctcc cactcagtcc	1980
tcggctatc cagtataaga gagcacaagt gaatgaggac attattgtga gcaagactag	2040
ccttggaaa gaaatgcagt acaacagaga cattatcatc cggttaagt ttggaaactc	2100
tatcgtgaag atggcggcc tgggtacaa atggcagaa atctcatata aagccgcaaa	2160
ctatcgtat aactacttga gagacggcga gcaggttaacc gcccacacaa catgtctgt	2220
caacggcggtt aataacttta gctacaacgg aggcttcctt cccacccact tcggtatcag	2280
ccggatgaa gtcataagg aaaattctta tgtgtacgtt gattactggg atgatagcaa	2340

agcgttccgc aacatggtgt atgttaggag cctggctgct aatctaatt ctgtgaagt	2400
tactgggttga tcatattatt tctcaattcc cgtggggct tggccagtc tgaatggcg	2460
ggcagtcctcc ctccattttgc ttggcgtgac gttgagact cagttaccg atttcgtgtc	2520
tctgaactcc ctgagggttcc gttttccct tactgtcgac gagccccat tcagcattct	2580
gcgtacaaga actgtcaacc tctacgggtt acctgccg aatccaaaca acggcaatga	2640
atactatgaa atttcgggcc gtttctttt gataagtctg gtaccaacta atgacgacta	2700
tcagacaccc atcatgaaca gcgtgactgt cagacaggac ctggaaagac aacttacaga	2760
tctgcggaa gaattcaatt ctctcagtca ggagattgca atggccaat tgatagatct	2820
tgccctactg ccctcgata ttttttagtat gtttccggc atcaaatcaa ctatagatct	2880
gacaaagagc atggctactt ctgtgtgaa gaagttcagg aaatcaaacc ttgccacgag	2940
catatcagaa atgacgaact ctctgagtta tgcatcgatca tcagcgtcac gcaacgtttc	3000
cattcggctcg aatctcagcg ccatcagcaa ctggacaaac gtgtccaacg acgtcagcaa	3060
cgtgaccaac tccttgaacg atatttctac ccagacgtca acgatcagta agaaactccg	3120
cttggaaagaa atgatcaccc agactgaggg aatgttttc gacgacattt ccggcccg	3180
gctaaaaacc aaaatcgata tgtctactca gatcgcaag aacactctgc cgatatcg	3240
aaccgaagcc tcggaaaagt ttatccctaa gcgcagctac agaatattga aagatgacga	3300
ggtcatggag atcaacacag aaggaaagtt cttcgcttat aagatcaaca ccttgacga	3360
ggttccgttt gacgtcaata agtttgcaga gtcgtgaca gatagttccag tgatctgc	3420
catcattgac tttaagactt tgaagaacct gaacgacaac tatggataaa cacggaccga	3480
agcgttgaac ctcattaagt ccaatccaa tatgttgccc aatttcattt accagaacaa	3540
tccaatcata agaaatagga ttgagcaatt aatcctgcaaa tgtaaactct gaaggcctat	3600
tttcttagt ttgaatttac ttttattcgg tgtgcatttc tatgttttgtt gagcggttt	3660
ctgtgctcag agtgtgtta ttttatgtaa tttaatttct ttgtgagctc ctgttttagca	3720
ggtcgtccct tcagcaagga cacaaaaaaga ttttaattttt attaaaaaaaaaaaaaaa	3780
agaccggaa ttgcataatca agcttatacg cctgcagatc gttcaaacat ttggcaataa	3840
agtttcttaa gattgaatcc ttttgccgtt cttgcgtatca ttatcatata atttctgtt	3900
aattacgtta agcatgtat aattaacatg taatgcata cgttattttt gagatggtt	3960
tttatgatta gagtcccgca attatacatt taatacgcga tagaaaacaa aatatagcgc	4020
gcaaaactagg ataaatttac gcgccgggtg tcatctatgt tactagat	4068

<210> 36

<211> 775

<212> PRT

<213> Rotavirus

<400> 36

Met Ala Ser Leu Ile Tyr Arg Gln Leu Leu Thr Asn Ser Tyr Ser Val

1 5 10 15

Asp Leu His Asp Glu Ile Glu Gln Ile Gly Ser Glu Lys Thr Gln Asn

20 25 30

Val Thr Ile Asn Pro Gly Pro Phe Ala Gln Thr Arg Tyr Ala Pro Val

35 40 45

Asn Trp Asp His Gly Glu Ile Asn Asp Ser Thr Thr Val Glu Pro Ile

50 55 60

Leu Asp Gly Pro Tyr Gln Pro Thr Thr Phe Thr Pro Pro Asn Asp Tyr

65 70 75 80

Trp Ile Leu Ile Asn Ser Asn Thr Asn Gly Val Val Tyr Glu Ser Thr

85 90 95

Asn Asn Ser Asp Phe Trp Thr Ala Val Val Ala Ile Glu Pro His Val

100 105 110

Asn Pro Val Asp Arg Gln Tyr Met Ile Phe Gly Glu Ser Lys Gln Phe

115 120 125

Asn Val Ser Asn Asp Ser Asn Lys Trp Lys Phe Leu Glu Met Phe Arg

130 135 140

Ser Ser Ser Gln Asn Glu Phe Tyr Asn Arg Arg Thr Leu Thr Ser Asp

145 150 155 160

Thr Arg Leu Val Gly Ile Phe Lys Tyr Gly Gly Arg Val Trp Thr Phe

165 170 175

His Gly Glu Thr Pro Arg Ala Thr Thr Asp Ser Ser Ser Thr Ala Asn

180 185 190

Leu Asn Asn Ile Ser Ile Thr Ile His Ser Glu Phe Tyr Ile Ile Pro

195 200 205

Arg Ser Gln Glu Ser Lys Cys Asn Glu Tyr Ile Asn Asn Gly Leu Pro

210 215 220

Pro Ile Gln Asn Thr Arg Asn Val Val Pro Leu Pro Leu Ser Ser Arg

225	230	235	240
-----	-----	-----	-----

Ser Ile Gln Tyr Lys Arg Ala Gln Val Asn Glu Asp Ile Ile Val Ser

245	250	255
-----	-----	-----

Lys Thr Ser Leu Trp Lys Glu Met Gln Tyr Asn Arg Asp Ile Ile Ile

260	265	270
-----	-----	-----

Arg Phe Lys Phe Gly Asn Ser Ile Val Lys Met Gly Gly Leu Gly Tyr

275	280	285
-----	-----	-----

Lys Trp Ser Glu Ile Ser Tyr Lys Ala Ala Asn Tyr Gln Tyr Asn Tyr

290	295	300
-----	-----	-----

Leu Arg Asp Gly Glu Gln Val Thr Ala His Thr Thr Cys Ser Val Asn

305	310	315	320
-----	-----	-----	-----

Gly Val Asn Asn Phe Ser Tyr Asn Gly Gly Phe Leu Pro Thr Asp Phe

325	330	335
-----	-----	-----

Gly Ile Ser Arg Tyr Glu Val Ile Lys Glu Asn Ser Tyr Val Tyr Val

340	345	350
-----	-----	-----

Asp Tyr Trp Asp Asp Ser Lys Ala Phe Arg Asn Met Val Tyr Val Arg

355	360	365
-----	-----	-----

Ser Leu Ala Ala Asn Leu Asn Ser Val Lys Cys Thr Gly Gly Ser Tyr

370	375	380
-----	-----	-----

Tyr Phe Ser Ile Pro Val Gly Ala Trp Pro Val Met Asn Gly Gly Ala

385	390	395	400
-----	-----	-----	-----

Val Ser Leu His Phe Ala Gly Val Thr Leu Ser Thr Gln Phe Thr Asp

405	410	415
-----	-----	-----

Phe Val Ser Leu Asn Ser Leu Arg Phe Arg Phe Ser Leu Thr Val Asp

420	425	430
-----	-----	-----

Glu Pro Pro Phe Ser Ile Leu Arg Thr Arg Thr Val Asn Leu Tyr Gly

435	440	445
-----	-----	-----

Leu Pro Ala Ala Asn Pro Asn Asn Gly Asn Glu Tyr Tyr Glu Ile Ser

450	455	460
-----	-----	-----

Gly Arg Phe Ser Leu Ile Ser Leu Val Pro Thr Asn Asp Asp Tyr Gln

465	470	475	480
-----	-----	-----	-----

Thr Pro Ile Met Asn Ser Val Thr Val Arg Gln Asp Leu Glu Arg Gln

485 490 495

Leu Thr Asp Leu Arg Glu Glu Phe Asn Ser Leu Ser Gln Glu Ile Ala

500 505 510

Met Ala Gln Leu Ile Asp Leu Ala Leu Leu Pro Leu Asp Met Phe Ser

515 520 525

Met Phe Ser Gly Ile Lys Ser Thr Ile Asp Leu Thr Lys Ser Met Ala

530 535 540

Thr Ser Val Met Lys Lys Phe Arg Lys Ser Lys Leu Ala Thr Ser Ile

545 550 555 560

Ser Glu Met Thr Asn Ser Leu Ser Asp Ala Ala Ser Ser Ala Ser Arg

565 570 575

Asn Val Ser Ile Arg Ser Asn Leu Ser Ala Ile Ser Asn Trp Thr Asn

580 585 590

Val Ser Asn Asp Val Ser Asn Val Thr Asn Ser Leu Asn Asp Ile Ser

595 600 605

Thr Gln Thr Ser Thr Ile Ser Lys Lys Leu Arg Leu Lys Glu Met Ile

610 615 620

Thr Gln Thr Glu Gly Met Ser Phe Asp Asp Ile Ser Ala Ala Val Leu

625 630 635 640

Lys Thr Lys Ile Asp Met Ser Thr Gln Ile Gly Lys Asn Thr Leu Pro

645 650 655

Asp Ile Val Thr Glu Ala Ser Glu Lys Phe Ile Pro Lys Arg Ser Tyr

660 665 670

Arg Ile Leu Lys Asp Asp Glu Val Met Glu Ile Asn Thr Glu Gly Lys

675 680 685

Phe Phe Ala Tyr Lys Ile Asn Thr Phe Asp Glu Val Pro Phe Asp Val

690 695 700

Asn Lys Phe Ala Glu Leu Val Thr Asp Ser Pro Val Ile Ser Ala Ile

705 710 715 720

Ile Asp Phe Lys Thr Leu Lys Asn Leu Asn Asp Asn Tyr Gly Ile Thr

725	730	735
Arg Thr Glu Ala Leu Asn Leu Ile Lys Ser Asn Pro Asn Met Leu Arg		
740	745	750

Asn Phe Ile Asn Gln Asn Asn Pro Ile Ile Arg Asn Arg Ile Glu Gln		
755	760	765
Leu Ile Leu Gln Cys Lys Leu		
770	775	

<210> 37		
<211> 51		
<212> DNA		
<213> Artificial sequence		

<220><223> Primer IF-Rtx_VP7(opt).s1+3c

<400> 37

aaatttgtcg gccccatgta cggcatcgag tataacaacaa ttttaattt c	51
--	----

<210> 38

<211> 54

<212> DNA

<213> Artificial sequence

<220><223> Primer IF-Rtx_VP7(opt).s1-4r

<400> 38

actaaagaaa ataggcctct aaacgcgata atagaaggct gctgagttca ggaa	54
---	----

<210> 39

<211> 326

<212> PRT

<213> Rotavirus

<400> 39

Met Tyr Gly Ile Glu Tyr Thr Thr Ile Leu Ile Phe Leu Ile Ser Ile

1	5	10	15
---	---	----	----

Ile Leu Leu Asn Tyr Ile Leu Lys Ser Val Thr Arg Ile Met Asp Tyr

20	25	30
----	----	----

Ile Ile Tyr Arg Ser Leu Leu Ile Tyr Val Ala Leu Phe Ala Leu Thr

35	40	45
----	----	----

Arg Ala Gln Asn Tyr Gly Leu Asn Leu Pro Ile Thr Gly Ser Met Asp

50	55	60
Thr Val Tyr Ala Asn Ser Thr Gln Glu Gly Ile Phe Leu Thr Ser Thr		
65	70	75
Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ser Thr Gln Ile Asn Asp Gly		
85	90	95
Glu Trp Lys Asp Ser Leu Ser Gln Met Phe Leu Thr Lys Gly Trp Pro		
100	105	110
Thr Gly Ser Val Tyr Phe Lys Glu Tyr Ser Ser Ile Val Asp Phe Ser		
115	120	125
Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn Leu Val Leu Met Lys Tyr		
130	135	140
Asp Gln Asn Leu Glu Leu Asp Met Ser Glu Leu Ala Asp Leu Ile Leu		
145	150	155
Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr Tyr Tyr Gln		
165	170	175
Gln Ser Gly Glu Ser Asn Lys Trp Ile Ser Met Gly Ser Ser Cys Thr		
180	185	190
Val Lys Val Cys Pro Leu Asn Thr Gln Met Leu Gly Ile Gly Cys Gln		
195	200	205
Thr Thr Asn Val Asp Ser Phe Glu Met Val Ala Glu Asn Glu Lys Leu		
210	215	220
Ala Ile Val Asp Val Val Asp Gly Ile Asn His Lys Ile Asn Leu Thr		
225	230	235
Thr Thr Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly Pro Arg Glu		
245	250	255
Asn Val Ala Val Ile Gln Val Gly Gly Ser Asn Val Leu Asp Ile Thr		
260	265	270
Ala Asp Pro Thr Thr Asn Pro Gln Thr Glu Arg Met Met Arg Val Asn		
275	280	285
Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Ile Val Asp Tyr Ile Asn		

290	295	300	
Gln Ile Val Gln Val Met Ser Lys Arg Ser Arg Ser Leu Asn Ser Ala			
305	310	315	320
Ala Phe Tyr Tyr Arg Val			
325			
<210> 40			
<211> 60			
<212> DNA			
<213> Artificial sequence			
<220><223> Primer IF-Rtx_VP7(opt).s2+4c			
<400> 40			
tctcagatct tcgcccagaa ctatggcctg aacttaccaa tcaccggttc aatggatacc			60
<210> 41			
<211> 4897			
<212> DNA			
<213> Artificial sequence			
<220><223> Construct 1192			
<400> 41			
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg			60
gacgtttta atgtactgaa ttaacgccga atccggct ggtatattta tatgttgtca			120
aataactcaa aaaccataaa agttaagt agcaagtgt tacattttt cttgaacaaa			
aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg			
ataagaacaa gagtagtgat atttgacaa caattttgtt gcaacatttg agaaaatttt			
gttggctct ctttcattt gtcaaaaaca atagagagag aaaaaggaag agggagaata			
aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaa atagttgtac			
aaatatcatt gaggaatttg aaaaaagcta cacaataag ggttaattgc tgtaaataaa			
taaggatgac gcatttagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga			
aagaataaat tattttaaa attaaaagtt gagtcatttg attaacatg tgattattta			
atgaattgtt gaaagagttt gattaaagtt gtattagtaa tttagatttg gtgtcaaatt			
taatttgaca ttgtatctt tcctatata tggcccatag agtcagttaa ctcattttt			
tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg			
gtatatttac taaaaaatct aagccacgta ggaggataac aggtccccg taggaggata			
600			
660			
720			
780			
840			

acatccaatc caaccaatca caacaatct gatgagataa cccacttaa gcccacgcat	900
ctgtggcaca tctacattat ctaaatcaca cattttcca cacatcttag ccacacaaaa	960
accaatccac atcttatca cccattctat aaaaatcac acttttgag tctacactt	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggAAC gagctataca aggaaacgac gctaggAAC aagctaACAG tgaacgttgg	1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagtgg	1200
actgagtggc ggctacataa cgatgagacg aattcaatc aagataatcc cttgggttc	1260
aaggaaagct ggggttcgg gaaagtgtt ttaagagat atctcagata cgacaggacg	1320
gaagctcac tgacagagt cttggact tggacggag attcggttaa ctatgcagca	1380
tctcgattt tcggttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt	1440
agtatcaccg ttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gttgccccaa atcgaagtgg aaagtaatgt atcaagagga	1560
tgcctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gtttctcgt	1620
ctcctattta taatatggtt tttttttttt aattttgtt ttgttagaaga gcttaattaa	1680
tcgttgttgt tatgaaatac tatttgatg agatgaactg gtgtatgtt attcattac	1740
ataagtggag tcagaatcag aatgtttctt ccataactaa cttagacatga agacctgccc	1800
cgtacaattt tcttatattt gaacaactaa aatttgcacat ctttgcac aactttataa	1860
gtggtaata tagctcaat atatggtaa gttcaataga ttaataatgg aaatatcagt	1920
tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccc	1980
ttgataatg atagtagacacc aatttaggaag gagcatgctc gcctaggaga ttgtcggttc	2040
ccgccttcag ttgcagact gctctagccg tttttttttt acgcaaacccg cctctcccg	2100
cgcgttggaa attactagcg cgtgtcgaca agcttgcattt ccggtaaca tggggagca	2160
cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaaggcaat	2220
tgagactttt caacaaaggg taatatccgg aaaccttcctc ggattccatt gcccagctat	2280
ctgtcactttt attgtgaaga tagtgaaaaa ggaagggtggc tcctacaat gccatcattg	2340
cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggccccaa aagatggacc	2400
cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt	2460
ggattgtatgtt gataacatgg tggagcacga cacacttgc tactccaaaa atatcaaaga	2520
tacagtctca gaagacaaaaa gggcaattga gactttcaa caaaggtaa tatccggaaa	2580

cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tgaaaagga	2640
aggggctcc tacaaatgcc atcattgcga taaaggaaag gccatcggt aagatgcctc	2700
tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcggtt aaaaagaaga	2760
cgttccaacc acgtttcaa agcaagtggta ttgatgttat atctccactg acgtaaggga	2820
tgacgcacaa tcccactatc cttcgcaaga cccttcctt atataaggaa gttcattca	2880
tttggagagg tattaaaatc ttaatagtt ttgataaaag cgaacgtgg gaaacccgaa	2940
ccaaaccttc ttctaaactc tctctcatct ctcttaaagc aaacttctt cttgttttc	3000
ttgcgtgagc gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgtttctt	3060
tcactgaagc gaaatcaaag atcttttgt ggacacgttag tgccgcgccca ttaaataacg	3120
tgtacttgc ctattttgtt cggtgtggc ttggaaaag aaagctgtt ggaggctgt	3180
gttcagcccc atacattact ttttacgatt ctgctgactt tcggcgggtt caatatctt	3240
acttctgtt gacgaggtat ttttgcgtt acttctttct ttttcttctt gctgattgg	3300
tctataagaa atctagtatt ttctttgaaa cagagtttc cctgtgtttt cgaacttgg	3360
gaaagattgt taagttctg tatattctgc ccaaattttt cggcccatg gcgaaaaacg	3420
ttgcgatttt cggttattt ttttcttctt ttgtgttggt tccttcctcgtt atcttcggc	3480
cggctcctca gccaaaacga caccccatc tgtctatcca ctggccctg gatctgtgc	3540
ccaaactaac tccatggtga ccctggatg cctggtaag ggctatttcc ctgagccagt	3600
gacagtgacc tggactctg gatccctgtc cagcgggtgtt cacaccccttcc cagctgtctt	3660
gcagcttgac ctctacactc tgagcagtc agtactgtc ccctccagca cctggccctt	3720
cgagaccgtc acctgcaacg ttgcccaccc ggccagcagc accaagggtgg acaagaaaat	3780
tgtgcccagg gattgtggtt gtaaggctt catatgtaca gtccagaag tatcatctgt	3840
cttcatcttc ccccaaagc ccaaggatgt gtcaccattt actctgactc ctaaggtcac	3900
gtgtgttgtt gtagacatca gcaaggatga tcccggatc cagttcagct gggttgtttaga	3960
tgtatgtggag gtgcacacag ctcagacgca accccggag gagcagttca acagcacttt	4020
ccgctcagtc agtgaacttc ccatcatgca ccaggactgg ctcaatggca aggagcgatc	4080
gctcaccatc accatcacca tcaccatcac cattaaaggc ctatccctt tagttgaat	4140
ttactgttat tcgggtgtca tttctatgtt tggtgagcgg tttctgtgc tcagagtgt	4200
tttattttt gtaatttaat ttctttgtga gtcctgtttt agcaggatgtt cccttcagca	4260
aggacacaaa aagatttaa ttttattaaa aaaaaaaaaaaa aaaaagaccg ggaattcgat	4320

atcaagctta tcgacctgca gatcggtcaa acattggca ataaaggttc ttaagattga	4380
atccgttgc cggcttgcg atgattatca tataatttct gttgaattac gtttagcatg	4440
taataattaa catgtaatgc atgacgttat ttatgagatg ggttttatg attagagtcc	4500
cgcaattata cattaatac gcgatagaaa acaaaatata ggcgcacaaac taggataat	4560
tatcgcgcc ggtgtcatct atgttactag atctctagag tctcaagctt ggcgcgccca	4620
cgtgactagt ggcactggcc gtcgtttac aacgtcgta ctggaaaac cctggcgtaa	4680
cccaacttaa tcgccttgca gcacatcccc ctttcggccag ctggcgtaat agcgaagagg	4740
cccgccaccga tcgccttcc caacagttgc gcagcctgaa tggcgaatgc tagaggcgt	4800
tgagcttggaa tcagattgtc gtttcccggcc ttcaattaa actatcgtt tttgacagga	4860
tatattggcg ggttaaaccta agagaaaaga gcgttta	4897
<210> 42	
<211> 2643	
<212> DNA	
<213> Artificial sequence	
<220><223> Expression cassette number 1735	
<400> 42	
gtcaacatgg tggagcacga cacacttgc tactccaaaa atatcaaaga tacagtctca	60
gaagacccaaa gggcaattga gactttcaa caaaggtaa tatccggaaa ctcctcgga	120
ttccattgcc cagctatctg tcaacttatt gtgaagatag tgaaaagga aggtggctcc	180
tacaaatgcc atcattgcga taaaggaaag gccatcggt aagatgcctc tgccgacagt	240
ggtcccaaaat atggaccccc acccacgagg agcatcggtt aaaaagaaga cgttccaacc	300
acgtcttcaa agcaagtggaa ttgatgtat aacatgggtt agcacgacac acttgtctac	360
tccaaaata tcaaagatac agtctcagaa gaccaaggaa caattgagac ttttcaacaa	420
aggtaatat ccggaaacct ctcggattt cattggccag ctatctgtca ctttattgtt	480
aagatagtgg aaaaggaagg tggctctac aaatgcctt attgcgataa aggaaaggcc	540
atcggttgaag atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagg	600
atcggtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattt atgtgtatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttccctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggtttt ataaaagcga	780
acgtggggaa acccgaacca aaccttcttca taaactctt ctcattcttc ttaaaggaaa	840
cttctctttt gtctttcttgcgtgagcgat cttcaacgtt gtcagatgtt gcttcggcac	900
cagtacaacg ttttcttca ctgaagcgaa atcaaagatc tctttgttggaa cacgtatgtc	960

ggccgcatta aataacgtgt acttgtccta ttcttgtcgg tgggtcttggaaaaaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgc gaggttattgt tgccctgtact tctttttct	1140
tcttcttgc gattgggtct ataagaatc tagtatttc tttgaaacag agtttcccg	1200
tgggtttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg	1260
gcccatggcg aaaaacgttg cgattttcgg cttattgtt tctcttcttgc ttgtgggtcc	1320
ttctcagatc ttcccaga actatggcct gaacttacca atcaccgtt caatggatac	1380
cgtttacgct aattccactc aagagggat atttctgaca agtaccctgt gcctgttata	1440
tccaacagaa gcctctaccc agatcaatga tggggagtgg aaggatagtc tctcacagat	1500
gttcctaacc aagggtggc ccaccgggtc cgtctacttc aaggaatact cttagtattgt	1560
cgacttctca gtgtaccccc agctttatttgcgactacaac ctggtactta tgaatatacg	1620
ccagaacctg gagctggata tgtccgagct ggctgacccgt atcctcaatg agtggctgt	1680
caacccatg gacatcacat tatattacta ccagcagtct ggagaatcca acaagtggat	1740
cagtagatggc tcaagttgca ccgtgaaggt gtgtcccttgc aacacccaaa tgctggcat	1800
tggtgtcag acaactaatg tggattcgaa tgaaatggta gccgaaaaacg agaagctggc	1860
tatagtggac gtatcgatg ggattaacca caagatcaat ctgactacca ccacttgtac	1920
catcagaacatgtaaaaaagc tcggcccccgggaaacgtc ggcgtgatcc aggtgggggg	1980
gagcaatgttctcgacatta ctggcggaccc taccaccaat ccacagacgg aacggatgt	2040
gagagtcaac tggaaagaaat ggtggcaggt cttttatacc attgtggact acattaacca	2100
gattgtgcaatgtcatgatgaaacgggtccag atccctgaac tcagcagcct tctattatcg	2160
cgttagagg cttttttctttagttgaa ttactgttaccaat ttcgggtgtc atttctatgt	2220
ttgggtgagcg gtttctgtcctcagatgt gtttattttatgtaaatggtaatggtaatggta	2280
agctcctgtt tagcaggtcg tcccttcagc aaggacacaa aaagatttatattttat	2340
aaaaaaaaaaaaaaa aaaaaagacc gggaaatcga tatcaagttt atcgacactgc agatcgatca	2400
aacatttggc aataaagttt cttaaagattt aatcctgttgc cgggtttgc gatgattatc	2460
atataatttc tggtaatttgcgtttaagcat gtaataatcatgtatgttacatgttacatgtt	2520
tttatgatgat gggttttat gattagatgc ccgcaattat acatttataa cgcgatagaa	2580
aacaaaatat agcgcgcggaaat ctaggataaa ttatcgccgc cgggtgtcatc tatgttactat	2640
gat	2643

<210> 43

<211> 300

<212> PRT

<213> Rotavirus

<400> 43

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val

1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Asn Tyr Gly Leu Asn Leu Pro

20 25 30

Ile Thr Gly Ser Met Asp Thr Val Tyr Ala Asn Ser Thr Gln Glu Gly

35 40 45

Ile Phe Leu Thr Ser Thr Leu Cys Leu Tyr Tyr Pro Thr Glu Ala Ser

50 55 60

Thr Gln Ile Asn Asp Gly Glu Trp Lys Asp Ser Leu Ser Gln Met Phe

65 70 75 80

Leu Thr Lys Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Ser

85 90 95

Ser Ile Val Asp Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn

100 105 110

Leu Val Leu Met Lys Tyr Asp Gln Asn Leu Glu Leu Asp Met Ser Glu

115 120 125

Leu Ala Asp Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile

130 135 140

Thr Leu Tyr Tyr Tyr Gln Gln Ser Gly Glu Ser Asn Lys Trp Ile Ser

145 150 155 160

Met Gly Ser Ser Cys Thr Val Lys Val Cys Pro Leu Asn Thr Gln Met

165 170 175

Leu Gly Ile Gly Cys Gln Thr Thr Asn Val Asp Ser Phe Glu Met Val

180 185 190

Ala Glu Asn Glu Lys Leu Ala Ile Val Asp Val Val Asp Gly Ile Asn

195 200 205

His Lys Ile Asn Leu Thr Thr Thr Cys Thr Ile Arg Asn Cys Lys

210	215	220
Lys Leu Gly Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser		
225	230	235
Asn Val Leu Asp Ile Thr Ala Asp Pro Thr Thr Asn Pro Gln Thr Glu		

245	250	255
Arg Met Met Arg Val Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr		
260	265	270
Ile Val Asp Tyr Ile Asn Gln Ile Val Gln Val Met Ser Lys Arg Ser		
275	280	285
Arg Ser Leu Asn Ser Ala Ala Phe Tyr Tyr Arg Val		

290	295	300
-----	-----	-----

<210> 44

<211> 4068

<212> DNA

<213> Artificial sequence

<220><223>

Expression cassette number 1730

<400> 44

gtcaacatgg tggagcacga cacacttgc tactccaaaa atatcaaaga tacagtctca	60
gaagacaaaa gggcaattga gactttcaa caaaggtaa tatccgaaa cctcctcgaa	120
ttccattgcc cagctatctg tcactttatt gtgaagatag tgaaaagga aggtggctcc	180
tacaaatgcc atcattgcga taaaggaaag gccatcggt aagatgcctc tgccgacagt	240
ggtcccaaag atggaccccc acccacgagg agcatcggtt aaaaagaaga cgttccaacc	300
acgtttcaa agcaagtggta ttgatgttat aacatggtg agcacgacac acttgttac	360
tccaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac tttcaacaa	420

aggtaatat ccggaaacct cctcgatttc cattgcccag ctatctgtca ctttatttg	480
aagatagtgg aaaaggaagg tggctctac aaatgccatc attgcgataa aggaaaggcc	540
atcggtgaag atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagc	600
atcggtggaa aagaagacgt tccaaccacg tcttcaaagc aagtggattt atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttccctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggtttt ataaaagcga	780

acgtgggaa acccgaacca aacttttc taaactctt ctcatctc ttaaagcaa	840
cttcctctt gtcttcttg cgtgagcgat cttcaacgtt gtcagatgt gcttcggcac	900
cagtagacaacg ttttttca ctgaagcgaa atcaaagatc tctttgtgga cacgtatgc	960
ggccattttt aataacgtgt acttgtccta ttctgtcggt tggtgttgg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctactact tctgcttgcac gaggtattgt tgctgtact tctttttct	1140
tcttcttgct gattggttct ataagaaatc tagtattttc ttgaaacag agttttcccg	1200
tggtttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcg	1260
gcccatggct agcctgatct acagacaact cttgaccaat tcataattctg tggatcttca	1320
tgacgaaatc gagcaggattt ggtccggagaa gaccggaaac gtgaccatca accctggacc	1380
ttttgctcag acccgctatg cccctgtgaa ttggatcac ggagaaatca acgacagtac	1440
gaccgtcgaa cccattctgg acggggccata ccaacccacc accttcaccc cacctaata	1500
ttattggatt ttaatcaact ccaacacaaa cggagtggtc tacgagtcca ctaataactc	1560
cgattttgg accggccgtt tagccatcga gccacacgtc aatcctgtcg atcggcagta	1620
tatgatattc ggcgagtc aacagttaa cgtttcaat gacagcaaca aatgaaatgtt	1680
tctggagatg ttgcagact cctctcagaat cgaattctat aatagacgga cccttaccc	1740
cgtatcacga ctctggta ttttaagta cggcggcagg gtgtggacat ttacgggtga	1800
aaccctcga gcaaccactg actccagtag cactgcaac ctgaacaata tatctattac	1860
catccacagc gaattctaca taatccaaat atctcaggaa agtaagtgtt acgaatata	1920
caacaacgga ctcccccaa ttcaatac acggAACgtg gtgcctctcc cactcgttc	1980
tcggtctatc cagtataaga gaggcacaatg gaatgaggac attattgtga gcaagactag	2040
ccttggaaa gaaatgcagt acaacagaga cattatcatc cggttaatgt ttggaaactc	2100
tatcgtgaag atggcggcc tgggtacaa atggcagaa atctcatata aagccgcaaa	2160
ctatcgtat aactacttga gagacggcga gcaggtaacc gcccacacaa catgtctgt	2220
caacggcgaa aataacttta gctacaacgg aggcttcctt cccaccgact tcggatcat	2280
ccgttatgaa gtcataagg aaaatttta tgtgtacgtt gattactggg atgatagcaa	2340
agcgttccgc aacatgggtt atgttaggac cctggctgtt aatctcaatt ctgtgaagtgt	2400
tactgggtgaa tcatattatt tctcaattcc cgtggggct tggccagtc tgaatggcgg	2460
ggcagtcctcc ctccatggcgtgac gttgagactt cagtttaccg atttcgtgtc	2520

tctgaactcc ctgaggttcc gggtttccct tactgtcgac gagccccat tcagcattct	2580
gcgtacaaga actgtcaacc tctacgggtt acctgccgat aatccaaaca acggcaatga	2640
atactatgaa atttcgggcc gcttctttt gataagtctg gtaccaacta atgacgacta	2700
tcagacaccc atcatgaaca gcgtgactgt cagacaggac ctgaaagac aacttacaga	2760
tctgcggaa gaattcaatt ctctcagta ggagattgca atggccaat tgatagatct	2820
tgcctactg cctctcgata tgtttagtat gttctccggc atcaaataa ctatagatct	2880
gacaaagagc atggctactt ctgtgatgaa gaagttcagg aaatcaaaac ttgccacgag	2940
catacagaa atgacgaact ctctgagtga tgcagcatca tcagcgac gcaacgttcc	3000
cattcggtcg aatctcagcg ccatcagca ctggacaac gtgtccaacg acgtcagcaa	3060
cgtgaccaac tccttgaacg atatttctac ccagacgtca acgtcagta agaaactccg	3120
cttggaaagaa atgatcaccc agactgaggg aatgttttc gacgacattt ccggccggt	3180
gctaaaaacc aaaatcgata tgtctactca gatcgcaag aacactctgc cgatgtatcg	3240
aaccgaagcc tccgaaaagt ttatccctaa ggcgacgtac agaatattga aagatgacga	3300
ggtcatggag atcaacacag aaggaaagtt cttcgcttat aagatcaaca ctttgacga	3360
ggttccgtt gacgtcaata agttgcaga gctcgac gatagttccag tgatttctgc	3420
catcattgac ttaagactt tgaagaacctt gaacgacaac tatggataa cacggaccga	3480
agcgttgaac ctcattaagt ccaatccaa tatgttgcgc aatttcatta accagaacaa	3540
tccaatcata agaaaatagga ttgagcaatt aatccgtcaa tggaaactct gaaggctat	3600
tttcttagt ttgaatttac ttttattcgg ttttgcatttc tttttttt ggggtttt	3660
ctgtgctcag agtgtgttta ttttatgtaa tttaatttct ttgtgagctc ctgttttagca	3720
ggtcgtccct tcagcaagga cacaataaga tttaattttt attaaaaaaa aaaaaaaaaa	3780
agaccggaa ttgcataatca agcttatacgat cctgcagatc gttcaaacat ttggcaataa	3840
agtttcttaa gattgaatcc ttttgcgggtt cttgcgtatca ttatcatata atttctgttgc	3900
aattacgtta agcatgtat aatatacatg taatgcgtat cgttattttt gagatgggtt	3960
tttatgatta gagtcccgca attatacatt taatacgtatc tagaaaaacaa aatatacgcc	4020
gcaaaactagg ataaatttac ggcgcgggtt tcattatgt tactagat	4068
<210> 45	
<211> 2673	
<212> DNA	
<213> Rotavirus	
<400> 45	

atggcatacc ggaagagagg agcaaagcgc gaaaacctgc cgcaacagaa cgagagactg	60
caagaaaaag agatagagaa agatgtcgac gtaacaatgg aaaacaagaa taacaatagg	120
aaacaacagc tgtccgacaa agttctgtcc cagaaggagg aaattatcac tgacgccag	180
gacgatatta aaattgccgg agaaataaag aagagctcg aagaagaatc taaacagctg	240
ctcgaaattc tgaaaacaaa agaagaccat cagaaagaga ttcaaatatga aattttgcaa	300
aaaacaatac ctacatttga gtccaaagaa agtacctca agaagcttga agacataaga	360
ccggagcagg caaaaaaaca gatgaaactc tttcgatt tcgagcaaa acagctccct	420
atatatcgcg ccaatggcga gaaggagcta cgcaaccgtt ggtactggaa gttgaaaaaa	480
gacaccctgc cagatggaga ttatgacgtc cgggagtatt tcctcaatct ctatgatcag	540
atcctcatcg aaatgcccga ctatctgctc ctcaaggaca tggccgtgga gaacaaaaat	600
agcagagacg ccggcaaagt tgtcgactct gagactgccat atatttgta tgccatcttc	660
caggatgagg agaccgaggg agtcgtccgt agattcatcg ctgatatgca gcaacaggtc	720
caggctgate gtaacattgt caattaccct tccatccttc accctatttga tcatgcattc	780
aatgagtatt ttcttaacca ccagttggtg gagccgtga acaatgagat aatctcaat	840
tacataccag agaggataag gaatgacgtg aattacatcc tgaacatgga tatgaatctg	900
ccatctacag ccaggtatat caggccaaac ttgttgagg atagactgaa tcttcacat	960
aatttigagt ccctgtggta taccatcaca acatccaact acattctggc caggtccgtc	1020
gttcccgatt tgaaggagaa ggagctggc tccacccgaag cacagatcca gaaaatgagc	1080
caggacctgc agctggaggg cctcaactatt cagagcgaga cacagtttt agccggatt	1140
aacagttagg ctgccaatga ttgtttcaag accctcatag ccgccccatgtctcaaaaga	1200
accatgtctt tggactttgt gaccacgaac tatatgagcc taatctccgg aatgtggcta	1260
cttacagtga ttcccaacgta tatgttcctc cgggagtcac tagtggcctg tgagctggcg	1320
atcatcaaca ccatcggtta tccagcattc ggaatgcaga gaatgcatta ccggaaatggc	1380
gaccctcaga cacccttcca gatcgcgaa cagcagatcc agaatttcca ggtggcgaac	1440
tggctccatt ttatataacaa taacagattc aggcaagttg tgattgtatgg agttctgaat	1500
cagactctga acgacaatatac acgaaatggc caggtcatca accagctgtat ggaagcattg	1560
atgcaactca gcagacagca gttcccccacg atgcctgtgg attacaaacg gaggatccaa	1620
cggggcattc tgcttctctc caataggctg gggcagctg tcgacttaac ccgactggc	1680
tcctataact acgagacgct aatggcttgcgt gtgaccatga acatgcagca cgtgcaaaacc	1740

ctgacaactg agaagttgca gtcacttct gtgacttcgc tttgtatgtt aattggtaac	1800
acaaccgtga ttccgtcccc acagacactg ttccactact acaacatcaa cgtgaattc	1860
cactccaatt ataatgagcg gatcaacgac gccgtcgcca taattaccgc agcaaatagg	1920
ctgaatcttt atcagaaaaa aatgaagtcc atagtgaaag actttctgaa acggctccag	1980
atttcgacg taccacgagt gcctgacgac caaatgtaca ggctgaggga tcgccttcgg	2040
ctcttacccg ttgaacggag acggcttgac atattcaact tgatcctgat gaatatggag	2100
cagatcgaac gcgcttctga taagattgct caggggtta tcatcgata ccgagatatg	2160
cagctggaac gcgacgagat gtacggatat gttaatattg cacggaatct tcatggctac	2220
cagcaaatta acttggagga actcatgcgc accggtgatt acggacaaat tacgaacatg	2280
cttctcaaca atcaacccgt tgcccttgcg ggtgcattgc cttcggtac ggactcatcc	2340
gtgatcagtc taatcgccaa gtcgacgca accgtttcg cttagatagt gaagctcagg	2400
aaagttgaca cactgaagcc catactgtac aaaataaact cggattccaa tgactttac	2460
cttgtggcca actacgactg gatccccaca agtacaacta aggtctacaa acaggtgcca	2520
caaccattcg actttagagc cagcatgcac atgctgactt ctaaccttac gtttaccgtc	2580
tactctgacc tactgtcatt tgttcagcg gacacggtag agcccattaa cgcatcgca	2640
ttcgacaata tgcgaaataat gaacgagett taa	2673
<210> 46	
<211> 1194	
<212> DNA	
<213> Rotavirus	
<400> 46	
atggagggtcc ttatagttct ctccaaaacg ctgaaggacg cttagggacaa gatcgtggag	60
ggtacacttt atagcaatgt cagcgaccta atacagcagt ttaatcaaat gatcggtaca	120
atgaatggga atgatttcca aactggcggt attggtaatc tgccgtgag gaactggaca	180
ttcgatttcg gcctgctggg cacgactctc cttaatctcg atgcaaaatata tgtagaaaac	240
gccagaacga ttatcgagta ctatcgat ttcatgtata acgtttgtat ggatgagatg	300
gcccgcgagt cacaacggaa cggagttgtcc cacagtcgg aggccttcg gaaactcgcc	360
ggcattaaatgt tcaagcgat taatttcgac aactctccg aatataataga gaactggaac	420
ttgcagaatc gtcgacagag aaccggcttc gtgtccata aacctaataat ctggatcgat	480
agcgcctcat tcaccctgaa taggagtcg cccatgcacg acaacctcat gggtacaatg	540
tggctgaatg cggggagtga aatacaggcgc gcccgggttcg attactcctg tgccattaaat	600

gcacccgcaa acatccagca gttcgaacat atcgtcaac taagacggc tctcacgacc	660
gcgacaatta cactcctgcc cgacgcccag cgcttctcct ttccccgcgt aatcaactca	720
gctgatggcg ccaccacttg gttctcaac cctgttatat tgccctaa caacgttagag	780
gtggagttc tcttaaacgg acagatcatc aatacctacc aagccagggtt cggcacgatt	840
attgcaagaa atttcgacgc tattcaggctg ctcttccaac tcatgggccc ccccaatatg	900
actcccgtg tgaacgctt gttccgcag gctcagccctt tccagcacca cgccaccgtc	960
ggcttgactc ttcaaataga gagcgcggc tgcaatcag tgctggcaga cgccaaacgag	1020
acgctgctgg caaacgttac cgccgtgcgg caagagtatg ccatccagt agggcctgtg	1080
tttccacccg gcatgaactg gactgaacta attactaact atagccatc cagagaagac	1140
aacttgcagc gggcttcac tgtggcctct atccggagta tggatcaa gtag	1194
<210> 47	
<211> 2328	
<212> DNA	
<213> Artificial sequence	
<220><223> Optimized coding sequence of Rotavirus A VP4 from strain	
RVA/Vaccine/USA/Rotarix-A41CB052A/1988/G1P1A[8]	
<400> 47	
atggctagcc tggatctacag acaactttt accaattcat attctgtgg tcttcatgac	60
gaaatcgagc agattggc cgagaagacc cagaacgtga ccatcaaccc tggaccttt	120
gctcagaccc gctatcccc tgtgaattgg gatcacggag aaatcaacga cagtacgacc	180
gtcgaaacca ttctggacgg gccataccaa cccaccacct tcacccacc taatgattat	240
tggatttaa tcaactccaa cacaacggg gtggctacg agtccactaa taactccgat	300
ttttggaccc ccgttgtacg catcgagcc cacgtcaatc ctgtcgatcg ccagttatag	360
atattcggcg agtccaaaca gtttaacgtt tccaaatgaca gcaacaaatg gaagttctg	420
gagatgttc gcagctcctc tcagaacggaa ttctataata gacggaccct tacctccgat	480
acacgactcg tggatatttt taagtacggc ggcagggtgt ggacattca cggtaaaacc	540
cctcgagcaa ccactgactc cagtagcact gcaaaccgtg acaatatac tattaccatc	600
cacagcgaat tctacataat cccaaagatct caggaaagta agtgtaacga atatatcaac	660
aacggactcc ccccaattca gaatacacgg aacgtggc ctctccact cagttctcg	720
tctatccagt ataagagagc acaagtgaat gaggacatata ttgtgagcaa gactgcctt	780
tggaaagaaa tgcgtacaa cagagacatt atcatccgtt ttaagttgg gaactctatc	840

gtgaagatgg gcccctggg gtacaaatgg tcagaaatct catataaagc cgccaactat	900
cagtataact acttgagaga cggcgagcag gtaaccgcc acacaacatg ctctgtcaac	960
ggcgttaata acttagcta caacggaggc ttcccccac ccgacttcgg tatcagccgg	1020
tatgaagtca tcaaggaaaa ttcttatgt tacgttagatt actggatga tagcaaagcg	1080
ttccgcacaa tgggtatgt taggagcctg gctgctaatac tcaattctgt gaagtgtact	1140
ggtggatcat attatttctc aattcccggt ggggcttggc cagtcataa tggcggggca	1200
gtctccctcc attttgctgg cgtgacgtt agcactcagt ttaccgattt cgtgtctctg	1260
aactccctga gttccgggtt ttcccttact gtcgacgagc cccattcag cattctgcgt	1320
acaagaactg tcaacctcta cgggttacct gccgcatac caaacaacgg caatgaatac	1380
tatgaaattt cggccgctt ctcttgata agtctggta caactaatga cgactatcag	1440
acacccatca tgaacagcgt gactgtcaga caggacctgg aaagacaact tacagatctg	1500
cgggaagaat tcaattctct cagtcaggag attgcaatgg cccaaatgtat agatcttgcc	1560
ctactgcctc tcatatgtt tagtatgttc tccggcatca aatcaactat agatctgaca	1620
aagagcatgg ctactctgt gatgaagaag ttcaagaaat caaaacttgc cacgagcata	1680
tcagaaatga cgaactctct gatgtatgca gcatcatcag cgtcacgcaa cgttccatt	1740
cggtcgaatc tcagcgccat cagcaactgg acaaacgtt ccaacgacgt cagcaacgt	1800
accaactcct tgaacgatat ttctacccag acgtcaacga tcagtaagaa actccgcitg	1860
aaagaaaatga tcacccagac tgagggatg tcttcgacg acatttccgc cgccgtgcta	1920
aaaacccaaaa tcgatatgtc tactcagatc ggcaagaaca ctctgcggat tatcgtaacc	1980
gaaggctccg aaaagtttat ccctaagcgc agctacagaa tattgaaaga tgacgaggtc	2040
atggagatca acacagaagg gaagttttc gcttataaga tcaacacctt tgacgaggtt	2100
ccgttgcacg tcaataagtt tgcagagctc gtgacagata gtccagtgtat ttctgccatc	2160
attgacttta agactttgaa gAACCTGAAC gacaactatg gaataaacacg gacggaaagcg	2220
ttgaacctca ttaagtccaa tcccaatatg ttgcgcatt tcattaacca gaacaatcca	2280
atcataagaa ataggattga gcaattaatc ctgcaatgta aactctga	2328
<210> 48	
<211> 981	
<212> DNA	
<213> Rotavirus	
<400> 48	
atgtacggca tcgagtatac aacaattttt atttcctga tttccatcat tctgttaaac	60

tacatcctta agtccgtgac cagaattatg gattatatta tctatcgtag ctcctcatc
 tacgtggccc ttttgccct gaccaggccc cagaactatg gcctgaactt accaatcacc
 gttcaatgg ataccgttta cgctaattcc actcaagagg ggatatttct gacaagtacc

ctgtgcctgt attatccaac agaaggctct acccagatca atgatgggaa gtggaaaggat
 agtctctcac agatgttctt aaccaaggcc tggcccaccc gttccgtcta cttcaaggaa
 tactctagta ttgtcgactt ctcagttgac cccagctt attgcgacta caacctggta
 ctatgaaat acgaccagaa cctggagctg gatatgtccg agctggctga cctgatcctc
 aatgagtggc tgtcaaccc catggacatc acattatatt actaccagca gtctggagaa
 tccaacaagt ggatcagttt gggctcaagt tgcaccgtga aggtgtgtcc cttgaacacc
 caaatgctgg gcattggttt tcagacaact aatgtggatt cgttgaaat ggtagccgaa

aacgagaagc tggctatagt ggacgttagtc gatgggatata accacaagat caatctgact
 accaccactt gtaccatcg aaactgtaaa aagctcgccc cccggagaa cgtcgccgt
 atccaggtgg gggggagcaa tgtgctcgac attactgccc accctaccac caatccacag
 acggaacgga tggatgagatg caactggaag aatggtggc aggtctttta taccattgt
 gactacatta accagattgt gcaagtcatg agtaaacggt ccagatccct gaactcagca
 gccttctatt atcgcttta g

<210> 49

<211> 903

<212> DNA

<213> Rotavirus

<400> 49

atggcgaaaa acgttgcgtt ttcggctta ttgtttctc ttcttgcgtt gtttccttct
 cagatcttcg cccagaacta tggcctgaac ttaccaatca cgggttcaat ggataaccgtt
 tacgctaatt ccactcaaga gggatattt ctgacaagta ccctgtgcct gtattatcca
 acagaaggcctt ctacccagat caatgtatggg gagtggaaagg atagtctctc acagatgttc
 ctaaccaagg gctggccac cggttccgtc tacttcaagg aatactctag tattgtcgac
 ttctcagttt acccccagct ttattgcgtac tacaacctgg tacttatgaa atacgaccag
 aacctggagc tggatatgtc cgagctggct gacctgatcc tcaatgagtg gctgtgcaac

ccccatggaca tcacattata ttactaccag cagtctggag aatccaacaa gtggatcagt
 atgggctcaa gttgcaccgtt gaaggtgtgtt cccttgaaca cccaaatgtt gggcatttgtt
 tgtcagacaa ctaatgtgaa ttctgttgcgtt atggtagccg aaaacgagaa gctggctata

gtggacgtag tcgatggat taaccacaag atcaatctga ctaccaccac ttgtaccatc	660
agaaaactgta aaaagctcgcc cccccggag aacgtcgccg tgatccaggt gggggggagc	720
aatgtgctcg acattactgc cgaccctacc accaatccac agacggaacg gatgatgaga	780
gtcaactgga agaaatggtg gcaggtctt tataccatg tggactacat taaccagatt	840
gtgcaagtca tgagtaaacg gtccagatcc ctgaacttag cagccttcta ttatcgctt	900
tag	903
<210> 50	
<211> 2331	
<212> DNA	
<213> Rotavirus	
<400> 50	
atggcttcgc tcatttatag acaattgctc acgaattctt atacagttaga tttatccat	60
gagatacaag agattggatc aactaaatca caaaatgtca caattaatcc tggaccattt	120
gchgcaaacag gttatgctcc agttaactgg ggacctggag aaattaatga ttctacgaca	180
gttgaaccat tgcgtggatgg gccttatcaa ccaacgacat tcaatccacc agtgcattat	240
tggatgttac tggctccaac gacacctggc gtaattgtt aaggtaaaaa taatacagat	300
agatggtag ccacaatttt aatcgagcca aatgttcagt ctgaaaatag aacttacact	360
atatttggta ttcaagaaca attaacggta tccaaatactt cacaagacca gtggaaattt	420
attgtatgtcg taaaaacaac tgcaaatttga agtataaggac aatatggacc attactatcc	480
agtccgaaat tatatgcagt tatgaagcat aatgaaaaat tatatacata tgaaggacag	540
acacctaacg ctaggacacg acattattca acaacgaatt atgattctgt taatatgact	600
gtcttttgtt acttttatat aattcctaga tctgaagagt ctaaatgtac ggaatacatt	660
aataatggat taccaccaat acaaaatact agaaaatgtt taccattatc gttgactgct	720
agagatgtaa tacactatag agctcaagct aatgaagata ttgtgatatc caagacatca	780
ttgtggaaag aatgcaata taatagagat ataactatata gatttaaatt tgcaaataca	840
attataaaat caggagggtt gggatataag tggtcagaaa tatcatatca gccagcgaat	900
tatcaataca catatactcg tcatggtaa gaagttaccg cacatactac ttgttcagtg	960
aatggcgta atgacttcag tttaatggta ggatattac caactgattt tggttatct	1020
aaattttaag aatattaaaga gaattcatac gtctatatcg attactggta tgattcacaa	1080
gcatttcgtt acgtgggtta tgcgtcgatcg ttagcagcaa acttgaattc agttatgtgt	1140
actggaggca gctataattt tagtctacca gttggacaat ggcctgttt aactggggga	1200

gcagtttctt tacattcagc tggtgtaaca ctatctactc aattcacaga tttcgatca	1260
ttaaattcat taagatttag attagacta gctgtcgaag aaccacactt taaactgact	1320
agaactagat tagatagatt gtatggtctg cctgctgcag atccaaataa tggtaaagaa	1380
tattatgaaa ttgctggacg atttcactt atatcattag tgccatcaa tgatgactat	1440
cagactccta tagcaaactc agttaactgtc cgacaagatt tagaaaggca gttaggagaa	1500

ctaagagaag agttaacgc ttgtctcaa gaaattgcaa tgtcgagtt aatcgattta	1560
gcgcttctac cattagatat gttctcaatg tttctggca taaaagtac tatttgtct	1620
gcaaaatcaa tggctactaa tgttatgaaa aaattcaaaa agtcaggatt agcgaattca	1680
gttcaacac tgacagattc tttatcagac gcagcatcat caaatatcaag aggttcatct	1740
atacggtcga ttggatcttc agcatcagca tggacggatg tatcaacaca aataactgat	1800
atatcgcat cagtaagttc agtttcgaca caaacgtcaa ctatcagtag aagattgaga	1860
ctaaaggaaa tggcaacaca aactgagggt atgaattttg atgatatac agcgcttt	1920

<210> 51

<211> 2331

<212> DNA

<213> Artificial sequence

<220><223> Optimized coding sequence of Rotavirus A VP4 from strain

RVA/Simian-tc/ZAF/SA11-H96/1958/G3P5B[2]

<400> 51

atggcttcat tgatatatcg ccagttgctg actaatagct atactgtgga tttgtcagac	60
gaaatccagg aaataggatc cacaaagagt cagaacgtga ccataaaccc cgaccgttc	120
gcccatgtt ggtatcccc cgtaaactgg ggcccccggcg agattaacga cagcaccacc	180
gtggagccac tgctggatgg accctaccaa cccactactt ttaatcctcc agtggactac	240
tggatgttgt tggctccac gacacctggt gtaattgttag agggcaccaa caataccgat	300

cgctggctgg cgacaatact gatagaaccc aacgtgcagt ccgagaacag aacctatacc	360
atttcggca tccaggaaca gctaaccgtg agcaatacga gccaggacca gtggaagttt	420
atcgatgtag taaaaactac ggccaatgga tctatcggc aatacggcc gctgctgtcc	480
tcaccaagc tcacgccgt gatgaaacat aatgagaacat tgtacactta cgagggccaa	540
accccaatg ccagaactgc ccactacagt acaaccaact atgactcggt gaacatgaca	600
gcgttctgtg attttatat tattccaaga tcagaagaat ccaagtgtac tgagtacatc	660
aacaatggac ttccacccat ccagaacact cgaaatgtcg tcccactgtc tctaactgct	720
cgggatgtga tccactatcg cgcccaagct aatgaggata tagtcatttc aaagacgagc	780
ttatgaaagg aaatgcagta taacagagac atcacaatca gttcaagtt cgccaaact	840
attattaagt cggggggact ggggtacaaa tggagtgaga tcagtttaa gcccgtAAC	900
tatcagtaca cctatactcg cgacggcgaa gaggttaaccg cccacacaaac ttgctcggtt	960
aatggcgtga acgattttag cttcaacggg ggctacctgc ctactgattt cgtggtgagc	1020
aagttgaag tcatacaagga aaattcctac gtgtatattg actactggaa tgatagccag	1080
gcctccgaa atgttgta tggtagatca ctggccgaa acctaattc agtcatgtgc	1140
acaggagggtt cttacaattt tagtcttccc gtcggcagt ggccagtgtcacagggggc	1200
gctgtgagct tgcattccgc cggagtcacc ttgagtactc agttcacaga ctttgttct	1260
ctgaatagcc taaggttcag gtttagactt gcagtagaaag agcctcaact taagctact	1320
cgtacgaggg tggatcggtgtacggcctg cccggcgtg atcccaataa cggcaaggaa	1380
tattacgaga tagccggag atttcgctg atcagtcgg tgccgtAAA cgatgattac	1440
cagaccccaa ttgccaacag tgtcaactgtc aggcaagatc tggagagaca acttggcgag	1500
ctgagagagg agtcaacgc cctgtctcaa gagatcgaa tgtctcagct cattgacctg	1560
gccctgttac ccctcgacat gttctcaatg ttctccggca taaaatccac tatcgacgt	1620
gcaaaagtcca tggccacaaa tgtgtatgaag aagttaaga agagcggctt ggcaaatagc	1680
gtgtctacgc tgaccgatag ttgtcggtt gcccggcgtt ccattagccg tggatccagc	1740
attaggtcca ttggctcttc cgcctctgtc tggactgacg tgagtacaca gataactgac	1800
atttccttt ctgtctccag tgtgagcaca caaacttcca cgatatacg acgactgagg	1860
ctcaaagaga tggcaacgca aacggaaggt atgaatttg atgacatcg cgcccgagt	1920
ttgaagacaa agatcgataa aagcactcaa attagccccat acgatccc tgacatttg	1980
actgaggcat ctgaaaagtt cattccaaac cgtgcttac gggcattaa caatgtatgt	2040
gtcttcgagg cccgcattcgatggcaagttttt tttgttata aagtggatac cttcgaggag	2100
attccttcg atgtacagaa gtttgctgac ctcgtacgg atagcccgat gataagcgcc	2160

attatagact tcaaaacatt gaaaaattt aacgataatt atggattac caagcagcag	2220
gccttaact tgtaagatc tgaccctcg gtgcagag agtttattaa ccaggacaac	2280
cccatcatca gaaacaggat cgagcagctg attatgcagt gtcgcctgt a	2331
<210> 52	
<211> 981	
<212> DNA	
<213> Artificial sequence	

<220><223> Coding sequence of Rotavirus A VP7 from strain

RVA/Simian-1c/ZAF/SA11-H96/1958/G3P5B[2]

<400> 52	
atgtatggta ttgaatatac cacagttcta accttctga tatcgattat tctactaaat	60
tacatactta aatcattaac tagaataatg gactttataa tttagatt tcttttata	120
atttgatat tgtcaccatt tctcagagca caaaattatg gtattaatct tccaatcaca	180
ggctccatgg acactgcata cgctaattca acgcaagaag aaacattcct cacttctaca	240
cttgcctat attatccgac tgaggctgctg actgaaataa acgataattc atgaaagac	300
acactgtcac aactatttct tacgaaaggg tggccaactg gatccgtata tttaaagaa	360

tatactaaca ttgcatcggt ttctgttgc ccgcaggatgtt attgtgatta taacgttaga	420
ctaatgaaat atgacgcgac gttcaattt gatatgtcag aacttgcgg tctaattat	480
aacgaatggt tgtatcc aatggatatt actctgtatt attatcagca aactgacgaa	540
gcgaataataat ggatataat gggctcatca tgtacaat aagtatgtcc acttaataca	600
caaactttt gaattggatg cttgacaact gatgctacaa ctttgaaga agttgcgaca	660
gctgaaaatg tgtaattac tgacgtggtt gatggcgta atcataagct ggtatgtcaca	720
acagcaacgt gtactattag aaactgtaa gaaattggac caagagaaaa cgtagccgtt	780

ataacaagttt gtggttctga catcctcgat ataactgctg atccaactac tgcaccacag	840
acagaacgga tcatgcgat taactggaaa aaatggtggc aagtttttta tactgttaga	900
gactatgttag atcagataat acaagttatg tccaaagat caagatcact aaattcagca	960
gcattttattt acagagtgtat g	981

<210> 53

<211> 981

<212> DNA

<213> Rotavirus

<400> 53

atgtacggaa tcgagttatac caccgttctg acatttctta ttagtattat cctcttgaac	60
tatattctga agtcacttac ccggataatg gattttatata tatatagggtt tctgttcatc	120
attgttaattc tgagccctt cctgagggcc cagaattacg gcataaacct accaatcacc	180
ggttctatgg ataccgctta tgctaactct acacaagagg agacattcct cacatcaacc	240
ctatgcctgt actatccgac tgaaggcagcc acagagataa acgataactc ttggaaagat	300
acattgagcc agctttcct gactaaggga tggcccaccc gatcggtcta cttaaggag	360
tacacaaaca tcgcaagttt cagcgtggat ccccgatgtt attgtgatta taacgttg	420
ctgtatgaaat acgacgcaac cctccagctt gacatgagcg agttggcaga cctaattcctc	480
aatgagtggc tgitgtaaccc aatggatata acactgtact attatcagca gaccgatgaa	540
gcaaacaataat ggatttaat gggaaagcagc tgtaccatca aagtttgtcc tctcaacacc	600
caaactctcg gcatagggtg tctgaccaca gacgctacta ccttgaaga agttgcgacc	660
gcggaaaagc tggttatcac agatgtggta gatggcgta accacaaatt ggacgtaacc	720
acagcaacat gcacaattag gaactgcaag aagctaggac ccaggaaaaa cgtagccgtc	780
atccaagtgg gccccgtgtga catccttagac atcaccgcag acccaacaac agcaccacaa	840
accgagagga tggatgcgtat taatttggaaat aaatggtggc aggtgttttta cactgtcg	900
gactatgtgg accagatcat tcagggtgtg agcaaggcgaa gtgcgtcatt gaatagtgt	960
gccttttattt acagagtcta a	981
<210> 54	
<211> 981	
<212> DNA	
<213> Artificial sequence	
<220><223> Optimized coding sequence of Rotavirus A VP7	
<400> 54	
atgtacggca tcgagttatac aacaattttt attttctgtat tttccatcat tctgttaaac	60
tacatcctta agtccgtgac cagaattatg gattatatta tctatcgtag cctccatc	120
taatgtggcc ttttgcctt gaccaggccc cagaactatg gcctgaactt accaatcacc	180
ggttcaatgg ataccgttta cgcttaattcc actcaagagg ggatatttct gacaagtacc	240
ctgtgcctgt attatccaac agaaggcctt acccagatca atgatggggaa gtggaaaggat	300
agtctctcac agatgttcctt aaccaaggcc tggcccaccc gttccgtcta cttcaaggaa	360
tactcttagta ttgtcgactt ctcaaggctt attgcgacta caacctggta	420
cttatgaaat acgaccagaa cctggagctg gatatgtccg agctggctga cctgatcctc	480

aatgagtggc tgtgcaaccc catggacatc acattatatt actaccagca gtctggagaa 540
tccaacaagt ggatcgtat gggctcaagt tgcaccgtga aggtgtgtcc cttgaacacc 600
caaatgctgg gcattggttg tcagacaact aatgtggatt cgttgaaat ggtagcccaa 660

aacgagaagc tggctatagt ggacgttagtc gatgggatta accacaagat caatctgact 720
accaccactt gtaccatcg aaactgtaaa aagctggcc cccggagaa cgtcgccgtg 780
atccaggtgg gggggagcaa tgtgctcgac attactgccg accctaccac caatccacag 840
acggaacgga tcatgagagt caactggaag aaatgggtgc aggtcttttta taccatttg 900
gactacatta accagattgt gcaagtcatg agtaaacggt ccagatccct gaactcagca 960
gccttctatt atcgcgttta g 981

<210> 55

<211> 50

<212> DNA

<213> Artificial sequence

<220><223> Primer IF-TrSP+Rtx_VP7(opt).s1+3c

<400> 55

aaatttgtcg gccccatgga ttatattatc tatcgttagcc tcctcatcta 50

<210> 56

<211> 54

<212> DNA

<213> Artificial sequence

<220><223> Primer IF-Rtx_VP7(opt).s1-4r

<400> 56

actaaagaaa ataggcctct aaacgcgata atagaaggct gctgagttca ggaa 54

<210> 57

<211> 981

<212> DNA

<213> Artificial sequence

<220><223> Optimized coding sequence of Rotavirus A VP7 from strain
RVA/Vaccine/USA/Rotarix-A41CB052A/1988/G1P1A[8]

<400> 57

atgtacggca tcgagtatac aacaatttttta attttctgtat tttccatcat tctgttaaac 60
tacatcctta agtccgtgac cagaattatg gattatatttta tctatcgtag cctccatc 120

tacgtggccc ttttgcctt gaccaggccc cagaactatg gcctgaactt accaatcacc	180
gttcaatgg ataccgtta cgctaattcc actcaagagg ggatatttct gacaagtacc	240
ctgtgcctgt attatccaac agaaggctct acccagatca atgatggga gtgaaaggat	300
agtctctcac agatgttcct aaccaaggcc tggccaccc gttccgtcta cttaaggaa	360
tactcttagta ttgtcgactt ctcagttac cccagctt attgcgacta caacctggta	420
cttatgaaat acgaccagaa cctggagctg gatatgtccg agctggctga cctgatcctc	480
aatagtgcc tgtcaaccc catggacatc acattatatt actaccagca gtctggagaa	540
tccaacaagt ggatcagtat gggctcaagt tgcaccgtga aggtgtgtcc cttgaacacc	600
caaatgtgg gcattggttg tcagacaact aatgtggatt cgttgaaat ggtagccaa	660
aacgagaagc tggctatagt ggacgttagtc gatgggatta accacaagat caatctgact	720
accacccactt gtaccatcg aaactgtaaa aagctcgcc cccggagaa cgtcgccgtg	780
atccaggtgg gggggagcaa tgtgctcgac attactgccg accctaccac caatccacag	840
acggaacgga tcatgagagt caactggaag aaatggtggc aggtctttta taccatttg	900
gactacatta accagattgt gcaagtcatg agtaaacggt ccagatccct gaactcagca	960
gccttctatt atcgcgttta g	981
<210> 58	
<211> 2634	
<212> DNA	
<213> Artificial sequence	
<220><223> Expression cassette number 1734	
<400> 58	
gtcaacatgg tggagcacga cacacttgc tachtccaaa atatcaaaga tacagtctca	60
gaagacccaaa gggcaattga gactttcaa caaaggtaa tatccggaaa cttccctcgga	120
ttccattgcc cagctatctg tcaatttatt gtgaagatag tgaaaagga aggtggctcc	180
tacaatgcc atcattgcga taaaggaaag gccatcggt aagatgcctc tgccgacagt	240
ggtcccaag atggaccccc acccacgagg agcatcggtt aaaaagaaga cgttccaacc	300
acgtcttcaa agcaagtgg ttgatgtat aacatgggtt agcacgacac acttgtctac	360
tccaaaata tcaaagatac agtctcagaa gaccaaggaa caattgagac ttttcaacaa	420
aggtaatat ccggaaacct cctcggtt cattgeccag ctatctgtca ctttattgt	480
aagatagtgg aaaaggaagg tggctctac aaatgccatc attgcgataa aggaaaggcc	540
atcggttgaag atgcctctgc cgacagtggt cccaaagatg gaccccccacc cacgaggagc	600

atcggtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattt atgtgatatac	660
tccactgacg taaggatga cgcacaatcc cactatccctt cgcaagaccc ttccctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggtttt ataaaaggcga	780
acgtggggaa acccgacca aaccccttc taaactctct ctcatctctc ttaaagcaaa	840
cttcctctt gtcttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac	900
cagtacaacg tttttttca ctgaagcgaa atcaaagatc tctttgtgga cacgtatgc	960
ggccatcca aataacgtgt acttgtccta ttcttgtcgg tgtggcttg ggaaaagaaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa taictctact tctgcttgc gaggtattgt tgccctgtact tctttctct	1140
tcttcgttctt gattggttct ataagaaatc tagtattttc tttgaaacag agtttcccg	1200
tggtttctga acttggagaa agattgttaa gcttctgtat attctgccca aatttgcgg	1260
gcccattggat tatattatct atcgtagcct cctcatctac gtggcccttt ttgcctgac	1320
caggccccag aactatggcc tgaacttacc aatcaccgt tcaatggata ccgttacgc	1380
taattccact caagagggga tatttctgac aagtaccctg tgcctgtatt atccaacaga	1440
agcctctacc cagatcaatg atggggagtg gaaggatagt ctctcacaga tgttcctaacc	1500
caagggtgg cccaccgggtt ccgtctactt caaggaatac tcttagtattt tcgacttctc	1560
agttgacccc cagctttattt gcgactacaa cctggactt atgaaatacg accagaacct	1620
ggagctggat atgtccgagc tggctgaccc tgcctcaat gagtggctgt gcaaccccat	1680
ggacatcaca ttatattactt accagcgtc tggagaatcc aacaagtggta tcagtatgg	1740
ctcaagttgc accgtgaagg tggccctt gaacacccaa atgctggca ttgggtgtca	1800
gacaactaat gtggattcgt ttgaaatggt agccgaaaac gagaagctgg ctatagtgaa	1860
cgtagtcgtt gggattaacc acaagatcaa tctgactacc accacttgtt ccatcagaaaa	1920
ctgtaaaaag ctggcccccc gggagaacgt cgccgtgtt caggtgggggg ggagcaatgt	1980
gctcgacatt actgcccacc ctaccacca tccacacacg gaacggatga tgagagtcaa	2040
ctggaaagaaa tggtggcagg tcttttatac cattgtggac tacattaacc agattgtca	2100
agtcatgagt aaacggtcca gatccctgaa ctcagcagcc ttcttatttgc gcgtttagag	2160
gccttatttgc tttagtttga attactgtt attcgggtgtt cattttatgc tttgggtgac	2220
gttttctgt gctcagatgt tggatattt atgtaatttta atttcttgc gagctccgtt	2280
ttagcagggtc gtcccttcag caaggacaca aaaagatttt aattttattt aaaaaaaaaaa	2340
aaaaaaaaagac cggaaattcg atatcaagct tatcgaccctg cagatcggtt aaacattgg	2400
caataaagtt tcttaagattt gaatccgtt gccggcttgc cgtgattat catataattt	2460

ctgttgaatt acgttaagca tctaataatt aacatgtaat gcatgacgtt atttatgaga 2520
 tgggtttta tgatttagt cccgcaatta tacatthaat acgcgataga aaacaaaata 2580
 tagcgcgcaa actaggataa attatcgccg gcgggtgtcat ctatgttact agat 2634
 <210> 59
 <211> 297
 <212> PRT
 <213> Rotavirus

<400> 59

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

Asn Asp Gly Glu Trp Lys Asp Ser Leu Ser Gln Met Phe Leu Thr Lys
 65 70 75 80
 Gly Trp Pro Thr Gly Ser Val Tyr Phe Lys Glu Tyr Ser Ser Ile Val
 85 90 95
 Asp Phe Ser Val Asp Pro Gln Leu Tyr Cys Asp Tyr Asn Leu Val Leu
 100 105 110
 Met Lys Tyr Asp Gln Asn Leu Glu Leu Asp Met Ser Glu Leu Ala Asp
 115 120 125

Leu Ile Leu Asn Glu Trp Leu Cys Asn Pro Met Asp Ile Thr Leu Tyr
 130 135 140
 Tyr Tyr Gln Gln Ser Gly Glu Ser Asn Lys Trp Ile Ser Met Gly Ser
 145 150 155 160
 Ser Cys Thr Val Lys Val Cys Pro Leu Asn Thr Gln Met Leu Gly Ile
 165 170 175
 Gly Cys Gln Thr Thr Asn Val Asp Ser Phe Glu Met Val Ala Glu Asn

180 185 190

Glu Lys Leu Ala Ile Val Asp Val Val Asp Gly Ile Asn His Lys Ile
195 200 205
Asn Leu Thr Thr Thr Cys Thr Ile Arg Asn Cys Lys Lys Leu Gly
210 215 220
Pro Arg Glu Asn Val Ala Val Ile Gln Val Gly Gly Ser Asn Val Leu
225 230 235 240
Asp Ile Thr Ala Asp Pro Thr Thr Asn Pro Gln Thr Glu Arg Met Met
245 250 255

Arg Val Asn Trp Lys Lys Trp Trp Gln Val Phe Tyr Thr Ile Val Asp
260 265 270
Tyr Ile Asn Gln Ile Val Gln Val Met Ser Lys Arg Ser Arg Ser Leu
275 280 285
Asn Ser Ala Ala Phe Tyr Tyr Arg Val
290 295