(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0069042
(43) 공개일자 2023년05월18일

(51) 국제특허분류(Int. Cl.)
C12N 15/113 (2010.01) *A61K 48/00* (2006.01)
C07K 14/47 (2006.01) *C07K 14/52* (2006.01)
C07K 14/54 (2006.01) *C07K 14/725* (2006.01)
C07K 16/00 (2006.01) *C12N 15/67* (2006.01)
C12N 15/85 (2006.01) *C12N 15/88* (2017.01)
(52) CPC특허분류
C12N 15/113 (2013.01)
A61K 48/0025 (2013.01)
(21) 출원번호 10-2022-7036619
(22) 출원일자(국제) 2021년03월22일
심사청구일자 없음
(85) 번역문제출일자 2022년10월20일
(86) 국제출원번호 PCT/US2021/023540
(87) 국제공개번호 WO 2021/189059
국제공개일자 2021년09월23일
(30) 우선권주장
62/992,518 2020년03월20일 미국(US)

(71) 출원인
오나 테라퓨틱스, 인코포레이티드
미국, 매사추세츠 02139, 캠브리지, 세컨드 플로어, 620 메모리얼 드라이브
(72) 발명자
웨셀호르트, 로버트 알렉산더
미국, 매사추세츠 02139, 캠브리지, 메모리얼 드라이브 620, 세컨드 플로어 사내
굿맨, 브라이언
미국, 매사추세츠 02139, 캠브리지, 메모리얼 드라이브 620, 세컨드 플로어 사내
(74) 대리인
김영철, 김 순 영

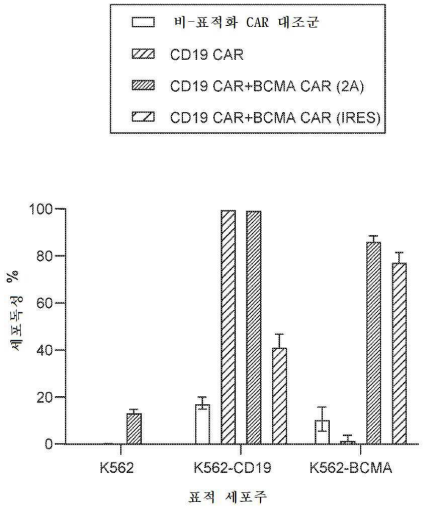
전체 청구항 수 : 총 126 항

(54) 발명의 명칭 **원형 RNA 조성물 및 방법**

(57) 요 약

관련 조성물 및 방법과 함께 원형 RNA가 본 명세서에 기재된다. 일부 실시형태에서, 본 발명의 원형 RNA는 포스트 스플라이싱 그룹 I 인트론 단편, 스페이서, IRES, 선택적인 듀플렉스 형성 영역 및 1종 초과의 발현 서열을 포함한다. 일부 실시형태에서, 발현 서열은 절단 부위를 암호화하는 1종 초과의 폴리뉴클레오타이드 서열에 의해서 분리되어 있다. 일부 실시형태에서, 본 발명의 원형 RNA는 선형 RNA와 비교할 때 개선된 발현, 기능적 안정성, 면역원성, 제조 용이성 및/또는 반감기를 갖는다. 일부 실시형태에서, 본 발명의 방법 및 작제물은 기존의 RNA 원형화 접근법과 비교할 때 개선된 원형화 효율, 스플라이싱 효율 및/또는 순도를 초래한다.

대 표 도 - 도44



(52) CPC특허분류
A61K 48/005 (2013.01)
C07K 14/4748 (2013.01)
C07K 14/521 (2013.01)
C07K 14/54 (2013.01)
C07K 14/7051 (2013.01)
C07K 16/00 (2013.01)
C12N 15/67 (2013.01)
C12N 15/85 (2013.01)
C12N 15/88 (2013.01)

명 세 서

청구범위

청구항 1

다음 순서로, 포스트 스플라이싱 3' 그룹 I 인트론 단편(post splicing 3’ group I intron fragment), 내부 리보솜 유입 부위(Internal Ribosome Entry Site: IRES), 제1 발현 서열, 제2 발현 서열 및 포스트 스플라이싱 5' 그룹 I 인트론 단편을 포함하는, 원형 RNA 폴리뉴클레오타이드.

청구항 2

제1항에 있어서, 상기 제1 발현 서열과 상기 제2 발현 서열 사이에 절단 부위를 암호화하는 폴리뉴클레오타이드 서열을 포함하는, 원형 RNA 폴리뉴클레오타이드.

청구항 3

제2항에 있어서, 상기 절단 부위는 자가-절단 스페이서인, 원형 RNA 폴리뉴클레오타이드.

청구항 4

제3항에 있어서, 상기 자가-절단 스페이서는 2A 자가-절단 펩타이드, 원형 RNA 폴리뉴클레오타이드.

청구항 5

제1항에 있어서, 상기 제1 발현 서열과 상기 제2 발현 서열 사이에 제2 IRES를 포함하는, 원형 RNA 폴리뉴클레오타이드.

청구항 6

제5항에 있어서, 상기 제1 IRES는 서열번호 1 내지 72 중 어느 하나에 따른 서열로 이루어지거나 상기 서열을 포함하는, 원형 RNA 폴리뉴클레오타이드.

청구항 7

제5항 또는 제6항에 있어서, 상기 제2 IRES는 서열번호 1 내지 72 중 어느 하나에 따른 서열로 이루어지거나 상기 서열을 포함하는, 원형 RNA 폴리뉴클레오타이드.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 제1 치료용 단백질을 암호화하고, 상기 제2 발현 서열은 제2 치료용 단백질을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열 또는 상기 제2 발현 서열은 항체를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 10

제1항 내지 제9항 중 어느 한 항에 있어서, 상기 제1 발현 서열 또는 상기 제2 발현 서열은 키메라 항원 수용체를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 상기 제1 발현 서열 또는 상기 제2 발현 서열은 전사 인자를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서, 상기 제1 발현 서열 또는 상기 제2 발현 서열은 사이토카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, 상기 제1 발현 서열 또는 상기 제2 발현 서열은 면역 저해 분자를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 상기 제1 발현 서열 또는 상기 제2 발현 서열은 공자극 분자의 효능제를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 15

제1항 내지 제14항 중 어느 한 항에 있어서, 상기 제1 발현 서열 또는 상기 제2 발현 서열은 면역 관문 분자의 저해제를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 16

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 T 세포 수용체(TCR)의 알파쇄를 암호화하고, 상기 제2 발현 서열은 T 세포 수용체(TCR)의 베타쇄를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 17

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 T 세포 수용체(TCR)의 베타쇄를 암호화하고, 상기 제2 발현 서열은 T 세포 수용체(TCR)의 알파쇄를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 18

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 T 세포 수용체(TCR)의 감마쇄를 암호화하고, 상기 제2 발현 서열은 T 세포 수용체(TCR)의 델타쇄를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 19

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 T 세포 수용체(TCR)의 델타쇄를 암호화하고, 상기 제2 발현 서열은 T 세포 수용체(TCR)의 감마쇄를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 20

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 T 세포 수용체(TCR)를 암호화하고, 상기 제2 발현 서열은 사이토카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 21

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 사이토카인을 암호화하고, 상기 제2 발현 서열은 T 세포 수용체(TCR)를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 22

제20항 또는 제21항에 있어서, 상기 사이토카인은 IL-2, IL-7, IL-12 및 IL-15로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 23

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 T 세포 수용체(TCR)를 암호화하고, 상기 제2 발현 서열은 케모카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 24

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 케모카인을 암호화하고, 상기 제2 발현 서열

은 T 세포 수용체(TCR)를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 25

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 T 세포 수용체(TCR)를 암호화하고, 상기 제2 발현 서열은 전사 인자를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 26

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 전사 인자를 암호화하고, 상기 제2 발현 서열은 T 세포 수용체(TCR)를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 27

제25항 또는 제26항에 있어서, 상기 전사 인자는 FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt 및 cd25로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 28

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 키메라 항원 수용체(CAR)를 암호화하고, 상기 제2 발현 서열은 PD1 또는 PDL1 길항제를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 29

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 PD1 또는 PDL1 길항제를 암호화하고, 상기 제2 발현 서열은 키메라 항원 수용체(CAR)를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 30

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 키메라 항원 수용체(CAR)를 암호화하고, 상기 제2 발현 서열은 사이토카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 31

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 사이토카인을 암호화하고, 상기 제2 발현 서열은 키메라 항원 수용체(CAR)를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 32

제30항 또는 제31항에 있어서, 상기 사이토카인은 IL-2, IL-7, IL-12 및 IL-15로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 33

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 키메라 항원 수용체(CAR)를 암호화하고, 상기 제2 발현 서열은 케모카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 34

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 케모카인을 암호화하고, 상기 제2 발현 서열은 키메라 항원 수용체(CAR)를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 35

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 전사 인자를 암호화하고, 상기 제2 발현 서열은 사이토카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 36

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 사이토카인을 암호화하고, 상기 제2 발현 서열은 전사 인자를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 37

제35항 또는 제36항에 있어서, 상기 전사 인자는 FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt 및 cd25로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 38

제35항 내지 제37항 중 어느 한 항에 있어서, 상기 사이토카인은 IL-10, IL-12 및 TGFβ로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 39

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 전사 인자를 암호화하고, 상기 제2 발현 서열은 케모카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 40

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 케모카인을 암호화하고, 상기 제2 발현 서열은 전사 인자를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 41

제39항 또는 제40항에 있어서, 상기 전사 인자는 FOXP3, STAT5B 및 HELIOS로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 42

제39항 내지 제41항 중 어느 한 항에 있어서, 상기 케모카인은 CC 케모카인, CXC 케모카인, C 케모카인 또는 CX3C 케모카인인, 원형 RNA 폴리뉴클레오타이드.

청구항 43

제39항 내지 제42항 중 어느 한 항에 있어서, 상기 케모카인은 CCL1, CCL2, CCL3, CCL4, CCL5, CCL6, CCL7, CCL8, CCL9/CCL10, CCL11, CCL12, CCL13, CCL14, CCL15, CCL16, CCL17, CCL18, CCL19, CCL20, CCL21, CCL22, CCL23, CCL24, CCL25, CCL26, CCL27, CCL28, CXCL 1, CXCL2, CXCL3, CXCL4, CXCL5, CXLC6, CXCL7, CXCL8, CXCL9, CXCL10, CXCL11, CXCL12, CXCL13, CXCL14, CXCL15, CXCL16, CXCL17, XCL1, XCL2 및 CX3CL1로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 44

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 종양 항원을 암호화하고, 상기 제2 발현 서열은 사이토카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 45

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 사이토카인을 암호화하고, 상기 제2 발현 서열은 종양 항원을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 46

제44항 또는 제45항에 있어서, 상기 항원은 신생항원인, 원형 RNA 폴리뉴클레오타이드.

청구항 47

제44항 내지 제46항 중 어느 한 항에 있어서, 상기 사이토카인은 IFNɣ인, 원형 RNA 폴리뉴클레오타이드.

청구항 48

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 CAR을 암호화하고, 상기 제2 발현 서열은 CAR을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 49

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 사이토카인을 암호화하고, 상기 제2 발현 서열은 사이토카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 50

제49항에 있어서, 상기 제1 또는 제2 발현 서열은 IL-10, TGFβ 및 IL-35로부터 선택된 사이토카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 51

제49항 또는 제50항에 있어서, 상기 제1 또는 제2 발현 서열은 IFNɣ, IL-2, IL-7, IL-15 및 IL-18로부터 선택된 사이토카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 52

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 T 세포 수용체(TCR)를 암호화하고, 상기 제2 발현 서열은 T 세포 수용체(TCR)를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 53

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 케모카인을 암호화하고, 상기 제2 발현 서열은 케모카인을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 54

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 또는 제2 발현 서열은 면역억제 효소를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 55

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 속도 제한 효소(rate limiting enzyme)를 암호화하고, 상기 제2 발현 서열은 플럭스-제한 효소(flux-limiting enzyme)를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 56

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 플럭스-제한 효소를 암호화하고, 상기 제2 발현 서열은 속도 제한 효소를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 57

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 전사 인자를 암호화하고, 상기 제2 발현 서열은 생존 인자를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 58

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 생존 인자를 암호화하고, 상기 제2 발현 서열은 전사 인자를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 59

제57항 또는 제58항에 있어서, 상기 전사 인자는 FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt 및 cd25로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 60

제57항 내지 제59항 중 어느 한 항에 있어서, 상기 생존 인자는 BCL-XL로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 61

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 또는 제2 발현 서열은 샤페론 단백질 또는 복합체를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 62

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 전사 인자를 암호화하고, 상기 제2 발현 서열은 샤페론 단백질 또는 복합체를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 63

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 샤페론 단백질 또는 복합체를 암호화하고, 상기 제2 발현 서열은 전사 인자를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 64

제61항 내지 제63항 중 어느 한 항에 있어서, 상기 샤페론 단백질 또는 복합체는 Skp, Spy, FkpA, SurA, Hsp60, Hsp70, GroEL, GroES, Hsp90, HtpG, Hsp100, ClpA, ClpX, ClpP 및 Hsp104로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 65

제61항 내지 제64항 중 어느 한 항에 있어서, 상기 전사 인자는 FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt 및 cd25로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 66

제1항 내지 제8항 중 어느 한 항에 있어서, 하나 또는 두 발현 서열은 신호전달 단백질을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 67

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 효소를 암호화하고, 상기 제2 발현 서열은 상기 제1 발현 서열의 음성 조절 저해제를 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 68

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 상기 제2 발현 서열에 암호화된 효소의 음성 조절 저해제 단백질을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 69

제67항 또는 제68항에 있어서, 상기 음성 조절 저해제는 p57kip2, BAX 저해제 및 TIPE2로부터 선택되는, 원형 RNA 폴리뉴클레오타이드.

청구항 70

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 우성 음성 단백질을 암호화하고, 상기 제2 발현 서열은 면역 단백질을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 71

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 면역 단백질을 암호화하고, 상기 제2 발현 서열은 우성 음성 단백질을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 72

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 또는 제2 발현 서열은 항-염증 단백질을 암호화하는, 원형 RNA 폴리뉴클레오타이드.

청구항 73

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 전사 인자를 암호화하고, 상기 제2 발현 서열은 5-플루오로사이토신(5-FC)을 5-플루오로우라실(5-FU)로 전환시킬 수 있는, 원형 RNA 폴리뉴클레오타이드.

청구항 74

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 제1 발현 서열은 5-플루오로사이토신(5-FC)을 5-플루오로우라실(5-FU)로 전환시킬 수 있고, 상기 제2 발현 서열은 전사 인자인, 원형 RNA 폴리뉴클레오타이드.

청구항 75

제73항 또는 제74항에 있어서, 5-플루오로사이토신(5-FC)을 5-플루오로우라실(5-FU)로 전환시킬 수 있는 상기 발현 서열은 사이토신 데아미나제인, 원형 RNA 폴리뉴클레오타이드.

청구항 76

제1항 내지 제75항 중 어느 한 항에 있어서, 상기 5' 듀플렉스 형성 영역과 상기 포스트 스플라이싱 3' 그룹 I 인트론 단편 사이의 제1 스페이서, 및 상기 포스트 스플라이싱 5' 그룹 I 인트론 단편과 상기 3' 듀플렉스 형성 영역 사이의 제2 스페이서를 포함하는, 원형 RNA 폴리뉴클레오타이드.

청구항 77

제76항에 있어서, 상기 제1 내부 스페이서 및 상기 제2 스페이서는 각각 약 10 내지 약 60개의 뉴클레오타이드의 길이를 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 78

제1항 내지 제77항 중 어느 한 항에 있어서, 상기 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 9 내지 약 19개 뉴클레오타이드 길이를 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 79

제77항 또는 제78항에 있어서, 상기 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 30개 뉴클레오타이드 길이를 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 80

제1항 내지 제79항 중 어느 한 항에 있어서, 상기 IRES는 타우라 증후군 바이러스(Taura syndrome virus), 트리아토마 바이러스(Triatoma virus), 타일러 뇌척수염 바이러스(Theiler's encephalomyelitis virus), 유인원 바이러스 40, 붉은 불개미 바이러스 1(Solenopsis invicta virus 1), 로팔로시품 파디 바이러스(Rhopalosiphum padi virus), 세망내피증 바이러스(Reticuloendotheliosis virus), 인간 폴리오바이러스 1, 갈색날개 노린재 장 바이러스(Plautia stali intestine virus), 카슈미르 벌 바이러스(Kashmir bee virus), 인간 리노바이러스 2(Human rhinovirus 2), 호말로디스카 코아굴라타 바이러스-1(Homalodisca coagulata virus-1), 인간 면역결핍 바이러스 타입 1, 호말로디스카 코아굴라타 바이러스-1(Homalodisca coagulata virus-1), 히메토비 P 바이러스(Himetobi P virus), C형 간염 바이러스, A형 간염 바이러스, 간염 GB 바이러스, 구제역 바이러스(Foot and mouth disease virus), 인간 엔테로바이러스 71(Human enterovirus 71), 말 비염 바이러스(Equine rhinitis virus), 엑트로피스 오블리쿠아 피코르나-유사 바이러스(Ectropis obliqua picorna-like virus), 뇌심근염 바이러스(Encephalomyocarditis virus), 초파리 C 바이러스(Drosophila C Virus), 인간 콕사키바이러스 B3(Human coxsackievirus B3), 크루시퍼 토바모바이러스(Crucifer tobamovirus), 귀뚜라미 마비병 바이러스(Cricket paralysis virus), 소 바이러스성 설사 바이러스 1(Bovine viral diarrhea virus 1), 블랙 퀸 세포 바이러스(Black Queen Cell Virus), 진딧물 치명 마비 바이러스(Aphid lethal paralysis virus), 조류 뇌척수염 바이러스(Avian encephalomyelitis virus), 급성 벌 마비 바이러스(Acute bee paralysis virus), 히비스커스 위황병 원형반점 바이러스(Hibiscus chlorotic ringspot virus), 돼지 열병 바이러스(Classical swine fever virus), 인간 FGF2, 인간 SFTPA1, 인간 AML1/RUNX1, 초파리 안테나페디아(Drosophila antennapedia), 인간 AQP4, 인간 AT1R, 인간 BAG-1, 인간 BCL2, 인간 BiP, 인간 c-IAP1, 인간 c-myc, 인간 eIF4G, 마우스 NDST4L, 인간 LEF1, 마우스 HIF1 알파, 인간 n.myc, 마우스 Gtx, 인간 p27kip1, 인간 PDGF2/c-sis, 인간 p53, 인간 Pim-1, 마우스

Rbm3, 초파리 리퍼(Drosophila reaper), 캐닌 스캠퍼(Canine Scamper), 초파리 Ubx, 인간 UNR, 마우스 UtrA, 인간 VEGF-A, 인간 XIAP, 털없는 초파리(Drosophila hairless), 에스. 세레비지애(S. cerevisiae) TFIID, 에스. 세레비지애 YAP1, 담배 식각 바이러스(tobacco etch virus), 터닙 주름 바이러스(turnip crinkle virus), EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, 피코비르나바이러스, HCV QC64, 인간 코사바이러스(Human Cosavirus) E/D, 인간 코사바이러스 F, 인간 코사바이러스 JMY, 리노바이러스(Rhinovirus) NAT001, HRV14, HRV89, HRVC-02, HRV-A21, 살리바이러스(Salivirus) A SH1, 살리바이러스 FHB, 살리바이러스 NG-J1, 인간 파레코바이러스(Human Parechovirus) 1, 크로히바이러스(Crohivirus) B, Yc-3, 로사바이러스(Rosavirus) M-7, 산바바이러스(Shanbavirus) A, 파시바이러스(Pasivirus) A, 파시바이러스 A 2, 에코바이러스(Echovirus) E14, 인간 파레코바이러스(Human Parechovirus) 5, 아이치 바이러스(Aichi Virus), A형 간염 바이러스 HA16, 포피바이러스(Phopivirus), CVA10, 엔테로바이러스(Enterovirus) C, 엔테로바이러스 D, 엔테로바이러스 J, 인간 페기바이러스(Human Pegivirus) 2, GBV-C GT110, GBV-C K1737, GBV-C 아이오와(Iowa), 페기바이러스(Pegivirus) A 1220, 파시바이러스(Pasivirus) A 3, 사펠로바이러스(Sapelovirus), 로사바이러스(Rosavirus) B, 바쿤사(Bakunsa) 바이러스, 트레모바이러스(Tremovirus) A, 돼지 파시바이러스(Swine Pasivirus) 1, PLV-CHN, 파시바이러스 A, 시시니바이러스(Sicinivirus), 헤파시바이러스(Hepacivirus) K, 헤파시바이러스 A, BVDV1, 보더병 바이러스(Border Disease Virus), BVDV2, CSFV-PK15C, SF573 디시스트로바이러스(Dicistrovirus), 후베이 피코르나-유사 바이러스(Hubei Picorna-like Virus), CRPV, 살리바이러스 A BN5, 살리바이러스 A BN2, 살리바이러스 A 02394, 살리바이러스 A GUT, 살리바이러스 A CH, 살리바이러스 A SZ1, 살리바이러스 FHB, CVB3, CVB1, 에코바이러스 7, CVB5, EVA71, CVA3, CVA12, EV24 또는 eIF4G에 대한 압타머로부터의 IRES의 서열을 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 81

제1항 내지 제80항 중 어느 한 항에 있어서, 자연 뉴클레오타이드로 이루어진, 원형 RNA 폴리뉴클레오타이드.

청구항 82

제1항 내지 제81항 중 어느 한 항에 있어서, 상기 발현 서열은 코돈 최적화된, 원형 RNA 폴리뉴클레오타이드.

청구항 83

제1항 내지 제82항 중 어느 한 항에 있어서, 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 마이크로RNA 결합 부위가 결여되도록 최적화된, 원형 RNA 폴리뉴클레오타이드.

청구항 84

제1항 내지 제83항 중 어느 한 항에 있어서, 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 엔도뉴클레아제 민감성 부위가 결여되도록 최적화된, 원형 RNA 폴리뉴클레오타이드.

청구항 85

제1항 내지 제84항 중 어느 한 항에 있어서, 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 RNA-편집 민감성 부위가 결여되도록 최적화된, 원형 RNA 폴리뉴클레오타이드.

청구항 86

제1항 내지 제85항 중 어느 한 항에 있어서, 상기 원형 RNA 폴리뉴클레오타이드는 약 100개 뉴클레오타이드 내지 약 15킬로베이스 길이인, 원형 RNA 폴리뉴클레오타이드.

청구항 87

제1항 내지 제86항 중 어느 한 항에 있어서, 적어도 약 20시간의 인간에서의 생체내 치료 효과 기간을 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 88

제1항 내지 제87항 중 어느 한 항에 있어서, 적어도 약 20시간의 기능성 반감기를 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 89

제1항 내지 제88항 중 어느 한 항에 있어서, 동일한 발현 서열을 포함하는 등가의 선형 RNA 폴리뉴클레오타이드의 인간 세포에서의 치료 효과 기간보다 더 크거나 동일한 인간 세포에서의 치료 효과 기간을 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 90

제1항 내지 제89항 중 어느 한 항에 있어서, 동일한 발현 서열을 포함하는 등가의 선형 RNA 폴리뉴클레오타이드의 인간 세포에서의 기능성 반감기보다 더 크거나 동일한 인간 세포에서의 기능성 반감기를 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 91

제1항 내지 제90항 중 어느 한 항에 있어서, 동일한 발현 서열을 갖는 등가의 선형 RNA 폴리뉴클레오타이드의 인간에서의 생체내 치료 효과 기간보다 더 큰 인간에서의 생체내 치료 효과 기간을 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 92

제1항 내지 제91항 중 어느 한 항에 있어서, 동일한 발현 서열을 갖는 등가의 선형 RNA 폴리뉴클레오타이드의 인간에서의 생체내 기능성 반감기보다 큰 인간에서의 생체내 기능성 반감기를 갖는, 원형 RNA 폴리뉴클레오타이드.

청구항 93

약제학적 조성물로서, 제1항 내지 제92항 중 어느 한 항의 원형 RNA 폴리뉴클레오타이드, 나노입자 및 선택적으로, 상기 나노입자에 작동 가능하게 연결된 표적화 모이어티를 포함하는, 약제학적 조성물.

청구항 94

제93항에 있어서, 상기 나노입자는 지질 나노입자, 코어-쉘 나노입자, 생분해성 나노입자, 생분해성 지질 나노입자, 중합체 나노입자 또는 생분해성 중합체 나노입자인, 약제학적 조성물.

청구항 95

제93항 또는 제94항에 있어서, 표적화 모이어티를 포함하되, 상기 표적화 모이어티는 세포 단리 또는 정제의 부재 하에서 수용체-매개된 엔도사이토시스 또는 선택된 세포 집단 또는 조직의 선택된 세포 내로의 직접 융합을 매개하는, 약제학적 조성물.

청구항 96

제93항 내지 제95항 중 어느 한 항에 있어서, 상기 나노입자에 작동 가능하게 연결된 표적화 모이어티를 포함하는, 약제학적 조성물.

청구항 97

제93항 내지 제96항 중 어느 한 항에 있어서, 상기 표적화 모이어티는 scFv, 나노바디, 펩타이드, 미니바디, 폴리뉴클레오타이드 압타머, 중쇄 가변 영역, 경쇄 가변 영역 또는 이의 단편인, 약제학적 조성물.

청구항 98

제93 내지 제97항 중 어느 한 항에 있어서, 상기 조성물 중의 상기 폴리뉴클레오타이드의 1중량% 미만은 이중 가닥 RNA, DNA 스플린트 또는 삼인산화된 RNA인, 약제학적 조성물.

청구항 99

제93항 내지 제98항 중 어느 한 항에 있어서, 상기 약제학적 조성물 중의 상기 폴리뉴클레오타이드 및 단백질의 1중량% 미만은 이중 가닥 RNA, DNA 스플린트, 삼인산화된 RNA, 포스파타제 단백질, 단백질 리가제 및 캡핑 효소

인, 약제학적 조성물.

청구항 100

치료를 필요로 하는 대상체의 치료 방법으로서, 제1항 내지 제92항 중 어느 한 항의 원형 RNA 폴리뉴클레오타이드, 나노입자 및 선택적으로, 상기 나노입자에 작동 가능하게 연결된 표적화 모이어티를 포함하는, 치료적 유효량의 조성물을 투여하는 단계를 포함하는, 치료 방법.

청구항 101

제100항에 있어서, 상기 표적화 모이어티는 scFv, 나노바디, 펩타이드, 미니바디, 중쇄 가변 영역, 경쇄 가변 영역 또는 이의 단편인, 치료 방법.

청구항 102

제100항 또는 제101항에 있어서, 상기 나노입자는 지질 나노입자, 코어-쉘 나노입자 또는 생분해성 나노입자인, 치료 방법.

청구항 103

제100항 내지 제102항 중 어느 한 항에 있어서, 상기 나노입자는 1종 이상의 양이온성 지질, 이온화 가능한 지질 또는 폴리 β-아미노 에스터를 포함하는, 치료 방법.

청구항 104

제100항 내지 제103항 중 어느 한 항에 있어서, 상기 나노입자는 1종 이상의 비-양이온성 지질을 포함하는, 치료 방법.

청구항 105

제100항 내지 제104항 중 어느 한 항에 있어서, 상기 나노입자는 1종 이상의 PEG-변형된 지질, 폴리글루탐산 지질 또는 히알루론산 지질을 포함하는, 치료 방법.

청구항 106

제100항 내지 제105항 중 어느 한 항에 있어서, 상기 나노입자는 콜레스테롤을 포함하는, 치료 방법.

청구항 107

제100항 내지 제106항 중 어느 한 항에 있어서, 상기 나노입자는 아라키돈산 또는 올레산을 포함하는, 치료 방법.

청구항 108

제100항 내지 제107항 중 어느 한 항에 있어서, 상기 조성물은 표적화 모이어티를 포함하되, 상기 표적화 모이어티는 선택적으로 세포 선택 또는 정제의 부재 하에서 선택된 세포 집단의 세포 내에서 수용체-매개된 엔도사이토시스를 매개하는, 치료 방법.

청구항 109

제100항 내지 제108항 중 어느 한 항에 있어서, 상기 나노입자는 1종 초과의 원형 RNA 폴리뉴클레오타이드를 포함하는, 치료 방법.

청구항 110

제100항 내지 제109항 중 어느 한 항에 있어서, 상기 대상체는 급성 림프구성 백혈병; 급성 골수성 백혈병(acute myeloid leukemia: AML); 포상횡문근육종; B 세포 악성종양; 방광암(예를 들어, 방광 암종); 골암; 뇌암(예를 들어, 수모세포종); 유방암; 항문, 항문관 또는 항문직장의 암; 안암; 간내담관의 암; 관절암; 두부암; 담낭암; 흉막의 암; 코, 비강 또는 중이의 암; 구강암; 음문암; 만성 림프구성 백혈병; 만성 골수성 암; 결장암; 식도암; 자궁경부암; 섬유육종; 위장 카르시노이드 종양; 두경부암(예를 들어, 두경부 편평 세포

암종); 호지킨 림프종; 하인두암; 신장암; 후두암; 백혈병; 액체 종양; 간암; 폐암(예를 들어, 비소세포 폐 암종 및 폐 선암종); 림프종; 중피종; 비만세포종; 흑색종; 다발성 골수종; 코인두암; 비호지킨 림프종; B-만성 림프구성 백혈병; 털세포 백혈병; 급성 림프구성 백혈병(acute lymphocytic leukemia: ALL); 버킷 림프종, 난소암; 췌장암; 복막암; 대망(omentum)암; 장간막암; 인두암; 전립선암; 직장암; 신장암; 피부암; 소장암; 연조직암; 고형 종양; 활막 육종; 위암; 고환암; 갑상선암; 및 요관암으로 이루어진 군으로부터 선택된 암을 갖는, 치료 방법.

청구항 111

제100항 내지 제109항 중 어느 한 항에 있어서, 상기 대상체는 경피증, 그레이브스병, 크론병, 쇼그렌병, 다발성 경화증, 하시모토병, 건선, 중증 근무력증, 자가면역 다발성 내분비병증 증후군(autoimmune polyendocrinopathy syndrome), 제I형 당뇨병(TIDM), 자가면역 위염, 자가면역 포도막염, 다발성근염, 대장염, 갑상선염 및 인간 루푸스로 대표되는 전신성 자가면역 질환(generalized autoimmune diseases typified by human Lupus)으로부터 선택되는 자가면역 장애를 갖는, 치료 방법.

청구항 112

원형 RNA 폴리뉴클레오타이드의 제조를 위한 벡터로서, 상기 원형 RNA 폴리뉴클레오타이드는 다음 순서로, 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 제2 발현 서열, 5' 그룹 I 인트론 단편 및 3' 듀플렉스 형성 영역을 포함하는, 벡터.

청구항 113

제112항에 있어서, 상기 제1 발현 서열과 상기 제2 발현 서열 사이에 절단 부위를 암호화하는 폴리뉴클레오타이드 서열을 포함하는, 벡터.

청구항 114

제112항 또는 제113항에 있어서, 상기 절단 부위는 자가-절단 스페이서인, 벡터.

청구항 115

제112항 내지 제114항 중 어느 한 항에 있어서, 상기 자가-절단 스페이서는 2A 자가-절단 펩타이드, 벡터.

청구항 116

제112항 내지 제115항 중 어느 한 항에 있어서, 상기 5' 듀플렉스 형성 영역과 상기 3' 그룹 I 인트론 단편 사이의 제1 스페이서 및 상기 5' 그룹 I 인트론 단편과 상기 3' 듀플렉스 형성 영역 사이의 제2 스페이서를 포함하는, 벡터.

청구항 117

제112항 내지 제116항 중 어느 한 항에 있어서, 상기 제1 스페이서 및 상기 제2 내부 스페이서는 각각 약 5 내지 약 60개의 뉴클레오타이드의 길이를 갖는, 벡터.

청구항 118

제112항 내지 제117항 중 어느 한 항에 있어서, 상기 제1 스페이서 및 제2 스페이서스페이서 각각은 적어도 5개 뉴클레오타이드 길이의 비구조화된 영역을 포함하는, 벡터.

청구항 119

제112항 내지 제118항 중 어느 한 항에 있어서, 상기 제1 스페이서 및 제2 스페이서스페이서 각각은 적어도 7개 뉴클레오타이드 길이의 구조화된 영역을 포함하는, 벡터.

청구항 120

제112항 내지 제119항 중 어느 한 항에 있어서, 상기 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역 각각은 약 9 내지 50개 뉴클레오타이드 길이를 갖는, 벡터.

청구항 121

제112항 내지 제120항 중 어느 한 항에 있어서, 상기 벡터는 코돈 최적화된, 벡터.

청구항 122

제112항 내지 제121항 중 어느 한 항에 있어서, 등가의 사전-최적화 폴리뉴클레오타이드에 존재하는 적어도 하나의 마이크로RNA 결합 부위가 결여된, 벡터.

청구항 123

제1항 내지 제92항 중 어느 한 항에 따른 원형 RNA 폴리뉴클레오타이드를 포함하는 진핵 세포.

청구항 124

제123항에 있어서, 상기 진핵 세포는 인간 세포인, 진핵 세포.

청구항 125

제124항에 있어서, 상기 진핵 세포는 면역 세포인, 진핵 세포.

청구항 126

제125항에 있어서, 상기 진핵 세포는 T 세포인, 진핵 세포.

발명의 설명

기 술 분 야

[0001] **관련 출원에 대한 상호 참조**

[0002] 본 출원은 미국 가출원 제62/992,518호(출원일: 2020년 3월 20일)의 이익 및 이에 대한 우선권을 주장하며, 상기 기초출원의 내용은 모든 목적을 위해서 전문이 참조에 의해 본 명세서에 원용된다.

배 경 기 술

[0003] 종래의 유전자 요법은 숙주 세포에 목적하는 유전자 정보를 삽입하기 위해서 DNA의 사용을 포함한다. 세포에 도입된 DNA는 일반적으로 하나 이상의 형질주입된 세포의 게놈에 어느 정도 통합되어, 숙주에서 도입된 유전 물질의 오래 지속되는 작용을 허용한다. 이러한 지속적인 작용에 상당한 이점이 있을 수 있지만, 외인성 DNA를 숙주 게놈에 통합하는 것도 많은 해로운 영향을 미칠 수 있다. 예를 들어, 도입된 DNA가 온전한 유전자에 삽입되어, 내인성 유전자의 기능을 방해하거나 심지어 완전히 제거하는 돌연변이를 초래할 수 있다. 따라서, DNA를 이용한 유전자 요법은 예를 들어, 필수 효소의 제거 또는 유해하게 감소된 생산 또는 세포 성장 조절에 중요한 유전자의 중단과 같이, 처리된 숙주에서 중요한 유전 기능의 손상을 초래할 수 있는데, 이는 조절되지 않는 또는 암성 세포 증식을 초래한다. 또한, 종래의 DNA 기반 유전자 요법에서는, 강력한 프로모터 서열을 포함시키기 위해서 목적하는 유전자 산물의 효과적인 발현이 필수적인데, 이는 다시 세포에서 정상적인 유전자 발현의 조절에 바람직하지 않은 변화로 이어질 수 있다. 또한 DNA 기반 유전 물질이 원치 않는 항-DNA 항체를 유도하여 아마도 치명적인 면역 반응을 촉발할 수도 있다. 바이러스 벡터를 사용하는 유전자 요법 접근법은 또한 역 면역 반응을 유발할 수 있다. 일부 상황에서, 바이러스 벡터는 숙주 게놈에 통합될 수도 있다. 또한, 임상 등급 바이러스 벡터의 생산도 비용이 많이 들고 시간 소모적이다. 바이러스 벡터를 사용하여 도입된 유전 물질의 표적화 전달 또한 제어하기 어려울 수 있다. 따라서, DNA 기반 유전자 요법은 바이러스 벡터를 사용하여 분비된 단백질의 전달에 대해 평가되었지만(미국 특허 제6,066,626호; 미국 공개 제2004/0110709호), 이러한 접근법은 이러한 다양한 이유로 제한될 수 있다.

[0004] DNA와 달리, RNA는 형질주입된 세포의 게놈에 안정적으로 통합될 위험이 없고, 따라서 도입된 유전 물질이 필수 유전자의 정상적인 기능을 방해하거나 또는 유해하거나 발암성 효과를 초래하는 돌연변이를 유발할 문제를 제거하고, 암호화된 단백질의 효과적인 번역을 위해 외인성 프로모터 서열이 필요하지 않아서, 다시 가능한 유해한 부작용을 피할 수 있기 때문에, 유전자 치료제로 RNA를 사용하는 것이 실질적으로 더 안전하다. 또한 mRNA가 기능을 수행하기 위해 핵으로 들어갈 필요는 없지만, DNA는 이러한 주요 장벽을 극복해야 한다.

[0005] 원형 RNA는 안정적인 형태의 RNA를 설계하고 생산하는 데 유용하다. RNA 분자의 원형화는 특히 불활성 입체배좌로 접히는 경향이 있는 분자의 경우에 RNA 구조 및 기능 연구에 이점을 제공한다(문헌[Wang and Ruffner, 1998]). 원형 RNA는 또한 생체내 응용, 특히 단백질 대체 요법 및 백신접종을 포함한 유전자 발현 및 치료제의 RNA-기반 제어 연구 분야에서 특히 흥미롭고 유용할 수 있다.

[0006] 본 발명 이전에, 원형화된 RNA를 시험관내에서 제조하기 위한 주요 기술인 스플린트-매개 방법, 순열 인트론-엑손 방법 및 RNA 리가제-매개 방법의 3가지 주요 기술이 존재하였다. 그러나, 기존의 방법론은 원형화될 수 있는 RNA의 크기에 의해 제한되어 치료적 적용에 제한적이다.

발명의 내용

[0007] 관련 조성물 및 방법과 함께 원형 RNA가 본 명세서에 기재된다. 일부 실시형태에서, 본 발명의 원형 RNA는 포스트 스플라이싱 그룹 I 인트론 단편(post splicing group I intron fragment), 스페이서, IRES, 선택적인 듀플렉스 형성 영역 및 1종 초과의 발현 서열을 포함한다. 일부 실시형태에서, 본 발명의 원형 RNA는 듀플렉스 형성 영역을 포함한다. 일부 실시형태에서, 발현 서열은 절단 부위를 암호화하는 1종 초과의 폴리뉴클레오타이드 서열에 의해서 분리되어 있다. 일부 실시형태에서, 절단 부위는 자가-절단 펩타이드이다. 일부 실시형태에서, 자가-절단 펩타이드는 2A 자가-절단 펩타이드이다. 일부 실시형태에서, 제1 발현 서열과 제2 발현 서열은 리보솜 스키핑 요소(ribosomal skipping element)에 의해서 분리되어 있다. 일부 실시형태에서, 각각의 발현 서열은 치료용 단백질을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 사이토카인 또는 이의 기능성 단편을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 전사 인자를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 면역 관문 저해제를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 키메라 항원 수용체를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 제1 T-세포 수용체(TCR) 쇄를 암호화하고, 제2 발현 서열은 제2 TCR 쇄를 암호화한다. 일부 실시형태에서, 본 발명의 원형 RNA는 선형 RNA와 비교할 때 개선된 발현, 기능적 안정성, 제조 용이성 및/또는 반감기를 갖는다. 일부 실시형태에서, 본 발명의 원형 RNA는 감소된 면역원성을 갖는다. 일부 실시형태에서, 본 발명의 방법 및 작제물은 기존의 RNA 원형화 접근법과 비교할 때 개선된 원형화 효율, 스플라이싱 효율 및/또는 순도를 초래한다.

[0008] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 하나 이상의 마이크로RNA 결합 부위를 포함한다. 마이크로RNA 결합 부위는 간에서 발현되는 마이크로RNA에 의해서 인식된다. 일부 실시형태에서, 마이크로RNA 결합 부위는 miR-122에 의해서 인식된다.

[0009] 본 출원의 일 양상은 다음 순서로, 포스트 스플라이싱 3' 그룹 I 인트론 단편, 내부 리보솜 유입 부위(Internal Ribosome Entry Site: IRES), 제1 발현 서열, 제2 발현 서열 및 포스트 스플라이싱 5' 그룹 I 인트론 단편을 포함하는, 원형 RNA 폴리뉴클레오타이드를 제공한다.

[0010] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 제1 발현 서열과 제2 발현 서열 사이에 절단 부위를 암호화하는 폴리뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 절단 부위는 자가-절단 스페이서이다. 일부 실시형태에서, 자가-절단 스페이서는 2A 자가-절단 펩타이드이다.

[0011] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 제1 발현 서열과 제2 발현 서열 사이에 제2 IRES를 포함한다. 일부 실시형태에서, 제1 IRES는 서열번호 1 내지 72 중 어느 하나에 따른 서열로 이루어지거나 상기 서열을 포함한다. 일부 실시형태에서, 제2 IRES는 서열번호 1 내지 72 중 어느 하나에 따른 서열로 이루어지거나 상기 서열을 포함한다.

[0012] 일부 실시형태에서, 제1 발현 서열은 제1 치료용 단백질을 암호화하고, 제2 발현 서열은 제2 치료용 단백질을 암호화한다. 일부 실시형태에서, 제1 발현 서열 또는 제2 발현 서열은 항체를 암호화한다. 일부 실시형태에서, 제1 발현 서열 또는 제2 발현 서열은 키메라 항원 수용체를 암호화한다. 일부 실시형태에서, 제1 발현 서열 또는 제2 발현 서열은 전사 인자를 암호화한다. 일부 실시형태에서, 제1 발현 서열 또는 제2 발현 서열은 사이토카인을 암호화한다. 일부 실시형태에서, 제1 발현 서열 또는 제2 발현 서열은 면역 저해 분자를 암호화한다. 일부 실시형태에서, 제1 발현 서열 또는 제2 발현 서열은 공자극 분자의 효능제를 암호화한다. 일부 실시형태에서, 제1 발현 서열 또는 제2 발현 서열은 면역 관문 분자의 저해제를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 T 세포 수용체(TCR)의 알파쇄를 암호화하고, 제2 발현 서열은 T 세포 수용체(TCR)의 베타쇄를 암호화한다.

[0013] 일부 실시형태에서, 제1 발현 서열은 T 세포 수용체(TCR)의 베타쇄를 암호화하고, 제2 발현 서열은 T 세포 수용체(TCR)의 알파쇄를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 T 세포 수용체(TCR)의 감마쇄를 암호화하

고, 제2 발현 서열은 T 세포 수용체(TCR)의 델타쇄를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 T 세포 수용체(TCR)의 델타쇄를 암호화하고, 제2 발현 서열은 T 세포 수용체(TCR)의 감마쇄를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 T 세포 수용체(TCR)를 암호화하고, 상기 제2 발현 서열은 케모카인을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 케모카인을 암호화하고, 제2 발현 서열은 T 세포 수용체(TCR)를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 키메라 항원 수용체(CAR)를 암호화하고, 제2 발현 서열은 PD1 또는 PDL1 길항제를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 PD1 또는 PDL1 길항제를 암호화하고, 제2 발현 서열은 키메라 항원 수용체(CAR)를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 키메라 항원 수용체(CAR)를 암호화하고, 제2 발현 서열은 케모카인을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 케모카인을 암호화하고, 제2 발현 서열은 키메라 항원 수용체(CAR)를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 전사 인자를 암호화하고, 제2 발현 서열은 사이토카인을 암호화한다.

[0014] 일부 실시형태에서, 제1 발현 서열은 T 세포 수용체(TCR)를 암호화하고, 상기 제2 발현 서열은 사이토카인을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 사이토카인을 암호화하고, 제2 발현 서열은 T 세포 수용체(TCR)를 암호화한다. 일부 실시형태에서, 사이토카인은 IL-2, IL-7, IL-12 및 IL-15로부터 선택된다.

[0015] 일부 실시형태에서, 제1 발현 서열은 T 세포 수용체(TCR)를 암호화하고, 제2 발현 서열은 전사 인자를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 전사 인자를 암호화하고, 제2 발현 서열은 T 세포 수용체(TCR)를 암호화한다. 일부 실시형태에서, 전사 인자는 FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt 및 cd25로부터 선택된다.

[0016] 일부 실시형태에서, 제1 발현 서열은 키메라 항원 수용체(CAR)를 암호화하고, 제2 발현 서열은 사이토카인을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 사이토카인을 암호화하고, 제2 발현 서열은 키메라 항원 수용체(CAR)를 암호화한다. 일부 실시형태에서, 사이토카인은 IL-2, IL-7, IL-12 및 IL-15로부터 선택된다.

[0017] 일부 실시형태에서, 제1 발현 서열은 사이토카인을 암호화하고, 제2 발현 서열은 전사 인자를 암호화한다. 일부 실시형태에서, 전사 인자는 FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt 및 cd25로부터 선택된다. 일부 실시형태에서, 사이토카인은 IL-10, IL-12 및 TGFβ로부터 선택된다.

[0018] 일부 실시형태에서, 제1 발현 서열은 전사 인자를 암호화하고, 제2 발현 서열은 케모카인을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 케모카인을 암호화하고, 제2 발현 서열은 전사 인자를 암호화한다. 일부 실시형태에서, 전사 인자는 FOXP3, STAT5B 및 HELIOS로부터 선택된다. 일부 실시형태에서, 케모카인은 CC 케모카인, CXC 케모카인, C 케모카인 또는 CX3C 케모카인이다. 일부 실시형태에서, 케모카인은 CCL1, CCL2, CCL3, CCL4, CCL5, CCL6, CCL7, CCL8, CCL9/CCL10, CCL11, CCL12, CCL13, CCL14, CCL15, CCL16, CCL17, CCL18, CCL19, CCL20, CCL21, CCL22, CCL23, CCL24, CCL25, CCL26, CCL27, CCL28, CXCL 1, CXCL2, CXCL3, CXCL4, CXCL5, CXLC6, CXCL7, CXCL8, CXCL9, CXCL10, CXCL11, CXCL12, CXCL13, CXCL14, CXCL15, CXCL16, CXCL17, XCL1, XCL2 및 CX3CL1로부터 선택된다.

[0019] 일부 실시형태에서, 제1 발현 서열은 종양 항원을 암호화하고, 제2 발현 서열은 사이토카인을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 사이토카인을 암호화하고, 제2 발현 서열은 종양 항원을 암호화한다. 일부 실시형태에서, 항원은 신생항원이다. 일부 실시형태에서, 사이토카인은 IFNɣ이다.

[0020] 일부 실시형태에서, 제1 발현 서열은 CAR을 암호화하고, 제2 발현 서열은 CAR을 암호화한다.

[0021] 일부 실시형태에서, 제1 발현 서열은 사이토카인을 암호화하고, 제2 발현 서열은 사이토카인을 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 IL-10, TGFβ 및 IL-35로부터 선택된 사이토카인을 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 IFNɣ, IL-2, IL-7, IL-15 및 IL-18로부터 선택된 사이토카인을 암호화한다.

[0022] 일부 실시형태에서, 제1 발현 서열은 T 세포 수용체(TCR)를 암호화하고, 제2 발현 서열은 T 세포 수용체(TCR)를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 케모카인을 암호화하고, 제2 발현 서열은 케모카인을 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 면역억제 효소를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 속도 제한 효소(rate limiting enzyme)를 암호화하고, 제2 발현 서열은 플럭스-제한 효소(flux-limiting enzyme)를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 플럭스-제한 효소를 암호화하고, 제2 발현 서열은 속도 제한 효소를 암호화한다.

[0023] 일부 실시형태에서, 제1 발현 서열은 전사 인자를 암호화하고, 제2 발현 서열은 생존 인자를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 생존 인자를 암호화하고, 제2 발현 서열은 전사 인자를 암호화한다. 일부 실시

형태에서, 전사 인자는 FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt 및 cd25로부터 선택된다. 일부 실시형태에서, 생존 인자는 BCL-XL로부터 선택된다.

[0024] 일부 실시형태에서, 제1 또는 제2 발현 서열은 샤페론 단백질 또는 복합체를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 전사 인자를 암호화하고, 제2 발현 서열은 샤페론 단백질 또는 복합체를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 샤페론 단백질 또는 복합체를 암호화하고, 제2 발현 서열은 전사 인자를 암호화한다. 일부 실시형태에서, 샤페론 단백질 또는 복합체는 Skp, Spy, FkpA, SurA, Hsp60, Hsp70, GroEL, GroES, Hsp90, HtpG, Hsp100, ClpA, ClpX, ClpP 및 Hsp104로부터 선택된다. 일부 실시형태에서, 전사 인자는 FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt 및 cd25로부터 선택된다.

[0025] 일부 실시형태에서, 하나 또는 두 발현 서열은 신호전달 단백질을 암호화한다.

[0026] 일부 실시형태에서, 제1 발현 서열은 효소를 암호화하고, 제2 발현 서열은 상기 제1 발현 서열의 음성 조절 저해제를 암호화한다. 일부 실시형태에서, 제1 발현 서열은 제2 발현 서열에 암호화된 효소의 음성 조절 저해제 단백질을 암호화한다. 일부 실시형태에서, 음성 조절 저해제는 p57kip2, BAX 저해제 및 TIPE2로부터 선택된다.

[0027] 일부 실시형태에서, 제1 발현 서열은 우성 음성 단백질을 암호화하고, 제2 발현 서열은 면역 단백질을 암호화한다. 일부 실시형태에서, 제1 발현 서열은 면역 단백질을 암호화하고, 제2 발현 서열은 우성 음성 단백질을 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 항-염증 단백질을 암호화한다.

[0028] 일부 실시형태에서, 제1 발현 서열은 전사 인자를 암호화하고, 제2 발현 서열은 5-플루오로사이토신(5-FC)을 5-플루오로우라실(5-FU)로 전환시킬 수 있다. 일부 실시형태에서, 제1 발현 서열은 5-플루오로사이토신(5-FC)을 5-플루오로우라실(5-FU)로 전환시킬 수 있고, 제2 발현 서열은 전사 인자이다. 일부 실시형태에서, 5-플루오로사이토신(5-FC)을 5-플루오로우라실(5-FU)로 전환시킬 수 있는 발현 서열은 사이토신 데아미나제이다.

[0029] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 5' 듀플렉스 형성 영역과 포스트 스플라이싱 3' 그룹 I 인트론 단편 사이에 제1 스페이서, 및 포스트 스플라이싱 5' 그룹 I 인트론 단편과 3' 듀플렉스 형성 영역 사이에 제2 스페이서를 포함한다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서는 각각 약 10 내지 약 60개의 뉴클레오타이드의 길이를 갖는다. 일부 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 9 내지 약 19개의 뉴클레오타이드의 길이를 갖는다. 일부 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 30개의 뉴클레오타이드의 길이를 갖는다. 일부 실시형태에서, IRES는 타우라 증후군 바이러스(Taura syndrome virus), 트리아토마 바이러스(Triatoma virus), 타일러 뇌척수염 바이러스(Theiler's encephalomyelitis virus), 유인원 바이러스 40, 붉은 불개미 바이러스 1(Solenopsis invicta virus 1), 로팔로시품 파디 바이러스(Rhopalosiphum padi virus), 세망내피증 바이러스(Reticuloendotheliosis virus), 인간 폴리오바이러스 1, 갈색날개 노린재 장 바이러스(Plautia stali intestine virus), 카슈미르 벌 바이러스(Kashmir bee virus), 인간 리노바이러스 2(Human rhinovirus 2), 호말로디스카 코아굴라타 바이러스-1(Homalodisca coagulata virus-1), 인간 면역결핍 바이러스 타입 1, 호말로디스카 코아굴라타 바이러스-1, 히메토비 P 바이러스, C형 간염 바이러스, A형 간염 바이러스, 간염 GB 바이러스, 구제역 바이러스(Foot and mouth disease virus), 인간 엔테로바이러스 71(Human enterovirus 71), 말 비염 바이러스(Equine rhinitis virus), 엑트로피스 오블리쿠아 피코르나-유사 바이러스(Ectropis obliqua picorna-like virus), 뇌심근염 바이러스(Encephalomyocarditis virus), 초파리 C 바이러스(Drosophila C Virus), 인간 콕사키바이러스 B3(Human coxsackievirus B3), 크루시퍼 토바모바이러스(Crucifer tobamovirus), 귀뚜라미 마비병 바이러스(Cricket paralysis virus), 소 바이러스성 설사 바이러스 1(Bovine viral diarrhea virus 1), 블랙 퀸 세포 바이러스(Black Queen Cell Virus), 진딧물 치명 마비 바이러스(Aphid lethal paralysis virus), 조류 뇌척수염 바이러스(Avian encephalomyelitis virus), 급성 벌 마비 바이러스(Acute bee paralysis virus), 히비스커스 위황병 원형반점 바이러스(Hibiscus chlorotic ringspot virus), 돼지 열병 바이러스(Classical swine fever virus), 인간 FGF2, 인간 SFTPA1, 인간 AML1/RUNX1, 초파리 안테나페디아(Drosophila antennapedia), 인간 AQP4, 인간 AT1R, 인간 BAG-1, 인간 BCL2, 인간 BiP, 인간 c-IAP1, 인간 c-myc, 인간 eIF4G, 마우스 NDST4L, 인간 LEF1, 마우스 HIF1 알파, 인간 n.myc, 마우스 Gtx, 인간 p27kip1, 인간 PDGF2/c-sis, 인간 p53, 인간 Pim-1, 마우스 Rbm3, 초파리 리퍼(Drosophila reaper), 캐닌 스캠퍼(Canine Scamper), 초파리 Ubx, 인간 UNR, 마우스 UtrA, 인간 VEGF-A, 인간 XIAP, 털없는 초파리(Drosophila hairless), 에스. 세레비지애(S. cerevisiae) TFIID, 에스. 세레비지애 YAP1, 담배 식각 바이러스(tobacco etch virus), 터닙 주름 바이러스(turnip crinkle virus), EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, 피코비르나바이러스, HCV QC64, 인간 코사바이러스(Human Cosavirus) E/D, 인간 코사바이러스 F, 인간 코사바이러스 JMY, 리노바이러스(Rhinovirus) NAT001,

HRV14, HRV89, HRVC-02, HRV-A21, 살리바이러스(Salivirus) A SH1, 살리바이러스 FHB, 살리바이러스 NG-J1, 인간 파레코바이러스(Human Parechovirus) 1, 크로히바이러스(Crohivirus) B, Yc-3, 로사바이러스(Rosavirus) M-7, 산바바이러스(Shanbavirus) A, 파시바이러스(Pasivirus) A, 파시바이러스 A 2, 에코바이러스(Echovirus) E14, 인간 파레코바이러스(Human Parechovirus) 5, 아이치 바이러스(Aichi Virus), A형 간염 바이러스 HA16, 포피바이러스(Phopivirus), CVA10, 엔테로바이러스(Enterovirus) C, 엔테로바이러스 D, 엔테로바이러스 J, 인간 페기바이러스(Human Pegivirus) 2, GBV-C GT110, GBV-C K1737, GBV-C 아이오와(Iowa), 페기바이러스(Pegivirus) A 1220, 파시바이러스(Pasivirus) A 3, 사펠로바이러스(Sapelovirus), 로사바이러스(Rosavirus) B, 바쿤사(Bakunsa) 바이러스, 트레모바이러스(Tremovirus) A, 돼지 파시바이러스(Swine Pasivirus) 1, PLV-CHN, 파시바이러스 A, 시시니바이러스(Sicinivirus), 헤파시바이러스(Hepacivirus) K, 헤파시바이러스 A, BVDV1, 보더병 바이러스(Border Disease Virus), BVDV2, CSFV-PK15C, SF573 디시스트로바이러스(Dicistrovirus), 후베이 피코르나-유사 바이러스(Hubei Picorna-like Virus), CRPV, 살리바이러스 A BN5, 살리바이러스 A BN2, 살리바이러스 A 02394, 살리바이러스 A GUT, 살리바이러스 A CH, 살리바이러스 A SZ1, 살리바이러스 FHB, CVB3, CVB1, 에코바이러스 7, CVB5, EVA71, CVA3, CVA12, EV24 또는 eIF4G에 대한 압타머로부터의 IRES의 서열을 갖는다.

[0030] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 자연 뉴클레오타이드를 포함한다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 자연 뉴클레오타이드로 이루어진다. 일부 실시형태에서, 발현 서열은 코돈 최적화된다.

[0031] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 마이크로RNA 결합 부위가 결여되도록 최적화된다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 엔도뉴클레아제 민감성 부위가 결여되도록 최적화된다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 RNA-편집 민감성 부위가 결여되도록 최적화된다.

[0032] 일부 실시형태에서, 상기 청구항 중 어느 한 항에 있어서, 상기 원형 RNA 폴리뉴클레오타이드는 약 100개 뉴클레오타이드 내지 약 15킬로베이스 길이인 원형 RNA 뉴클레오타이드이다.

[0033] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 인간에서 적어도 약 20시간의 생체내 치료 효과 기간을 갖는다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 적어도 약 20시간의 기능성 반감기를 갖는다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 동일한 발현 서열을 포함하는 등가의 선형 RNA 폴리뉴클레오타이드의 인간 세포에서의 치료 효과 기간보다 더 크거나 동일한 인간 세포에서의 치료 효과 기간을 갖는다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 동일한 발현 서열을 포함하는 등가의 선형 RNA 폴리뉴클레오타이드의 인간 세포에서의 기능성 반감기보다 더 크거나 동일한 인간 세포에서의 기능성 반감기를 갖는다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 동일한 발현 서열을 갖는 등가의 선형 RNA 폴리뉴클레오타이드의 인간에서의 생체내 치료 효과 기간보다 큰 인간에서의 생체내 치료 효과 기간을 갖는다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 동일한 발현 서열을 갖는 등가의 선형 RNA 폴리뉴클레오타이드의 인간에서의 생체내 기능성 반감기보다 큰 인간에서의 생체내 기능성 반감기를 갖는다.

[0034] 또 다른 양상에서, 본 출원은 본 명세서에 기재된 바와 같은 원형 RNA 폴리뉴클레오타이드, 나노입자 및 선택적으로, 나노 입자에 작동 가능하게 연결된 표적화 모이어티를 포함하는 약제학적 조성물을 제공한다. 일부 실시형태에서, 나노입자는 지질 나노입자, 코어-쉘 나노입자, 생분해성 나노입자, 생분해성 지질 나노입자, 중합체 나노입자 또는 생분해성 중합체 나노입자이다.

[0035] 일부 실시형태에서, 약제학적 조성물은 표적화 모이어티를 포함하되, 표적화 모이어티는 세포 단리 또는 정제의 부재 하에서 수용체-매개된 엔도사이토시스 또는 선택된 세포 집단 또는 조직의 선택된 세포 내로의 직접 융합을 매개한다. 일부 실시형태에서, 약제학적 조성물은 나노입자에 작동 가능하게 연결된 표적화 모이어티를 포함한다. 일부 실시형태에서, 표적화 모이어티는 scFv, 나노바디, 펩타이드, 미니바디, 폴리뉴클레오타이드 압타머, 중쇄 가변 영역, 경쇄 가변 영역 또는 이의 단편이다.

[0036] 일부 실시형태에서, 약제학적 조성물은 이중 가닥 RNA, DNA 플린트 또는 삼인산화된 RNA인 조성물 중에 1중량% 미만의 폴리뉴클레오타이드를 포함한다. 일부 실시형태에서, 약제학적 조성물은 이중 가닥 RNA, DNA 스플린트, 삼인산화된 RNA, 포스파타제 단백질, 단백질 리가제 및 캡핑 효소인 약제학적 조성물 중에 1중량% 미만의 폴리뉴클레오타이드 및 단백질을 포함한다.

[0037] 또 다른 양상에서, 본 출원은 치료를 필요로 하는 대상체의 치료 방법으로서, 이 방법은 본 명세서에 기재된 원형 RNA 폴리뉴클레오타이드, 나노입자 및 선택적으로, 상기 나노입자에 작동 가능하게 연결된 표적화 모이어티를 포함하는, 치료적 유효량의 조성물을 투여하는 단계를 포함한다.

[0038] 일부 실시형태에서, 조성물은 표적화 모이어티를 포함하되, 표적화 모이어티는 선택적으로 세포 선택 또는 정제의 부재 하에서 선택된 세포 집단의 세포 내에서 수용체-매개된 엔도사이토시스를 매개한다. 일부 실시형태에서, 표적화 모이어티는 scFv, 나노바디, 펩타이드, 미니바디, 중쇄 가변 영역, 경쇄 가변 영역 또는 이의 단편이다. 일부 실시형태에서, 나노입자는 지질 나노입자, 코어-쉘 나노입자 또는 생분해성 나노입자이다.

[0039] 일부 실시형태에서, 나노입자는 1종 이상의 양이온성 지질, 이온화 가능한 지질 또는 폴리 β-아미노 에스터를 포함한다. 일부 실시형태에서, 나노입자는 1종 이상의 비-양이온성 지질을 포함한다. 일부 실시형태에서, 나노입자는 1종 이상의 PEG-변형된 지질, 폴리글루탐산 지질 또는 히알루론산 지질을 포함한다. 일부 실시형태에서, 나노입자는 콜레스테롤을 포함한다. 일부 실시형태에서, 나노입자는 아라키돈산 또는 올레산을 포함한다. 일부 실시형태에서, 나노입자는 1종 초과의 원형 RNA 폴리뉴클레오타이드를 포함한다.

[0040] 일부 실시형태에서, 대상체는 급성 림프구성 백혈병; 급성 골수성 백혈병(acute lymphocytic leukemia: AML); 포상횡문근육종; B 세포 악성종양; 방광암(예를 들어, 방광 암종); 골암; 뇌암(예를 들어, 수모세포종); 유방암; 항문, 항문관 또는 항문직장의 암; 안암; 간내담관의 암; 관절암; 두부암; 담낭암; 흉막의 암; 코, 비강 또는 중이의 암; 구강암; 음문암; 만성 림프구성 백혈병; 만성 골수성 암; 결장암; 식도암; 자궁경부암; 섬유육종; 위장 카르시노이드 종양; 두경부암(예를 들어, 두경부 편평 세포 암종); 호지킨 림프종; 하인두암; 신장암; 후두암; 백혈병; 액체 종양; 간암; 폐암(예를 들어, 비소세포 폐 암종 및 폐 선암종); 림프종; 중피종; 비만세포종; 흑색종; 다발성 골수종; 코인두암; 비호지킨 림프종; B-만성 림프구성 백혈병; 털세포 백혈병; 급성 림프구성 백혈병(acute lymphocytic leukemia: ALL); 버킷 림프종, 난소암; 췌장암; 복막암; 대망(omentum)암; 장간막암; 인두암; 전립선암; 직장암; 신장암; 피부암; 소장암; 연조직암; 고형 종양; 활막 육종; 위암; 고환암; 갑상선암; 및 요관암으로 이루어진 군으로부터 선택된 암을 갖는다. 일부 실시형태에서, 대상체는 경피증, 그레이브스병, 크론병, 쇼그렌병, 다발성 경화증, 하시모토병, 건선, 중증 근무력증, 자가면역 다발성 내분비병증 증후군(autoimmune polyendocrinopathy syndrome), 제I형 당뇨병(TIDM), 자가면역 위염, 자가면역 포도막염, 다발성근염, 대장염, 갑상선염 및 인간 루푸스로 대표되는 전신성 자가면역 질환(generalized autoimmune diseases typified by human Lupus)으로부터 선택되는 자가면역 장애를 갖는다.

[0041] 또 다른 양상에서, 본 출원은 원형 RNA 폴리뉴클레오타이드의 제조를 위한 벡터를 제공하며, 원형 RNA 폴리뉴클레오타이드는 다음 순서로, 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 제2 발현 서열, 5' 그룹 I 인트론 단편 및 3' 듀플렉스 형성 영역을 포함한다.

[0042] 일부 실시형태에서, 벡터는 제1 발현 서열과 제2 발현 서열 사이에 절단 부위를 암호화하는 폴리뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 절단 부위는 자가-절단 스페이서이다. 일부 실시형태에서, 자가-절단 스페이서는 2A 자가-절단 펩타이드이다.

[0043] 일부 실시형태에서, 벡터는 5' 듀플렉스 형성 영역과 3' 그룹 I 인트론 단편 사이의 제1 스페이서 및 5' 그룹 I 인트론 단편과 3' 듀플렉스 형성 영역 사이의 제2 스페이서를 포함한다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서는 각각 약 5 내지 약 60개의 뉴클레오타이드의 길이를 갖는다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서스페이서 각각은 적어도 5개 뉴클레오타이드 길이의 비구조화된 영역을 포함한다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서스페이서 각각은 적어도 7개 뉴클레오타이드 길이의 구조화된 영역을 포함한다. 일부 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 9 내지 약 50개 뉴클레오타이드의 길이를 갖는다.

[0044] 일부 실시형태에서, 벡터는 코돈 최적화된다. 일부 실시형태에서, 벡터는 등가의 사전 최적화 폴리뉴클레오타이드에 존재하는 적어도 1개의 마이크로RNA 결합 부위가 결여되어 있다.

[0045] 또 다른 양상에서, 본 출원은 본 명세서에 기재된 바와 같은 원형 RNA 폴리뉴클레오타이드를 포함하는 진핵 세포를 제공한다. 일부 실시형태에서, 진핵 세포는 인간 세포이다. 일부 실시형태에서, 진핵 세포는 면역 세포이다. 일부 실시형태에서, 진핵 세포는 T 세포이다.

[0046] 양상에서, 본 명세서에는 다음 순서로, 포스트 스플라이싱 3' 그룹 I 인트론 단편, 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, 제2 발현 서열 및 포스트 스플라이싱 5' 그룹 I 인트론 단편을 포함하는 원형 RNA 폴리뉴클레오타이드가 제공된다. 양상에서, 본 명세서에는 다음

순서로, 포스트 스플라이싱 3' 그룹 I 인트론 단편, 제1 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 제2 IRES, 제2 발현 서열 및 포스트 스플라이싱 5' 그룹 I 인트론 단편을 포함하는 원형 RNA 폴리뉴클레오타이드가 제공된다. 일부 실시형태에서, 제1 발현 서열 및 제2 발현 서열은 상이한 치료용 단백질을 암호화한다. 일부 실시형태에서, 제1 발현 서열 및 제2 발현 서열은 동일한 치료용 단백질을 암호화한다.

[0047] 양상에서, 본 명세서에는 다음 순서로, 선택적인 5' 듀플렉스 형성 영역, 포스트 스플라이싱 3' 그룹 I 인트론 단편, 내부 리보솜 유입 부위(IRES), 발현 서열, 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, 제2 발현 서열, 5' 그룹 I 인트론 단편 및 선택적인 3' 듀플렉스 형성 영역을 포함하는 벡터의 전사로부터 생산된 원형 RNA 폴리뉴클레오타이드가 제공된다. 양상에서, 본 명세서에는 다음 순서로, 선택적인 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 제1 내부 리보솜 유입 부위(IRES), 발현 서열, 제2 IRES, 제2 발현 서열, 5' 그룹 I 인트론 단편, 및 선택적인 3' 듀플렉스 형성 영역을 포함하는 벡터의 전사로부터 생산된 원형 RNA 폴리뉴클레오타이드가 제공된다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드 또는 벡터는 3' 및 5' 듀플렉스 형성 영역을 포함한다.

[0048] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 5' 듀플렉스 형성 영역과 포스트 스플라이싱 3' 그룹 I 인트론 단편 사이에 제1 스페이서, 및 포스트 스플라이싱 5' 그룹 I 인트론 단편과 3' 듀플렉스 형성 영역 사이에 제2 스페이서를 포함한다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서는 각각 약 10 내지 약 60개의 뉴클레오타이드의 길이를 갖는다. 특정 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 9 내지 약 19개의 뉴클레오타이드의 길이를 갖는다. 특정 다른 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 30개의 뉴클레오타이드의 길이를 갖는다.

[0049] 특정 실시형태에서, IRES는 표 17로부터 선택되거나, IRES의 서열을 갖거나, 이의 기능성 단편 또는 변이체이다. 일부 실시형태에서, IRES는 타우라 증후군 바이러스, 트리아토마 바이러스, 타일러 뇌척수염 바이러스, 유인원 바이러스 40, 붉은 불개미 바이러스 1, 로팔로시품 파디 바이러스, 세망내피증 바이러스, 인간 폴리오바이러스 1, 갈색날개 노린재 장 바이러스, 카슈미르 벌 바이러스, 인간 리노바이러스 2, 호말로디스카 코아굴라타 바이러스-1(Homalodisca coagulata virus-1), 인간 면역결핍 바이러스 타입 1, 호말로디스카 코아굴라타 바이러스-1(Homalodisca coagulata virus-1), 히메토비 P 바이러스(Himetobi P virus), C형 간염 바이러스, A형 간염 바이러스, 간염 GB 바이러스, 구제역 바이러스, 인간 엔테로바이러스 71, 말 비염 바이러스, 엑트로피스 오블리쿠아 피코르나-유사 바이러스, 뇌심근염 바이러스, 초파리 C 바이러스, 인간 콕사키바이러스 B3, 크루시퍼 토바모바이러스, 귀뚜라미 마비병 바이러스, 소 바이러스성 설사 바이러스 1, 블랙 퀸 세포 바이러스, 진딧물 치명 마비 바이러스, 조류 뇌척수염 바이러스, 급성 벌 마비 바이러스, 히비스커스 위황병 원형반점 바이러스, 돼지 열병 바이러스, 인간 FGF2, 인간 SFTPA1, 인간 AML1/RUNX1, 초파리 안테나페디아, 인간 AQP4, 인간 AT1R, 인간 BAG-1, 인간 BCL2, 인간 BiP, 인간 c-IAP1, 인간 c-myc, 인간 eIF4G, 마우스 NDST4L, 인간 LEF1, 마우스 HIF1 알파, 인간 n.myc, 마우스 Gtx, 인간 p27kip1, 인간 PDGF2/c-sis, 인간 p53, 인간 Pim-1, 마우스 Rbm3, 초파리 리퍼, 캐닌 스캠퍼 초파리 Ubx, 인간 UNR, 마우스 UtrA, 인간 VEGF-A, 인간 XIAP, 털없는 초파리, 에스. 세레비지애 TFIID, 에스. 세레비지애 YAP1, 담배 식각 바이러스, 터닙 주름 바이러스, EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, 피코비르나바이러스, HCV QC64, 인간 코사바이러스 E/D, 인간 코사바이러스 F, 인간 코사바이러스 JMY, 리노바이러스 NAT001, HRV14, HRV89, HRVC-02, HRV-A21, 살리바이러스 A SH1, 살리바이러스 FHB, 살리바이러스 NG-J1, 인간 파레코바이러스 1, 크로히바이러스 B, Yc-3, 로사바이러스 M-7, 산바바이러스 A, 파시바이러스 A, 파시바이러스 A 2, 에코바이러스 E14, 인간 파레코바이러스 5, 아이치 바이러스, A형 간염 바이러스 HA16, 포피바이러스, CVA10, 엔테로바이러스 C, 엔테로바이러스 D, 엔테로바이러스 J, 인간 페기바이러스 2, GBV-C GT110, GBV-C K1737, GBV-C 아이오와, 페기바이러스 A 1220, 파시바이러스 A 3, 사펠로바이러스, 로사바이러스 B, 바쿤사 바이러스, 트레모바이러스 A, 돼지 파시바이러스 1, PLV-CHN, 파시바이러스 A, 시시니바이러스, 헤파시바이러스 K, 헤파시바이러스 A, BVDV1, 보더병 바이러스, BVDV2, CSFV-PK15C, SF573 디시스트로바이러스, 후베이 피코르나-유사 바이러스, CRPV, 살리바이러스 A BN5, 살리바이러스 A BN2, 살리바이러스 A 02394, 살리바이러스 A GUT, 살리바이러스 A CH, 살리바이러스 A SZ1, 살리바이러스 FHB, CVB3, CVB1, 에코바이러스 7, CVB5, EVA71, CVA3, CVA12, EV24 또는 eIF4G에 대한 압타머로부터의 IRES의 서열을 갖는다.

[0050] 일부 실시형태에서, 제1 스페이서 및 제2 스페이서polyA 서열은 각각 15 내지 50nt의 길이를 갖는다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서polyA 서열은 각각 약 20 내지 25nt의 길이를 갖는다.

[0051] 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 자연 발생 뉴클레오타이드로 이루어진다. 일부 실시형태에서, 원형 RNA는 적어도 약 80%, 적어도 90%, 적어도 약 95%, 또는 적어도 약 99%의 자연 발생 뉴클레오타이드를

함유한다. 특정 실시형태에서, 발현 서열은 코돈-최적화된다. 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 마이크로RNA 결합 부위가 결여되도록 최적화된다. 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 엔도뉴클레아제 민감성 부위가 결여되도록 최적화된다. 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 등가의 사전 최적화된 폴리뉴클레오타이드에 존재하는 적어도 하나의 RNA-편집 민감성 부위가 결여되도록 최적화된다.

[0052] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 약 100개 뉴클레오타이드 내지 약 15킬로베이스 길이이다. 특정 실시형태에서, 보 개시내용의 원형 RNA 폴리뉴클레오타이드는 인간에서 적어도 약 20시간의 생체내 치료 효과 기간을 갖는다. 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 적어도 약 20시간의 기능성 반감기를 갖는다. 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 동일한 발현 서열을 포함하는 등가의 선형 RNA 폴리뉴클레오타이드의 인간 세포에서의 치료 효과 기간보다 더 크거나 동일한 인간 세포에서의 치료 효과 기간을 갖는다. 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 동일한 발현 서열을 포함하는 등가의 선형 RNA 폴리뉴클레오타이드의 인간 세포에서의 기능성 반감기보다 더 크거나 동일한 인간 세포에서의 기능성 반감기를 갖는다. 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 동일한 발현 서열을 갖는 등가의 선형 RNA 폴리뉴클레오타이드의 인간에서의 생체내 기능성 반감기보다 큰 인간에서의 생체내 기능성 반감기를 갖는다. 일부 실시형태에서, 참조 선형 RNA 폴리뉴클레오타이드는 cap1 구조 및 polyA 테일을 포함하는 적어도 80nt 길이의 선형, 비변형된 또는 뉴클레오사이드-변형된, 완전-가공된 mRNA이다.

[0053] 일부 실시형태에서, 약제학적 조성물은 미리 결정된 역치 값의 것과 동일하거나 더 큰 인간 세포에서의 기능성 반감기를 갖는다. 일부 실시형태에서, 약제학적 조성물은 인간에서 미리 결정된 역치 값의 것보다 더 큰 생체내 기능성 반감기를 갖는다. 일부 실시형태에서, 기능성 단백질 검정은 시험과내 루시퍼라제 검정이다. 일부 실시형태에서, 기능성 단백질 검정은 환자 혈청 또는 조직 샘플에서 원형 RNA 폴리뉴클레오타이드의 발현 서열에 의해서 암호화된 단백질의 수준을 측정하는 것을 포함한다. 일부 실시형태에서, 미리 결정된 역치 값은 원형 RNA 폴리뉴클레오타이드와 동일한 발현 서열을 포함하는 참조 선형 RNA 폴리뉴클레오타이드의 기능성 반감기이다. 일부 실시형태에서, 약제학적 조성물은 적어도 약 20시간의 기능성 반감기를 갖는다.

[0054] 양상에서, 본 명세서에는 원형 RNA 폴리뉴클레오타이드, 나노입자 및 선택적으로, 나노입자에 작동 가능하게 연결된 표적화 모이어티를 포함하는 약제학적 조성물이 제공된다. 일부 실시형태에서, 나노입자는 지질 나노입자, 코어-쉘 나노입자, 생분해성 나노입자, 생분해성 지질 나노입자, 중합체 나노입자 또는 생분해성 중합체 나노입자이다. 특정 실시형태에서, 나노입자는 군 C12-200, MC3, DLinDMA, DLinkC2DMA, cKK-E12, ICE(이미다졸계), HGT5000, HGT5001, DODAC, DDAB, DMRIE, DOSPA, DOGS, DODAP, DODMA 및 DMDMA, DODAC, DLenDMA, DMRIE, CLinDMA, CpLinDMA, DMOBA, DOcarbDAP, DLinDAP, DLincarbDAP, DLinCDAP, KLin-K-DMA, DLin-K-XTC2-DMA, HGT4003 및 이들의 조합물로부터 선택된 1종 이상의 양이온성 지질이다.

[0055] 특정 실시형태에서, 약제학적 조성물은 표적화 모이어티를 포함하되, 표적화 모이어티는 세포 단리 또는 정제의 부재 하에서 수용체-매개된 엔도사이토시스 또는 선택된 세포 집단 또는 조직의 선택된 세포 내로의 직접 융합을 매개한다. 일부 실시형태에서, 표적화 모이어티는 scFv, 나노바디, 펩타이드, 미니바디, 폴리뉴클레오타이드 압타머, 중쇄 가변 영역, 경쇄 가변 영역 또는 이의 단편이다. 특정 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 자가면역 장애 또는 암의 치료를 필요로 하는 인간 대상체에서 자가면역 장애 또는 암을 치료하기에 효과적인 양이다. 특정 실시형태에서, 약제학적 조성물은 동일한 발현 서열을 암호화하는 외인성 DNA를 포함하는 벡터를 포함하는 약제학적 조성물과 비교하는 경우 향상된 안전성 프로파일을 갖는다.

[0056] 특정 실시형태에서, 조성물 중의 상기 폴리뉴클레오타이드의 1중량% 미만은 이중 가닥 RNA, DNA 스플린트 또는 삼인산화된 RNA이다. 특정 실시형태에서, 약제학적 조성물 중의 폴리뉴클레오타이드 및 단백질의 1중량% 미만은 이중 가닥 RNA, DNA 스플린트, 삼인산화된 RNA, 포스파타제 단백질, 단백질 리가제 및 캡핑 효소이다.

[0057] 양상에서, 본 명세서에는 치료를 필요로 하는 대상체의 치료 방법이 제공되며, 이 방법은 원형 RNA 폴리뉴클레오타이드, 나노입자 및 선택적으로, 상기 나노입자에 작동 가능하게 연결된 표적화 모이어티를 포함하는, 치료적 유효량의 조성물을 투여하는 단계를 포함한다. 일부 실시형태에서, 대상체는 자가면역 장애 또는 암을 갖는다.

[0058] 일부 실시형태에서, 표적화 모이어티는 scFv, 나노바디, 펩타이드, 미니바디, 중쇄 가변 영역, 경쇄 가변 영역 또는 이의 단편이다. 일부 실시형태에서, 조성물은 표적화 모이어티를 포함하되, 표적화 모이어티는 세포 단리 또는 정제의 부재 하에서 선택된 세포 집단의 선택적 세포 내에서 수용체-매개된 엔도사이토시스를 매개한다.

[0059] 일부 실시형태에서, 나노입자는 지질 나노입자, 코어-쉘 나노입자 또는 생분해성 나노입자이다. 일부 실시형태에서, 나노입자는 1종 이상의 양이온성 지질, 이온화 가능한 지질 또는 폴리 β-아미노 에스터를 포함한다. 일부 실시형태에서, 나노입자는 1종 이상의 비-양이온성 지질을 포함한다. 일부 실시형태에서, 나노입자는 1종 이상의 PEG-변형된 지질, 폴리글루탐산 지질, 구조 지질 또는 히알루론산 지질을 포함한다. 특정 실시형태에서, 나노입자는 콜레스테롤을 포함한다. 일부 실시형태에서, 구조 지질은 베타-시토스테롤이다. 일부 실시형태에서, 구조 지질은 베타 시토스테롤이 아니다. 일부 실시형태에서, 나노입자는 아라키돈산 또는 올레산을 포함한다. 일부 실시형태에서, 나노입자는 1종 초과의 원형 RNA 폴리뉴클레오타이드를 포함한다.

[0060] 일부 실시형태에서, 구조 지질은 C1q에 결합하고/하거나 구조 지질이 없는 대조군 전달 비히클과 비교하여 C1q에 대한 상기 지질을 포함하는 전달 비히클의 결합을 촉진하고/하거나 구조 지질이 없는 대조군 전달 비히클과 비교하여 면역 세포 내로의 C1q-결합된 전달 비히클의 흡수를 증가시킨다.

[0061] 일부 실시형태에서, PEG-변형된 지질은 DSPE-PEG, DMG-PEG 또는 PEG-1이다. 일부 실시형태에서, PEG-변형된 지질은 DSPE-PEG(2000)이다.

[0062] 일부 실시형태에서, 약제학적 조성물은 헬퍼 지질을 추가로 포함한다. 일부 실시형태에서, 헬퍼 지질은 DSPC 또는 DOPE이다.

[0063] 일부 실시형태에서, 약제학적 조성물은 DOPE, 콜레스테롤 및 DSPE-PEG를 포함한다.

[0064] 일부 실시형태에서, 전달 비히클은 몰비로 약 0.5% 내지 약 4%의 PEG-변형된 지질을 포함한다. 일부 실시형태에서, 전달 비히클은 몰비로 약 1% 내지 약 2%의 PEG-변형된 지질을 포함한다.

[0065] 일부 실시형태에서, 이온화 가능한 지질:DSPC:콜레스테롤:DSPE-PEG(2000)의 몰비는 62:4:33:1이다.

[0066] 일부 실시형태에서, 전달 비히클은 이온화 가능한 지질, DOPE, 콜레스테롤 및 DSPE-PEG(2000)를 포함한다.

[0067] 일부 실시형태에서, 이온화 가능한 지질:DSPC:콜레스테롤:DSPE-PEG(2000)의 몰비는 50:10:38.5:1.5이다.

[0068] 일부 실시형태에서, 전달 비히클은 약 3 내지 약 6의 질소:포스페이트(N:P) 비율을 갖는다.

[0069] 일부 실시형태에서, 전달 비히클은 원형 RNA 폴리뉴클레오타이드의 엔도좀 방출을 위해서 제형화된다.

[0070] 일부 실시형태에서, 전달 비히클은 APOE에 결합할 수 있다. 일부 실시형태에서, 전달 비히클은 원형 RNA 폴리뉴클레오타이드와 동일한 발현 서열을 갖는 참조 선형 RNA가 로딩된 등가의 전달 비히클보다 더 적은 아포지질단백질 E(APOE)과 상호작용한다. 일부 실시형태에서, 전달 비히클의 외부 표면은 APOE 결합 부위가 실질적으로 없다.

[0071] 일부 실시형태에서, 전달 비히클은 약 120nm 미만의 직경을 갖는다. 일부 실시형태에서, 전달 비히클은 300nm 초과의 직경을 갖는 응집물을 형성하지 않는다.

[0072] 일부 실시형태에서, 전달 비히클은 약 120nm 미만의 직경을 갖는다. 일부 실시형태에서, 전달 비히클은 300nm 초과의 직경을 갖는 응집물을 형성하지 않는다.

[0073] 일부 실시형태에서, 전달 비히클은 약 30시간 미만의 생체내 반감기를 갖는다.

[0074] 일부 실시형태에서, 전달 비히클은 세포 내로의 저밀도 지질단백질 수용체(LDLR) 의존적 흡수가 가능하다. 일부 실시형태에서, 전달 비히클은 세포 내로의 LDLR 의존적 흡수가 가능하다.

[0075] 일부 실시형태에서, 약제학적 조성물은 선형 RNA가 실질적으로 없다.

[0076] 일부 실시형태에서, 약제학적 조성물은 전달 비히클에 작동 가능하게 연결된 표적화 모이어티를 추가로 포함한다. 일부 실시형태에서, 표적화 모이어티는 면역 세포 항원에 또는 간접적으로 특이적으로 결합한다. 일부 실시형태에서, 면역 세포 항원은 T 세포 항원이다. 일부 실시형태에서, T 세포 항원은 CD2, CD3, CD5, CD7, CD8, CD4, 베타7 인테그린, 베타2 인테그린 및 C1q로 이루어진 군으로부터 선택된다.

[0077] 일부 실시형태에서, 약제학적 조성물은 전달 비히클 결합 모이어티 및 세포 결합 모이어티를 포함하는 어댑터 분자를 추가로 포함하되, 표적화 모이어티는 전달 비히클 결합 모이어티에 특이적으로 결합하고, 세포 결합 모이어티는 표적 세포 항원에 특이적으로 결합한다. 일부 실시형태에서, 표적 세포 항원은 면역 세포 항원이다. 일부 실시형태에서, 면역 세포 항원은 T 세포 항원, NK 세포, NKT 세포, 대식세포 또는 호중구이다. 일부 실시형태에서, T 세포 항원은 CD2, CD3, CD5, CD7, CD8, CD4, 베타7 인테그린, 베타2 인테그린, CD25, CD39, CD73,

A2a 수용체, A2b 수용체 및 C1q로 이루어진 군으로부터 선택된다. 일부 실시형태에서, 면역 세포 항원은 대식세포 항원이다. 일부 실시형태에서, 대식세포 항원은 만노스 수용체, CD206 및 C1q로 이루어진 군으로부터 선택된다.

[0078] 일부 실시형태에서, 표적화 모이어티는 소분자이다. 일부 실시형태에서, 소분자는 면역 세포 상의 엑토효소에 결합하되, 엑토효소는 CD38, CD73, 아데노신 2a 수용체 및 아데노신 2b 수용체로 이루어진 군으로부터 선택된다. 일부 실시형태에서, 소분자는 만노스, 레시틴, 아시비신, 바이오틴 또는 다이곡시게닌이다.

[0079] 일부 실시형태에서, 표적화 모이어티는 단일쇄 Fv(scFv) 단편, 나노바디, 펩타이드, 펩타이드-기재 마크로사이클, 미니바디, 소분자 리간드, 예컨대, 엽산염, 아르기닐글리실아스파트산(RGD) 또는 페놀-용해성 모둘린 알파 1 펩타이드(PSMA1), 중쇄 가변 영역, 경쇄 가변 영역 또는 이의 단편이다.

[0080] 일부 실시형태에서, 이온화 가능한 지질은 약 2주 미만의 세포 막에서의 반감기를 갖는다. 일부 실시형태에서, 이온화 가능한 지질은 약 1주 미만의 세포 막에서의 반감기를 갖는다. 일부 실시형태에서, 이온화 가능한 지질은 약 30시간 미만의 세포 막에서의 반감기를 갖는다. 일부 실시형태에서, 이온화 가능한 지질은 원형 RNA 폴리뉴클레오타이드의 기능성 반감기보다 짧은 세포 막에서의 반감기를 갖는다.

[0081] 또 다른 양상에서, 본 출원은 유효량의 본 명세서에 개시된 약제학적 조성물을 투여하는 단계를 포함하는, 질환, 장애 또는 병태를 치료 또는 예방하는 방법을 제공한다. 일부 실시형태에서, 질환, 장애 또는 병태는 표 27 또는 표 28로부터 선택된 폴리펩타이드의 비정상적인 발현, 활성 또는 국지화와 연관된다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 치료용 단백질을 암호화한다. 일부 실시형태에서, 비장에서 치료용 단백질 발현은 간에서의 치료용 단백질 발현보다 더 높다. 일부 실시형태에서, 비장에서 치료용 단백질 발현은 간에서의 치료용 단백질 발현의 적어도 약 2.9x이다. 일부 실시형태에서, 치료용 단백질은 간에서 기능성 수준으로 발현되지 않는다. 일부 실시형태에서, 치료용 단백질은 간에서 검출 가능한 수준으로 발현되지 않는다. 일부 실시형태에서, 비장에서 치료용 단백질 발현은 총 치료용 단백질 발현의 적어도 약 63%이다.

[0082] 일부 실시형태에서, 선형 RNA 폴리뉴클레오타이드는 3' 아나베나 그룹 I 인트론 단편 및 5' 아나베나 그룹 I 인트론 단편을 포함한다. 일부 실시형태에서, 참조 RNA 폴리뉴클레오타이드는 참조 3' 아나베나 그룹 I 인트론 단편 및 참조 5' 아나베나 그룹 I 인트론 단편을 포함한다. 일부 실시형태에서, 참조 3' 아나베나 그룹 I 인트론 단편 및 참조 5' 아나베나 그룹 I 인트론 단편은 L6-5 순열(permutation) 부위를 사용하여 생성되었다. 일부 실시형태에서, 3' 아나베나 그룹 I 인트론 단편 및 5' 아나베나 그룹 I 인트론 단편은 L6-5 순열 부위를 사용하여 생성되지 않았다. 일부 실시형태에서, 3' 아나베나 그룹 I 인트론 단편은 서열번호 112 내지 123 및 125 내지 150으로부터 선택된 서열을 포함하거나 이들로 이루어진다. 일부 실시형태에서, 5' 아나베나 그룹 I 인트론 단편은 서열번호 73 내지 84 및 86 내지 111로부터 선택된 상응하는 서열을 포함한다. 일부 실시형태에서, 5' 아나베나 그룹 I 인트론 단편은 서열번호 73 내지 84 및 86 내지 111로부터 선택된 서열을 포함하거나 이들로 이루어진다. 일부 실시형태에서, 3' 아나베나 그룹 I 인트론 단편은 서열번호 112 내지 124 및 125 내지 150으로부터 선택된 상응하는 서열을 포함하거나 이들로 이루어진다.

[0083] 일부 실시형태에서, IRES는 서열번호 348 내지 351로부터 선택된 뉴클레오타이드 서열을 포함한다. 일부 실시형태에서, 참조 IRES는 CVB3이다. 일부 실시형태에서, IRES는 CVB3이 아니다. 일부 실시형태에서, IRES는 서열번호 1 내지 64 및 66 내지 72로부터 선택된 서열을 포함한다.

[0084] 또 다른 양상에서, 본 출원은 본 명세서에 개시된 선형 RNA로부터 생산된 원형 RNA 폴리뉴클레오타이드를 개시한다.

[0085] 또 다른 양상에서, 본 출원은 5'에서부터 3'로, 3' 그룹 I 인트론 단편, IRES, 발현 서열 및 5' 그룹 I 인트론 단편을 포함하는 원형 RNA를 개시하되, IRES는 서열번호 348 내지 351로부터 선택된 뉴클레오타이드 서열을 포함한다.

[0086] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 3' 그룹 I 인트론 단편과 IRES 사이에 스페이서를 추가로 포함한다.

[0087] 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드는 듀플렉스를 형성할 수 있는 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역을 추가로 포함한다. 일부 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 9 내지 약 19개 뉴클레오타이드의 길이를 갖는다. 일부 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 30개의 뉴클레오타이드의 길이를 갖는다.

[0088] 양상에서, 본 명세서에는 다음 순서로, 선택적인 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, 제2 발현 서열, 5' 그룹 I 인트론 단편 및 선택적인 3' 듀플렉스 형성 영역을 포함하는, 원형 RNA 폴리뉴클레오타이드를 제조하기 위한 벡터가 제공된다. 양상에서, 본 명세서에는 다음 순서로, 선택적인 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 제1 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 제2 IRES, 제2 발현 서열, 5' 그룹 I 인트론 단편 및 선택적인 3' 듀플렉스 형성 영역을 포함하는, 원형 RNA 폴리뉴클레오타이드를 제조하기 위한 벡터가 제공된다. 일부 실시형태에서, 폴리뉴클레오타이드는 3' 듀플렉스 형성 영역 및 5' 듀플렉스 형성 영역을 함유한다. 특정 실시형태에서, 벡터는 5' 듀플렉스 형성 영역과 3' 그룹 I 인트론 단편 사이의 제1 스페이서 및 5' 그룹 I 인트론 단편과 3' 듀플렉스 형성 영역 사이의 제2 스페이서를 포함한다. 특정 실시형태에서, 제1 스페이서 및 제2 스페이서는 각각 약 5 내지 약 60개의 뉴클레오타이드의 길이를 갖는다. 특정 실시형태에서, 제1 스페이서 및 제2 스페이서는 각각 약 8 내지 약 60개의 뉴클레오타이드의 길이를 갖는다. 특정 실시형태에서, 제1 스페이서 및 제2 스페이서스페이서 각각은 적어도 5개 뉴클레오타이드 길이의 비구조화된 영역을 포함한다. 특정 실시형태에서, 제1 스페이서 및 제2 스페이서스페이서 각각은 적어도 7개 뉴클레오타이드 길이의 구조화된 영역을 포함한다. 특정 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 각각 약 9 내지 50개의 뉴클레오타이드의 길이를 갖는다. 특정 실시형태에서, 벡터는 코돈 최적화된다. 특정 실시형태에서, 벡터는 등가의 사전 최적화 폴리뉴클레오타이드에 존재하는 적어도 1개의 마이크로RNA 결합 부위가 결여되어 있다.

[0089] 양상에서, 본 명세서에는 다음 순서로, 선택적인 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, 제2 발현 서열, 5' 그룹 I 인트론 단편 및 선택적인 3' 듀플렉스 형성 영역을 포함하는 원형 RNA 폴리뉴클레오타이드를 제조하기 위한 벡터를 포함하는 원핵 세포가 제공된다. 일부 실시형태에서, 폴리뉴클레오타이드는 3' 듀플렉스 형성 영역 및 5' 듀플렉스 형성 영역을 함유한다. 양상에서, 본 명세서에는 다음 순서로, 선택적인 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 제1 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 제2 IRES, 제2 발현 서열, 5' 그룹 I 인트론 단편 및 선택적인 3' 듀플렉스 형성 영역을 포함하는, 원형 RNA 폴리뉴클레오타이드를 제조하기 위한 벡터를 포함하는 원핵 세포가 제공된다. 일부 실시형태에서, 폴리뉴클레오타이드는 3' 듀플렉스 형성 영역 및 5' 듀플렉스 형성 영역을 함유한다.

[0090] 양상에서, 본 명세서에는 본 개시내용의 원형 RNA 폴리뉴클레오타이드를 포함하는 진핵 세포가 제공된다. 일부 실시형태에서, 진핵 세포는 인간 세포이다. 일부 실시형태에서, 진핵 세포는 면역 세포이다. 일부 실시형태에서, 진핵 세포는 T 세포, NK 세포, NKT 세포, 대식세포 또는 호중구이다.

도면의 간단한 설명

[0091] **도 1**은 가우시아 루시퍼라제 발현 서열 및 다양한 IRES 서열을 포함하는 원형 RNA로 형질주입시킨 지 24시간 후 HEK293(도 1A, 도 1D 및 도 1E), HepG2(도 1B), 또는 1C1C7(도 1C) 세포의 상청액에서의 발광을 도시한 도면.

도 2는 가우시아 루시퍼라제 발현 서열 및 상이한 길이를 갖는 다양한 IRES 서열을 포함하는 원형 RNA로 형질주입시킨 지 24시간 후 HEK293(도 2A), HepG2(도 2B), 또는 1C1C7(도 2C) 세포의 상청액에서의 발광을 도시한 도면.

도 3은 발광에 의해 측정된 바와 같은 3일에 걸친 HepG2(도 3A) 또는 1C1C7(도 3B) 세포에서 선택된 IRES 작제물의 안정성을 도시한 도면.

도 4A 및 **도 4B**는 세포 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광에 의해 측정된 바와 같은, 주카트 세포에서 선택된 IRES 작제물로부터의 단백질 발현을 도시한 도면.

도 5A 및 **도 5B**는 발광에 의해 측정된 바와 같은 3일 동안 주카트 세포에서 선택된 IRES 작제물의 안정성을 도시한 도면.

도 6은 가우시아 루시퍼라제를 암호화하는 변형된 선형, 비정제된 원형 또는 정제된 원형 RNA의 24시간 발광(도 6A) 또는 3일에 걸친 상대 발광(도 6B)의 비교를 도시한 도면.

도 7은 변형된 선형, 비정제된 원형 또는 정제된 원형 RNA로 주카트 세포를 전기천공한 후 IFNγ(도 7A), IL-6(도 7B), IL-2(도 7C), RIG-I(도 7D), IFN-β1(도 7E) 및 TNFα(도 7F)의 전사체 유도를 도시한 도면.

도 8은 인간 1차 단핵구(도 8A) 및 대식세포(도 8B 및 도 8C)에서 가우시아 루시퍼라제를 암호화하는 원형 RNA

및 변형된 선형 RNA의 발광의 비교를 도시한 도면.

도 9는 가우시아 루시퍼라제 발현 서열 및 다양한 IRES 서열을 포함하는 원형 RNA를 형질도입한 후 1차 T 세포의 상청액에서 3일에 걸친 상대 발광(도 9A) 또는 24시간 발광(도 9B)을 도시한 도면.

도 10은 가우시아 루시퍼라제 발현 서열을 포함하는 원형 RNA 또는 변형된 선형 RNA를 형질도입한 후 1차 T 세포의 상청액에서 24시간 발광(도 10A) 또는 3일에 걸친 상대 발광(도 10B) 및 PBMC에서 24시간 발광(도 10C)을 도시한 도면.

도 11은 상이한 순열 부위를 갖는 RNA 작제물의 HPLC 크로마토그램(도 11A) 및 원형화 효율(도 11B)을 도시한 도면.

도 12는 상이한 인트론 및/또는 순열 부위를 갖는 RNA 작제물의 HPLC 크로마토그램(도 12A) 및 원형화 효율(도 12B)을 도시한 도면.

도 13은 상동성 아암을 갖거나 갖지 않는 3개의 RNA 작제물의 HPLC 크로마토그램(도 13A) 및 원형화 효율(도 13B)을 도시한 도면.

도 14는 상동성 암이 없거나 다양한 길이 및 GC 함량을 갖는 상동성 아암을 갖지 않거나 상동성 아암을 갖는 3개의 RNA 작제물의 원형화 효율을 도시한 도면.

도 15A 및 **도 15B**는 개선된 스플라이싱 효율에 대한 강력한 상동성 아암의 기여, 선택된 작제물에서 원형화 효율과 니킹(nicking) 사이의 관계 및 개선된 원형화 효율을 입증하는 것으로 가정된 순열 부위와 상동성 아암의 조합을 도시한 HPLC 크로마토그램을 도시한 도면.

도 16은 모의 전기천공되거나(좌측) 또는 CAR을 암호화하는 원형 RNA로 전기천공된(우측)되고, GFP 및 파이어플라이 루시퍼라제를 발현하는 라지(Raji) 세포와 공배양된 T 세포의 형광 영상을 도시한 도면.

도 17은 모의 전기천공되거나(상단) 또는 CAR을 암호화하는 원형 RNA로 전기천공된(하단)되고, GFP 및 파이어플라이 루시퍼라제를 발현하는 라지 세포와 공배양된 T 세포의 명시야(좌측), 형광(중앙) 및 오버레이(우측) 영상을 도시한 도면.

도 18은 모의 전기천공되거나 상이한 CAR 서열을 암호화하는 원형 RNA로 전기천공된 T 세포에 의한 라지 표적 세포의 특이적 용해를 도시한 도면.

도 19는 가우시아 루시퍼라제 발현 서열 및 다양한 IRES 서열(도 19A)을 포함하는 선형 또는 원형 RNA를 형질도입한 지 24시간 후 주카트 세포(좌측) 또는 휴지기 1차 인간 CD3+ T 세포(우측)의 상청액에서의 발광 및 3일에 걸친 상대 발광(도 19B)을 도시한 도면.

도 20은 변형된 선형, 비정제된 원형, 또는 정제된 원형 RNA로 인간 CD3+ T 세포를 전기천공한 후 IFN-β1(도 20A), RIG-I(도 20B), IL-2(도 20C), IL-6(도 20D), IFNɣ(도 20E) 및 TNFα(도 20F)의 전사체 유도를 도시한 도면.

도 21은 파이어플라이 발광의 검출에 의해 결정된 바와 같은 CAR을 암호화하는 circRNA로 전기천공된 인간 1차 CD3+ T 세포에 의한 라지 표적 세포의 특이적 용해(도 21A) 및 상이한 양의 CAR 서열을 암호화하는 원형 또는 선형 RNA로 천기천공한 지 24시간 후 IFNɣ 전사체 유도(도 21B)를 도시한 도면.

도 22는 파이어플라이 발광의 검출에 의해 결정된 바와 같은 상이한 E:T 비에서 CAR을 암호화하는 원형 또는 선형 RNA로 전기천공된 인간 1차 CD3+ T 세포에 의한 표적 또는 비-표적 세포의 특이적 용해(도 22A 및 도 22B)를 도시한 도면.

도 23은 전기천공 후 1, 3, 5 및 7일에 CAR을 암호화하는 RNA로 전기천공된 인간 CD3+ T 세포에 의한 표적 세포의 특이적 용해를 도시한 도면.

도 24는 CD19 또는 BCMA 표적화 CAR을 암호화하는 원형 RNA로 전기천공된 인간 CD3+ T 세포에 의한 표적 세포의 특이적 용해를 도시한 도면.

도 25는 FLuc를 암호화하고, 50% 지질 15(표 10b), 10% DSPC, 1.5% PEG-DMG 및 38.5% 콜레스테롤과 제형화된 원형 RNA가 투여된 CD-1 마우스로부터 수거된 기관의 총 플럭스를 도시한 도면.

도 26은 FLuc를 암호화하고, 50% 지질 15(표 10b), 10% DSPC, 1.5% PEG-DMG 및 38.5% 콜레스테롤과 제형화된

원형 RNA가 투여된 CD-1 마우스로부터 수거된 기관의 발광을 강조한 영상을 도시한 도면.

도 27은 표 10a로부터의 지질 26 및 27의 분자 특징규명을 도시한 도면. **도 27A**는 지질 26의 양성자 핵자기 공명(NMR) 스펙트럼을 도시한 도면. **도 27B**는 액체 크로마토그래피-질량 분광광도법(LC-MS)에 의해서 측정된 지질 26의 체류 시간을 도시한 도면. **도 27C**는 지질 26의 질량 스펙트럼을 도시한 도면. **도 27D**는 지질 27의 양성자 NMR 스펙트럼을 도시한 도면. **도 27E**는 LC-MS에 의해서 측정된 지질 27의 체류 시간을 도시한 도면. **도 27F**는 지질 27의 질량 스펙트럼을 도시한 도면.

도 28은 지질 22-S14 및 이의 합성 중간체의 분자 특징규명을 도시한 도면. **도 28A**는 2-(테트라데실티오)에탄-1-올의 NMR 스펙트럼을 도시한 도면. **도 28B**는 2-(테트라데실티오)에틸 아크릴레이트의 NMR 스펙트럼을 도시한 도면. **도 28C**는 비스(2-(테트라데실티오)에틸) 3,3'-((3-(2-메틸-1H-이미다졸-1-일)프로필)아잔다이일)다이프로피오네이트(지질 22-S14)의 NMR 스펙트럼을 도시한 도면.

도 29는 비스(2-(테트라데실티오)에틸) 3,3'-((3-(1H-이미다졸-1-일)프로필)아잔다이일)다이프로피오네이트(지질 93-S14)의 NMR 스펙트럼을 도시한 도면.

도 30은 헵타데칸-9-일 8-((3-(2-메틸-1H-이미다졸-1-일)프로필)(8-(노닐옥시)-8-옥소옥틸)아미노)옥탄오에이트(표 10a로부터의 지질 54)의 분자 특징규명을 도시한 도면. **도 30A**는 지질 54의 양성자 NMR 스펙트럼을 도시한 도면. **도 30B**는 LC-MS에 의해서 측정된 지질 54의 체류 시간을 도시한 도면. **도 30C**는 지질 54의 질량 스펙트럼을 도시한 도면.

도 31은 헵타데칸-9-일 8-((3-(1H-이미다졸-1-일)프로필)(8-(노닐옥시)-8-옥소옥틸)아미노)옥탄오에이트(표 10a로부터의 지질 53)의 분자 특징규명을 도시한 도면. **도 31A**는 지질 53의 양성자 NMR 스펙트럼을 도시한 도면. **도 31B**는 LC-MS에 의해서 측정된 지질 53의 체류 시간을 도시한 도면. **도 31C**는 지질 53의 질량 스펙트럼을 도시한 도면.

도 32A는 파이어플라이 루시퍼라제(FLuc)를 암호화하고, 16:1:4:1 또는 62:4:33:1 몰비의 중량비의 관심 이온화 가능한 지질, DSPC, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사(Avanti Polar Lipids Inc.))과 제형화된 원형 RNA가 투여된 CD-1 마우스로부터 수거된 비장 및 간의 총 플럭스를 도시한 도면. **도 32B**는 단백질 발현의 생체분포에 대한 평균 라디언스를 도시한 도면.

도 33A는 FLuc를 암호화하고, 16:1:4:1 또는 62:4:33:1 몰비의 중량비로 이온화 가능한 지질 22-S14, DSPC, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사(Avanti Polar Lipids Inc.))과 제형화된 원형 RNA가 투여된 CD-1 마우스로부터 수거된 기간의 발광을 강조한 영상을 도시한 도면. **도 33B**는 FLuc를 암호화하고, 16:1:4:1 또는 62:4:33:1 몰비의 중량비로 이온화 가능한 지질 22-S14, DSPC, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사(Avanti Polar Lipids Inc.))과 제형화된 원형 RNA가 투여된 CD-1 마우스의 전신 IVIS 영상을 도시한 도면.

도 34A는 FLuc를 암호화하고, 16:1:4:1 또는 62:4:33:1 몰비의 중량비로 이온화 가능한 지질 93-S14, DSPC, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사)과 제형화된 원형 RNA가 투여된 CD-1 마우스로부터 수거된 기간의 발광을 강조한 영상을 도시한 도면. **도 34B**는 FLuc를 암호화하고, 16:1:4:1 또는 62:4:33:1 몰비의 중량비로 이온화 가능한 지질 93-S14, DSPC, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사)과 제형화된 원형 RNA가 투여된 CD-1 마우스의 전신 IVIS 영상을 도시한 도면.

도 35A는 FLuc를 암호화하고, 16:1:4:1 또는 62:4:33:1 몰비의 중량비의 표 10a로부터의 이온화 가능한 지질 26, DSPC, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사)과 제형화된 원형 RNA가 투여된 CD-1 마우스로부터 수거된 기간의 발광을 강조한 영상을 도시한 도면. **도 35B**는 FLuc를 암호화하고, 16:1:4:1 또는 62:4:33:1 몰비의 중량비로 이온화 가능한 지질 26, DSPC, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사)과 제형화된 원형 RNA가 투여된 CD-1 마우스의 전신 IVIS 영상을 도시한 도면.

도 36은 FLuc를 암호화하고, 표 10b로부터의 지질 15(도 **36A**), 표 10a로부터의 지질 53(도 **36B**) 또는 표 10a로부터의 지질 54(도 **36C**)로 형성된 지질 나노입자에 캡슐화된 원형 RNA가 투여된 c57BL/6J 마우스로부터 수거된 기간의 발광을 강조한 영상. PBS를 대조군으로 사용하였다(도 **36D**).

도 37A 및 도 37B는 파이어플라이 루시퍼라제를 암호화하는 원형 RNA를 함유하는 시험 지질 나노입자와 함께 24시간 인큐베이션시킨 후 인간 PBMC의 용해물에서의 상대적인 발광을 도시한 도면.

도 38은 GFP 또는 CD19 CAR을 암호화하는 원형 RNA를 함유하는 시험 지질 나노입자와 함께 인큐베이션시킨 후

인간 PBMC에서 GFP(도 **38A**) 및 CD19 CAR(도 **38B**)의 발현을 도시한 도면.

도 39는 항-뮤린 CD19 CAR 발현 서열 및 다양한 IRES 서열을 포함하는 원형 RNA가 지질형질주입된(lipotransfected) 1C1C7 세포에서 항-뮤린 CD19 CAR의 발현을 도시한 도면.

도 40은 뮤린 T 세포에 대한 항-뮤린 CD19 CAR의 세포독성을 도시한 도면. CD19 CAR은 뮤린 T 세포에 전기천공된 원형 RNA에 의해서 암호화되고, 이로부터 발현된다.

도 41은 항-뮤린 CD19 CAR을 암호화한 원형 RNA를 캡슐화한 시험 지질 나노입자를 격일로 주사한 C57BL/6J 마우스에서 말초 혈액(도 **40A** 및 **도 40B**) 또는 비장(도 **40C**)에서 B 세포 계수치를 도시한 도면.

도 42A 및 **도 42B**는 원형 RNA로부터 발현된 항-인간 CD19 CAR의 발현 수준을 선형 mRNA로부터 발현된 것과 비교한 도면.

도 43A 및 도 43B는 원형 RNA로부터 발현된 항-인간 CD19 CAR의 세포독성 효과를 선형 mRNA로부터 발현된 것과 비교한 도면.

도 44는 T 세포에서 단일 원형 RNA로부터 발현된 2개의 CAR(항-인간 CD19 CAR 및 항-인간 BCMA CAR)의 세포독성을 도시한 도면.

도 45A는 표 10a로부터의 지질 27 또는 26 또는 표 10b로부터의 지질 15로 형성된 LNP로의 처리 후 다양한 비장 면역 세포 하위세트에서 tdTomato 발현의 빈도와 함께 대표적인 FACS 플롯을 도시한 도면. **도 45B**는 Cre 원형 RNA가 성공적으로 형질주입된 각각의 세포 집단의 비율과 동등한, tdTomato를 발현하는 골수 세포, B 세포 및 T 세포의 비율의 정량(평균 + 표준 편차, n = 3)을 도시한 도면. **도 45C**는 지질 27 및 26으로의 처리 후 tdTomato를 발현하는, NK 세포, 전통적인 단핵구, 비전통적인 단핵구, 호중구 및 수지상 세포를 비롯한 추가 비장 면역 세포 집단의 비율(평균 + 표준 편차, n = 3)을 도시한 도면.

도 46A는 인트론 내에 빌트-인 polyA 서열을 갖는 예시적인 RNA 작제물을 도시한 도면. **도 46B**는 정제되지 않은 원형 RNA의 크로마토그래피 트레이스. **도 46C**는 친화성-정제된 원형 RNA의 크로마토그래피 트레이스. **도 46D**는 다양한 IVT 조건 및 정제 방법을 사용하여 제조된 원형 RNA의 면역원성을 도시한 도면(Commercial = 상업적인 IVT 혼합물; Custom = 항문, 항문관형 IVT 혼합물; Aff = 친화성 정제; Enz = 효소 정제; GMP:GTP 비율 = 8, 12.5 또는 13.75).

도 47A는 혼성화 정제를 위한 polyA에 대한 대안으로서 전용 결합 서열을 갖는 예시적인 RNA 작제물 설계를 도시한 도면. **도 47B**는 정제되지 않은 원형 RNA의 크로마토그래피 트레이스. **도 46C**는 친화성-정제된 원형 RNA의 크로마토그래피 트레이스.

도 48A는 디스트로핀을 암호화하는 정제되지 않은 원형 RNA의 크로마토그래피 트레이스. **도 48B**는 디스트로핀을 암호화하는 효소-정제된 원형 RNA의 크로마토그래피 트레이스.

도 49는 주카트 세포에서 3' 인트론 단편/5' 내부 듀플렉스 영역과 IRES 사이에 상이한 5' 스페이서를 갖는 정제된 circRNA의 발현(도 **49A**) 및 안정성(도 **49B**)을 비교한 도면. (AC = A 및 C만 스페이서 서열에 사용함; UC = U 및 C만 스페이서 서열에 사용됨).

도 50은 제시된 본래 또는 변형된 IRES 요소를 함유하는 원형 RNA로부터의 1차 T 세포에서의 발광 발현 수준 및 안정성을 도시한 도면.

도 51은 제시된 본래 또는 변형된 IRES 요소를 함유하는 원형 RNA로부터의 HepG2 세포에서의 발광 발현 수준 및 안정성을 도시한 도면.

도 52는 제시된 본래 또는 변형된 IRES 요소를 함유하는 원형 RNA로부터의 1C1C7 세포에서의 발광 발현 수준 및 안정성을 도시한 도면.

도 53은 삽입된 미번역 영역(UTR)을 갖는 IRES 요소 또는 혼성 IRES 요소를 함유하는 원형 RNA로부터의 HepG2 세포에서의 발광 발현 수준 및 안정성을 도시한 도면. "Scr"은 대조군으로 사용된 스크램블드(Scrambled)를 의미한다.

도 54는 *가우시아* 루시퍼라제 암호 서열에 작동 가능하게 연결된 가변 중단 코돈 카세트 및 IRES를 함유하는 원형 RNA로부터의 1C1C7 세포에서의 발광 발현 수준 및 안정성을 도시한 도면.

도 55는 가우시안 루시퍼라제 암호 서열의 시작 코돈 이전에 삽입된 가변 미번역 영역(UTR) 및 IRES를 함유하는

원형 RNA로부터의 1C1C7 세포에서의 발광 발현 수준 및 안정성을 도시한 도면.

도 56은 hEPO 암호 서열로부터의 하류에 2개의 miR-122 표적 부위를 함유하는 원형 RNA로부터의 Huh7 세포에서의 인간 에리트로포이에틴(hEPO)의 발현 수준을 도시한 도면.

도 57은 항-CD19 CAR을 발현하는 원형 RNA를 캡슐화한 LNP로 처리되는 경우 말초 혈액(도 **57A**) 및 비장(도 **57B**)에서 CAR 발현 수준을 도시한 도면. 항-CD20(aCD20) 및 루시퍼라제를 암호화하는 원형 RNA(oLuc)를 비교를 위해서 사용하였다.

도 58은 항-CD19 CAR 발현의 전체 빈도, 세포의 표면 상에서의 항-CD19 CAR 발현의 빈도 및 T-세포 상의 항-CD19 CAR을 암호화하는 IRES 특이적 원형 RNA의 항종양 반응에 대한 효과를 도시한 도면. **도 58A**는 항-CD19 CAR 기하 평균 형광 강도를 도시하고, **도 58B**는 항-CD19 CAR 발현 백분율을 도시하고, **도 58C**는 항-CD19 CAR에 의해 수행된 표적 세포 용해 백분율을 도시한 도면(CK = 염소 코브바이러스; AP = 등줄쥐 피코르나바이러스; CK* = 코돈 최적화를 갖는 염소 코브바이러스; PV = 파라보바이러스; SV = 살리바이러스).

도 59는 IRES 특이적 원형 RNA 작제물로 처리되는 경우 A20 FLuc 표적 세포의 CAR 발현 수준을 도시한 도면.

도 60은 1차 인간 T-세포에서 원형 RNA로부터의 사이토솔(도 **60A**) 및 표면(도 **60B**) 단백질에 대한 발광 발현 수준을 도시한 도면.

도 61은 IRES 특이적 원형 작제물로 처리되는 경우 인간 T-세포에서 발광 발현을 도시한 도면. 원형 RNA 작제물에서의 발현을 선형 mRNA와 비교하였다. **도 61A**, **도 61B** 및 **도 61G**는 다수의 공여자 세포에서 *가우시아* 루시퍼라제 발현을 제공한 도면. **도 61C**, **도 61D**, **도 61E** 및 **도 61F**는 다수의 공여자 세포에서 파이어플라이 루시퍼라제 발현을 제공한 도면.

도 62는 파이어플라이 루시퍼라제 발현 K562 세포를 항-CD19 또는 항-BCMA CAR을 암호화하는 원형 RNA를 포함하는 지질 나노입자로 처리한 후 인간 T-세포에서 항-CD19 CAR(도 **62A** 및 **도 62B**) 및 항-BCMA CAR(도 **62B**) 발현을 도시한 도면.

도 63은 항-CD19 CAR을 암호화하는 원형 RNA의 시험관내 전기천공을 통한 전달로부터 발생하는 특이적 항원-의존성 방식의 항-CD19 CAR 발현 수준을 도시한 도면. **도 63A**는 항-CD19 CAR로의 Nalm6 세포 용해를 도시한 도면. **도 63B**는 항-CD19 CAR로의 K562 세포 용해를 도시한 도면.

도 64는 LNP 및 녹색 형광 단백질(GFP)을 발현하는 원형 RNA를 함유하는 용액에서 ApoE3의 사용에 의해서 매개되는 LNP의 형질주입을 도시한 도면. **도 64A**는 생-사 결과를 도시한 도면. **도 64B, 도 61C, 도 61D, 및 도 64E**는 다수의 공여자에 대한 발현 빈도를 제공한 도면.

발명을 실시하기 위한 구체적인 내용

[0092] 본 발명은 특히 원형 RNA 요법에 기초하여 자가면역 장애 또는 암을 치료하기 위한 방법 및 조성물을 제공한다. 특히, 본 발명은 치료를 필요로 하는 대상체에게 2종의 치료용 단백질을 암호화하는 RNA를 포함하는 조성물을 자가면역 장애 또는 암의 적어도 하나의 증상 또는 특징이 강도, 중증도 또는 빈도가 감소되거나 발병이 지연되도록 하는 유효 용량 및 투여 간격으로 투여함으로써 자가면역 장애 또는 암을 치료하는 방법을 제공한다.

[0093] 특정 실시형태에서, 본 명세서에는 원형 RNA를 제조하기 위한 벡터가 제공되며, 이 벡터는 선택적인 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 선택적으로 제1 스페이서, 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, 제2 발현 서열, 선택적으로 제2 스페이서, 5' 그룹 I 인트론 단편 및 선택적인 3' 듀플렉스 형성 영역을 포함한다. 특정 실시형태에서, 본 명세서에는 원형 RNA를 제조하기 위한 벡터가 제공되며, 이 벡터는 선택적인 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 선택적으로 제1 스페이서, 제1 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 제2 IRES, 제2 발현 서열, 선택적으로 제2 스페이서, 5' 그룹 I 인트론 단편 및 선택적인 3' 듀플렉스 형성 영역을 포함한다. 일부 실시형태에서, 이들 요소는 상기 순서로 벡터에 위치된다. 일부 실시형태에서, 폴리뉴클레오타이드는 3' 듀플렉스 형성 영역 및 5' 듀플렉스 형성 영역을 함유한다. 일부 실시형태에서, 벡터는 3' 그룹 I 인트론 단편과 IRES 사이의 내부 5' 듀플렉스 형성 영역 및 발현 서열과 5' 그룹 I 인트론 단편 사이의 내부 3' 듀플렉스 형성 영역을 추가로 포함한다. 일부 실시형태에서, 내부 듀플렉스 형성 영역은 서로 간에 듀플렉스를 형성할 수 있지만, 외부 듀플렉스 형성 영역과는 형성할 수 없다. 일부 실시형태에서, 내부 듀플렉스 형성 영역은 제1 스페이서 및 제2 스페이서스페이서의 일부이다. 추가 실시형태는 본 명세서에 제공된 벡터를 사용하여 제조된 원형 RNA 폴리뉴클레오타이드를

포함하는 원형 RNA 폴리뉴클레오타이드, 이러한 원형 RNA를 포함하는 조성물, 이러한 원형 RNA를 포함하는 세포, 이러한 벡터, 원형 RNA, 조성물 및 세포의 사용 및 제조 방법을 포함한다.

[0094] 일부 실시형태에서, 본 명세서에는 유용한 단백질의 요법 또는 생산을 위해 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드를 세포 내로 투여하는 것을 포함하는 방법이 제공된다. 일부 실시형태에서, 방법은 리보뉴클레아제에 대한 원형 RNA의 내성으로 인해, 선형 RNA보다 긴 반감기를 갖는 진핵 세포 내부에서 목적하는 폴리펩타이드의 생산을 제공하는 데 유리하다.

[0095] 원형 RNA 폴리뉴클레오타이드는 엑소뉴클레아제-매개 분해에 필요한 유리 단부가 결여되어 있어서, RNA 분해의 몇몇 기전에 내성을 갖게 하고, 등가의 선형 RNA와 비교할 때 연장된 반감기를 부여한다. 원형화는 일반적으로 짧은 반감기로 고통받는 RNA 폴리뉴클레오타이드의 안정화를 허용할 수 있고, 다양한 응용에서 외인성 mRNA의 전반적인 효능을 개선시킬 수 있다. 일 실시형태에서, 단백질 합성에 의해서 평가된 바와 같은 진핵 세포(예를 들어, 인간 세포와 같은 포유동물 세포)에서 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드의 기능성 반감기는 적어도 20시간(예를 들어, 적어도 80시간)이다.

[0096] **1. 정의**

[0097] 본 명세서에서 사용된 바와 같이, 용어 "circRNA" 또는 "원형 폴리리보뉴클레오타이드" 또는 "원형 RNA"는 상호교환적으로 사용되며, 공유 결합을 통해 원형 구조를 형성하는 폴리리보뉴클레오타이드를 지칭한다.

[0098] 본 명세서에서 사용된 바와 같이, 용어 "3' 그룹 I 인트론 단편"은 스플라이스 부위 다이뉴클레오타이드 및 선택적으로 자연 엑손 서열의 스트레치를 포함하는 자연 그룹 I 인트론의 3'-근위 단부와 75% 이상의 유사성을 갖는 서열을 지칭한다.

[0099] 본 명세서에서 사용된 바와 같이, 용어 "5' 그룹 I 인트론 단편"은 스플라이스 부위 다이뉴클레오타이드 및 선택적으로 자연 엑손 서열의 스트레치를 포함하는 자연 그룹 I 인트론의 5'-근위 단부와 75% 이상의 유사성을 갖는 서열을 지칭한다.

[0100] 본 명세서에서 사용된 바와 같이, 용어 "순열 부위"는 절단이 인트론의 순열 이전에 만들어지는 그룹 I 인트론 내의 부위를 지칭한다. 이 절단은 원형화될 전구체 RNA 스트레치의 양측 상에 존재하도록 순열되는 3' 및 5' 그룹 I 인트론 단편을 생성한다.

[0101] 본 명세서에서 사용된 바와 같이, 용어 "스플라이스 부위"는 그룹 I 인트론에 부분적으로 또는 완전히 포함되고 RNA 원형화 동안 포스포다이에스터 결합이 절단되는 사이에 존재하는 다이뉴클레오타이드를 지칭한다.

[0102] 폴리뉴클레오타이드 작제물에서 발현 서열은 발현 서열에 의해 암호화된 폴리펩타이드가 일단 번역되면 세포에 의해 별도로 발현될 수 있게 하는 "절단 부위" 서열에 의해 분리될 수 있다.

[0103] "자가-절단 펩타이드"는 2개의 인접한 아미노산 사이에 펩타이드 결합 없이 번역되거나, 폴리펩타이드 및 자가-절단 펩타이드를 포함하는 폴리펩타이드가 생산될 때, 그것이 임의의 외부 절단 활성에 대한 필요성 없이 별개의 구분되는 제1 폴리펩타이드 및 제2 폴리펩타이드로 즉시 절단 또는 분리되도록 기능하는 펩타이드를 지칭한다.

[0104] 본 명세서에서 사용된 바와 같이, 용어 "치료용 단백질"은 번역된 핵산의 형태로 대상체에게 직접 또는 간접적으로 투여되는 경우 치료, 진단 및/또는 예방 효과를 갖고/갖거나 목적하는 생물학적 및/또는 약리학적 효과를 도출하는 임의의 단백질을 지칭한다.

[0105] αβ TCR의 α 및 β 쇄는 일반적으로 각각 2개의 도메인 및 영역, 즉, 가변 및 불변 도메인/영역을 갖는 것으로 간주된다. 가변 도메인은 가변 영역과 결합 영역의 연결부(concatenation)로 이루어진다. 따라서, 본 명세서 및 청구범위에서, 용어 "TCR 알파 가변 도메인"은 TRAV 영역과 TRAJ 영역의 연결부를 지칭하고, 용어 TCR 알파 불변 도메인은 세포외 TRAC 영역 또는 C-말단 절단된 TRAC 서열을 지칭한다. 마찬가지로, 용어 "TCR 베타 가변 도메인"은 TRBV 및 TRBD/TRBJ 영역의 연결부를 지칭하고, 용어 TCR 베타 불변 도메인은 세포외 TRBC 영역 또는 C-말단 절단된 TRBC 서열을 지칭한다.

[0106] 본 명세서에서 사용된 바와 같이, 용어 "면역원성"은 물질에 대한 면역 반응을 유도할 수 있는 가능성을 지칭한다. 면역 반응은 유기체의 면역계 또는 특정 유형의 면역 세포가 면역원성 물질에 노출될 때 유도될 수 있다. 용어 "비면역원성"은 물질에 대한 검출 가능한 역치를 초과하는 면역 반응의 결여 또는 부재를 지칭한다. 유기체의 면역계 또는 특정 유형의 면역 세포가 비-면역원성 물질에 노출될 때 어떠한 면역 반응도 검출되지

않는다. 일부 실시형태에서, 본 명세서에 제공된 바와 같은 비-면역원성 원형 폴리리보뉴클레오타이드는 면역원성 검정에 의해 측정될 때 미리 결정된 역치를 초과하는 면역 반응을 유도하지 않는다. 일부 실시형태에서, 유기체 또는 특정 유형의 면역 세포의 면역계가 본 명세서에 제공된 바와 같은 비-면역원성 원형 폴리리보뉴클레오타이드에 노출될 때 선천성 면역 반응이 검출되지 않는다. 일부 실시형태에서, 유기체 또는 특정 유형의 면역 세포의 면역계가 본 명세서에 제공된 바와 같은 비-면역원성 원형 폴리리보뉴클레오타이드에 노출될 때 적응성 면역 반응이 검출되지 않는다.

[0107] 본 명세서에서 사용된 바와 같이, 용어 "원형화 효율"은 선형 출발 물질과 비교할 때 생성된 원형 폴리리보뉴클레오타이드의 측정치를 지칭한다.

[0108] 본 명세서에서 사용된 바와 같이, 용어 "번역 효율"은 리보뉴클레오타이드 전사체로부터의 단백질 또는 펩타이드 생산의 속도 또는 양을 지칭한다. 일부 실시형태에서, 번역 효율은 단백질 또는 펩타이드를 암호화하는 전사체의 주어진 양당 생산된 단백질 또는 펩타이드의 양으로 표현될 수 있다.

[0109] 용어 "뉴클레오타이드"는 리보뉴클레오타이드, 데옥시리보뉴클레오타이드, 이들의 변형된 형태 또는 이들의 유사체를 지칭한다. 뉴클레오타이드는 퓨린, 예를 들어, 아데닌, 하이포잔틴, 구아닌, 및 이들의 유도체 및 유사체, 뿐만 아니라 피리미딘, 예를 들어, 사이토신, 우라실, 티민 및 이들의 유도체 및 유사체를 포함하는 종을 포함한다. 뉴클레오타이드 유사체는 5'-위치 피리미딘 변형, 8'-위치 퓨린 변형, 사이토신 엑소사이클릭 아민에서의 변형 및 5-브로모-우라실을 포함하지만 이들로 제한되지 않는 염기, 당 및/또는 인산염의 화학적 구조에 변형; 및 2'-OH가 H, OR, R, 할로, SH, SR, NH_2, NHR, NR_2 또는 CN(식 중, R은 본 명세서에 정의된 바와 같은 알킬 모이어티임)과 같은 기로 대체된 당-변형된 리보뉴클레오타이드를 포함하지만 이들로 제한되지 않는 2'-위치 당 변형을 갖는 뉴클레오타이드를 포함한다. 뉴클레오타이드 유사체는 또한 이노신, 퀴오신, 잔틴과 같은 염기를 갖는 뉴클레오타이드; 2'-메틸 리보스 등과 같은 당; 메틸포스포네이트, 포스포로티오에이트 및 펩타이드 연결부와 같은 비-자연 포스포다이에스터 연결부를 갖는 뉴클레오타이드를 포함하는 의미이다. 뉴클레오타이드 유사체는 5-메톡시우리딘, 1-메틸슈도우리딘 및 6-메틸아데노신을 포함한다.

[0110] 용어 "핵산" 및 "폴리뉴클레오타이드"는 임의의 길이, 예를 들어, 약 2개 초과의 염기, 약 10개 초과의 염기, 약 100개 초과의 염기, 약 500개 초과의 염기, 1000개 초과의 염기 또는 최대 약 10,000개 초과 염기의 중합체를 설명하기 위해 상호 교환적으로 사용되며, 이것은 뉴클레오타이드, 예를 들어, 데옥시리보뉴클레오타이드 또는 리보뉴클레오타이드로 구성되고 효소적으로 또는 합성적으로 생성될 수 있고(예를 들어, 미국 특허 제5,948,902호 및 이 문헌에 인용된 참고문헌에 기재된 바와 같음), 이것은 2개의 자연 발생 핵산과 유사한 서열 특이적 방식으로 자연 발생 핵산과 혼성화할 수 있고, 예를 들어, 왓슨-크릭 염기 짝지움 상호작용에 참여할 수 있다. 자연 발생 핵산은 구아닌, 사이토신, 아데닌, 티민 및 우라실(각각G, C, A, T 및 U)을 포함하는 뉴클레오타이드로 구성된다.

[0111] 본 명세서에서 사용된 바와 같은 용어 "리보핵산" 및 "RNA"는 리보뉴클레오타이드로 구성된 중합체를 의미한다.

[0112] 본 명세서에서 사용된 바와 같은 용어 "데옥시리보핵산" 및 "DNA"는 데옥시리보뉴클레오타이드로 구성된 중합체를 의미한다.

[0113] "단리된" 또는 "정제된"은 일반적으로 물질(예를 들어, 일부 실시형태에서 화합물, 폴리뉴클레오타이드, 단백질, 폴리펩타이드, 폴리뉴클레오타이드 조성물 또는 폴리펩타이드 조성물)이 그것이 존재하는 샘플의 상당한 백분율(예를 들어, 1% 초과, 2% 초과, 5% 초과, 10% 초과, 20% 초과, 50% 초과, 또는 일반적으로 최대 약 90% 내지 100%)을 차지하도록 하는 물질의 단리를 지칭한다. 특정 실시형태에서, 실질적으로 정제된 성분은 샘플의 적어도 50%, 80% 내지 85%, 또는 90% 내지 95%를 차지한다. 관심 폴리뉴클레오타이드 및 폴리펩타이드를 정제시키기 위한 기술은 당업계에 널리 공지되어 있으며, 예를 들어, 이온 교환 크로마토그래피, 친화성 크로마토그래피 및 밀도에 따른 침강을 포함한다. 일반적으로, 물질은 샘플의 다른 성분에 비해 자연적으로 발견되는 양보다 많은 양이 샘플에 존재할 때 정제된다.

[0114] 본 명세서에서 사용된 바와 같은 용어 "듀플렉스화", "이중 가닥" 또는 "혼성화"는 상보적 서열을 함유하는 핵산의 2개의 단일 가닥의 혼성화에 의해 형성된 핵산을 지칭한다. 대부분의 경우 게놈 DNA는 이중 가닥이다. 서열은 완전히 상보적이거나 부분적으로 상보적일 수 있다.

[0115] 본 명세서에서 사용된 바와 같이, RNA와 관련하여 "비구조화된"은 그 자체 또는 동일한 RNA 분자 내의 다른 서열과 함께 구조(예를 들어, 헤어핀 루프)를 형성하는 것으로 RNAFold 소프트웨어 또는 유사한 예측 도구에 의해 예측되지 않은 RNA 서열을 지칭한다. 일부 실시형태에서, 비구조화된 RNA는 뉴클레아제 보호 검정을 사용하여

기능적으로 특징규명될 수 있다.

[0116] 본 명세서에서 사용된 바와 같이, RNA와 관련하여 "구조화된"은 그 자체 또는 동일한 내의 다른 서열과 함께 구조(예를 들어, 헤어핀 루프)를 형성하는 것으로 RNAFold 소프트웨어 또는 유사한 예측 도구에 의해 예측된 RNA 서열을 지칭한다.

[0117] 본 명세서에서 사용된 바와 같이, 2개의 "듀플렉스 형성 영역", "상동성 아암" 또는 "상동성 영역"은 2개의 영역이 서로의 역상보체에 대해 충분한 수준의 서열 동일성을 공유하여 혼성화 반응에 대한 기질로서 작용하는 경우 서로 보완하거나 상보적이다. 본 명세서에서 사용된 바와 같이 폴리뉴클레오타이드 서열은 역상보체 또는 "상보적" 서열과 동일하거나 서열 동일성을 공유할 때 "상동성"을 갖는다. 듀플렉스 형성 영역과 대응 듀플렉스 형성 영역의 역상보체 사이의 백분율 서열 동일성은 혼성화가 일어나도록 하는 서열 동일성의 임의의 백분율일 수 있다. 일부 실시형태에서, 본 발명의 내부 듀플렉스 형성 영역은 또 다른 내부 듀플렉스 형성 영역과 듀플렉스를 형성할 수 있고, 외부 듀플렉스 형성 영역과 듀플렉스를 형성하지 않는다.

[0118] 선형 핵산 분자는 핵산 포스포다이에스터 연결부가 치환체 모노뉴클레오타이드의 당 모이어티의 5' 탄소 및 3' 탄소에서 발생하기 때문에 "5'-말단"(5' 단부) 및 "3'-말단"(3' 단부)을 갖는다고 한다. 새로운 연결부가 5' 탄소에 있을 폴리뉴클레오타이드의 단부 뉴클레오타이드는 이의 5' 말단 뉴클레오타이드이다. 새로운 연결부가 3' 탄소에 있을 폴리뉴클레오타이드의 단부 뉴클레오타이드는 이의 3' 말단 뉴클레오타이드이다. 본 명세서에서 사용된 바와 같은 말단 뉴클레오타이드는 3'- 또는 5'-말단의 단부 위치에 존재하는 뉴클레오타이드이다.

[0119] "전사"는 DNA 분자를 주형으로 사용하는 RNA 중합효소에 의한 RNA 분자의 형성 또는 합성을 의미한다. 본 발명은 전사에 사용되는 RNA 중합효소에 대해 제한되지 않는다. 예를 들어, 일부 실시형태에서, T7-타입 RNA 중합효소가 사용될 수 있다.

[0120] "번역"은 RNA 주형에 기초한 리보솜에 의한 폴리펩타이드 분자의 형성을 의미한다.

[0121] 본 명세서에서 사용된 용어는 단지 특정 실시형태를 설명하기 위한 것이며, 제한하려는 의도가 아님을 이해해야 한다. 본 명세서 및 첨부된 특허청구범위에 사용된 바와 같이, 단수 형태는 내용이 명백하게 달리 지시하지 않는 한 복수의 지시 대상을 포함한다. 따라서, 예를 들어, "세포"에 대한 언급은 2개 이상의 세포의 조합, 또는 세포의 전체 배양물을 포함하고; "폴리뉴클레오타이드"에 대한 언급은 실질적으로 그 폴리뉴클레오타이드의 다수의 카피를 포함한다. 구체적으로 언급되거나 문맥상 명백하지 않는 한, 본 명세서에서 사용된 용어 "또는"은 포괄적인 것으로 이해된다. 본 명세서의 상기에서 본 명세서 및 하기에 정의되지 않는 한, 본 명세서에 사용된 모든 기술 및 과학 용어는 당업자가 일반적으로 이해하는 것과 동일한 의미를 갖는다.

[0122] 구체적으로 언급되지 않거나 문맥상 명백하지 않은 한, 본 명세서에서 사용되는 용어 "약"은 당업계에서 통상적인 허용 오차 범위 내, 예를 들어, 평균의 2 표준 편차 내로 이해된다. "약"은 제시된 값의 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 0.1%, 0.09%, 0.08%, 0.07%, 0.06%, 0.05%, 0.04%, 0.03%, 0.02% 또는 0.01% 이내인 것으로 이해될 수 있다. 문맥상 달리 명백하지 않는 한, 본 명세서에 제공된 모든 수치는 "약"이라는 용어로 수식된다.

[0123] 본 명세서에 사용된 용어 "암호화하다"는 중합체 매크로분자(polymeric macromolecule)의 정보가 첫 번째 분자와 상이한 두 번째 분자의 생산을 지시하는 데 사용되는 임의의 과정을 광범위하게 지칭한다. 두 번째 분자는 첫 번째 분자의 화학적 성질과 상이한 화학 구조를 가질 수 있다.

[0124] "공동 투여"는 본 명세서에 제공된 치료제가 하나 이상의 추가 치료제의 효과를 향상시킬 수 있도록 또는 그 역이도록 시간이 충분히 근접한 하나 이상의 추가 치료제와 함께 본 명세서에 제공된 치료제를 투여하는 것을 의미한다.

[0125] 본 명세서에 사용된 용어 "치료하다" 및 "예방하다" 및 이로부터 파생된 단어는 반드시 100% 또는 완전한 치료 또는 예방을 의미하는 것은 아니다. 오히려, 당업자가 잠재적인 이점 또는 치료 효과를 갖는 것으로 인식하는 다양한 정도의 치료 또는 예방이 존재한다. 본 명세서에 개시된 방법에 의해 제공되는 치료 또는 예방은 질환의 하나 이상의 상태 또는 증상의 치료 또는 예방을 포함할 수 있다. 또한, 본 명세서의 목적을 위해, "예방"은 질환 또는 이의 증상 또는 병태의 발병을 지연시키는 것을 포함할 수 있다.

[0126] 본 명세서에서 사용되는 바와 같이, "자가면역"은 포유동물 및 인간 내에 침입하여 지속하는 박테리아, 바이러스, 진균 또는 기생 유기체로부터의 감염성 비-자가-항원과 구별되는, 비-감염성 자가-항원에 대한 지속적이고 점진적인 면역 반응으로 정의된다. 자가면역 병태는 경피증, 그레이브스병, 크론병, 쇼그렌병, 다발성 경화증,

하시모토병, 건선, 중증 근무력증, 자가면역 다발성 내분비병증 증후군, 제I형 당뇨병(TIDM), 자가면역 위염, 자가면역 포도막염, 다발성근염, 대장염 및 갑상선염, 뿐만 아니라 인간 루푸스로 대표되는 전신성 자가면역 질환을 포함한다. 본 명세서에서 사용되는 바와 같은 "자가항원" 또는 "자가-항원"은 포유동물에서 천연이고, 상기 포유동물에서 면역원성인 항원 또는 에피토프를 지칭한다.

[0127] 본 명세서에 사용된 용어 "발현 서열"은 산물, 예를 들어, 펩타이드 또는 폴리펩타이드, 조절 핵산, 또는 비-암호 핵산을 암호화하는 핵산 서열을 지칭할 수 있다. 펩타이드 또는 폴리펩타이드를 암호화하는 예시적인 발현 서열은 복수의 뉴클레오타이드 트라이어드를 포함할 수 있으며, 이들 각각은 아미노산을 암호화할 수 있고 "코돈"으로 지칭된다.

[0128] 본 명세서에서 사용된 바와 같이, "스페이서"는 폴리뉴클레오타이드 서열을 따라 2개의 다른 요소를 분리하는 1개 뉴클레오타이드 내지 수백 또는 수천 개의 뉴클레오타이드 범위의 폴리뉴클레오타이드 서열 영역을 지칭한다. 서열은 정의되거나 무작위일 수 있다. 스페이서는 일반적으로 비-암호화된다. 일부 실시형태에서, 스페이서는 듀플렉스 형성 영역을 포함한다.

[0129] 본 명세서에서 사용된 바와 같이, "내부 리보솜 유입 부위" 또는 "IRES"는 10nt 내지 1000nt 또는 그 초과의 크기 범위의 RNA 서열 또는 구조 요소이며, 이것은 전형적인 RNA 캡 구조의 부재 하에서의 폴리펩타이드의 번역을 개시할 수 있다. IRES는 전형적으로 약 500nt 내지 약 700nt 길이이다.

[0130] 본 명세서에서 사용된 바와 같이, "miRNA 부위"는 자연 miRNA 서열의 적어도 8개의 뉴클레오타이드와 듀플렉스를 형성할 수 있는 폴리뉴클레오타이드 내의 뉴클레오타이드의 스트레치를 지칭한다.

[0131] 본 명세서에서 사용된 바와 같이, "엔도뉴클레아제 부위"는 엔도뉴클레아제 단백질에 의해 인식 및 절단될 수 있는 폴리뉴클레오타이드 내의 뉴클레오타이드의 스트레치를 지칭한다.

[0132] 본 명세서에서 사용된 바와 같이, "바이시스트론성(bicistronic) RNA"는 2개의 구별되는 단백질을 암호화하는 2개의 발현 서열을 포함하는 폴리뉴클레오타이드를 지칭한다. 이들 발현 서열은 종종 2A 부위 또는 IRES 서열과 같은 절단 가능한 펩타이드에 의해 분리된다. 이들은 또한 리보솜 스키핑 요소 또는 프로테아제 절단에 의해서 분리될 수 있다.

[0133] 본 명세서에서 사용되는 바와 같이, 용어 "리보솜 스키핑 요소"는 하나의 RNA 분자의 번역으로부터 2개의 펩타이드 쇄의 생성을 초래할 수 있는 짧은 펩타이드 서열을 암호화하는 뉴클레오타이드 서열을 지칭한다. 이론에 얽매이고자 함은 아니지만, 리보솜 스키핑 요소는 다음에 의해서 기능한다고 가정된다: (1) 제1 펩타이드 쇄의 말단 번역 및 제2 펩타이드 쇄의 재-초기 번역; 또는 (2) 암호화된 펩타이드의 내인성 프로테아제 활성에 의한 또는 환경(예를 들어, 사이토솔) 내의 또 다른 프로테아제에 의한 리보솜 스크핑 요소에 의해서 암호화된 펩타이드 서열 내의 펩타이드 결합의 절단.

[0134] 본 명세서에서 사용된 바와 같이, 용어 "공동제형화하다"는 2개 이상의 핵산 또는 핵산 및 다른 활성 약물 물질을 포함하는 나노입자 제형을 지칭한다. 전형적으로, 비율은 등몰이거나 둘 이상의 핵산 또는 핵산 및 다른 활성 약물 물질의 비율적 양으로 정의된다.

[0135] 본 명세서에서 사용된 바와 같이, "전달 비히클"은 일반적으로 핵산을 포함하는 생물학적 활성제의 투여와 관련하여 사용하도록 의도된 임의의 표준 약제학적 담체, 희석제, 부형제 등을 포함한다.

[0136] 본 명세서에서 사용된 바와 같이, "지질 나노입자"라는 어구는 하나 이상의 지질(예를 들어, 일부 실시형태에서 양이온성 지질, 비-양이온성 지질 및 PEG-변형된 지질)을 포함하는 전달 비히클을 지칭한다.

[0137] 본 명세서에서 사용된 바와 같이, "양이온성 지질"이라는 어구는 생리 pH와 같은 선택된 pH에서 순 양전하를 갖는 다수의 지질 종 중 임의의 것을 지칭한다.

[0138] 본 명세서에서 사용된 바와 같이, "비-양이온성 지질"이라는 어구는 임의의 중성, 양쪽이온성 또는 음이온성 지질을 지칭한다.

[0139] 본 명세서에서 사용된 바와 같이, "음이온성 지질"이라는 어구는 생리 pH와 같은 선택된 pH에서 순 음전하를 갖는 다수의 지질 종 중 임의의 것을 지칭한다.

[0140] 본 명세서에서 사용된 바와 같이, "이온화 가능한 지질"이라는 문구는 생리 pH 4와 같은 선택된 pH에서 순 양전하를 갖고 생리 pH 7과 같은 다른 pH에서 중성 전하를 갖는 다수의 지질 종 중 임의의 것을 지칭한다.

[0141] 일부 실시형태에서, 본 명세서에 개시된 지질, 예를 들어, 이온화 가능한 지질은 하나 이상의 절단 가능한 기를 포함한다. 용어 "절단하다" 및 "절단 가능한"은 본 명세서에서 대상 작용기 내의 또는 이에 인접한 원자 사이의 하나 이상의 화학 결합(예를 들어, 공유 결합, 수소 결합, 반데르발스힘 및/또는 이온성 상호작용 중 하나 이상)이 파괴되거나(예를 들어, 가수분해되거나), 선택된 조건에 노출될 때(예를 들어, 효소 조건에 노출될 때) 파괴될 수 있음을 의미하는 데 사용된다. 특정 실시형태에서, 절단 가능한 기는 이황화 작용기이고, 특정 실시형태에서 선택된 생물학적 조건(예를 들어, 세포내 조건)에 노출될 때 절단될 수 있는 이황화기이다. 특정 실시형태에서, 절단 가능한 기는 선택된 생물학적 조건에 노출될 때 절단될 수 있는 에스터 작용기이다. 예를 들어, 이황화기는 효소에 의해 또는 가수분해, 산화 또는 환원 반응에 의해 절단될 수 있다. 이러한 이황화 작용기의 절단 시, 이에 결합된 하나 이상의 작용기 모이어티 또는 기(예를 들어, 헤드기 및/또는 테일기 중 하나 이상)가 해방될 수 있다. 예시적인 절단 가능한 기는 이황화기, 에스터기, 에터기 및 이들의 임의의 유도체(예를 들어, 알킬 및 아릴 에스터)를 포함할 수 있지만 이들로 제한되지 않는다. 특정 실시형태에서, 절단 가능한 기는 에스터기 또는 에터기가 아니다. 일부 실시형태에서, 절단 가능한 기는 하나 이상의 작용성 모이어티 또는 기(예를 들어, 적어도 하나의 헤드기 및 적어도 하나의 테일기)에 결합된다(예를 들어, 수소 결합, 반 데르 발스힘, 이온성 상호작용 및 공유 결합 중 하나 이상에 의해 결합된다). 특정 실시형태에서, 작용성 모이어티 또는 기 중 적어도 하나는 친수성이다(예를 들어, 이미다졸, 구아니디늄, 아미노, 이민, 엔아민, 선택적으로-치환된 알킬 아미노 및 피리딜 중 하나 이상을 포함하는 친수성 헤드기).

[0142] 본 명세서에서 사용된 바와 같이, 용어 "친수성"은 작용기가 물-선호적이며, 전형적으로 이러한 기가 수용성임을 정성적 용어로 나타내기 위해 사용된다. 예를 들어, 본 명세서에는 하나 이상의 친수성 기(예를 들어, 친수성 헤드기)에 결합된 절단 가능한 이황화(S-S) 작용기를 포함하는 화합물이 개시되며, 여기서 이러한 친수성 기는 이미다졸, 구아니디늄, 아미노, 이민, 엔아민, 선택적으로 치환된 알킬 아미노(예를 들어, 알킬 아미노, 예컨대, 다이메틸아미노) 및 피리딜을 포함하거나 이들로 이루어진 군으로부터 선택된다.

[0143] 특정 실시형태에서, 본 명세서에 개시된 화합물을 포함하는 모이어티의 작용기 중 적어도 하나는 자연에서 소수성이다(예를 들어, 콜레스테롤과 같은 자연 발생 지질을 포함하는 소수성 테일기). 본 명세서에서 사용된 바와 같이, 용어 "소수성"은 작용기가 물-회피성이며, 전형적으로 이러한 기가 수용성이 아님을 정성적 용어로 나타내기 위해 사용된다. 예를 들어, 본 명세서에는 하나 이상의 소수성 기에 결합된 절단 가능한 작용기(예를 들어, 이황화(S-S) 기)를 포함하는 화합물이 개시되며, 여기서 이러한 소수성 기는 콜레스테롤과 같은 하나 이상의 자연 발생 지질 및/또는 선택적으로 치환된, 가변적으로 포화 또는 불포화된 C_6-C_{20} 알킬 및/또는 선택적으로 치환된, 가변적으로 포화 또는 불포화된 C_6-C_{20} 아실을 포함한다.

[0144] 본 명세서에 기재된 화합물은 하나 이상의 동위원소 치환을 또한 포함할 수 있다. 예를 들어, H는 ^{1}H, ^{2}H(D 또는 중수소) 및 ^{3}H(T 또는 삼중수소)를 비롯한 임의의 동위원소 형태로 존재할 수 있고; C는 ^{12}C, ^{13}C 및 ^{14}C를 비롯한 임의의 동위원소 형태로 존재할 수 있고; O는 ^{16}O 및 ^{18}O를 비롯한 임의의 동위원소 형태로 존재할 수 있고; F는 ^{18}F a및 ^{19}F를 비롯한 임의의 동위원소 형태로 존재할 수 있고; 그 등등이다.

[0145] 화합물 및 이의 약제학적으로 허용 가능한 염, 이러한 화합물을 함유하는 약제학적 조성물 및 이러한 화합물 및 조성물의 사용 방법을 포함할 수 있는 본 발명을 기재하는 경우, 하기 용어는 존재하는 경우 달리 제시되지 않는 한 다음 의미를 갖는다. 또한, 본 명세서에 기술될 때 하기에 정의된 임의의 모이어티는 다양한 치환체로 치환될 수 있으며, 각각의 정의는 하기에 제시된 범주 내에서 그러한 치환된 모이어티를 포함하도록 의도된다는 것을 또한 이해해야 한다. 달리 명시되지 않는 한, 용어 "치환된"은 다음과 같이 정의된다. 용어 "기" 및 "라디칼"은 본 명세서에서 사용되는 경우 상호 교환 가능한 것으로 간주될 수 있음을 추가로 이해해야 한다.

[0146] 값의 범위가 열거되는 경우, 그 범위 내의 각각의 값 및 하위 범위를 포함하도록 의도된다. 예를 들어, "C_{1-6} 알킬"은 C_1, C_2, C_3, C_4, C_5, C_6, C_{1-6}, C_{1-5}, C_{1-4}, C_{1-3}, C_{1-2}, C_{2-6}, C_{2-5}, C_{2-4}, C_{2-3}, C_{3-6}, C_{3-5}, C_{3-4}, C_{4-6}, C_{4-5} 및 C_{5-6} 알킬을 포함하도록 의도된다.

[0147] 특정 실시형태에서, 본 명세서에 개시된 화합물은 예를 들어, 각각이 적어도 하나의 절단 가능한 기에 결합되어 이러한 화합물을 양친매성으로 만드는 적어도 하나의 친수성 헤드기 및 적어도 하나의 소수성 테일기를 포함한다. 화합물 또는 조성물을 설명하기 위해 본 명세서에서 사용된 바와 같이, 용어 "양친매성"은 극성(예를 들어, 물) 및 비극성(예를 들어, 지질) 환경 모두에서 용해되는 능력을 의미한다. 예를 들어, 특정 실시형태에서, 본 명세서에 개시된 화합물은 각각 절단 가능한 기(예를 들어, 이황화)에 결합된 적어도 하나의 친유성 테일기(예

를 들어, 콜레스테롤 또는 C_6-C_{20} 알킬) 및 적어도 하나의 친수성 헤드기(예를 들어, 이미다졸)를 포함한다.

[0148] 사용되는 바와 같은 용어 "헤드기" 및 "테일기"는 본 발명의 화합물을 기술하고, 특히 이러한 화합물을 포함하는 특정 작용기는 다른 작용기에 비해 하나 이상의 작용기의 배향을 기재하기 위해서 참조의 용이성을 위해서 사용된다는 것을 주목해야 한다. 예를 들어, 특정 실시형태에서 친수성 헤드기(예를 들어, 구아니디늄)는 절단 가능한 작용기(예를 들어, 이황화기)에 (수소 결합, 반 데르 발스 힘, 이온성 상호작용 및 공유 결합 중 하나 이상에 의해서) 결합되고, 그 다음 소수성 테일기(예를 들어, 콜레스테롤)에 결합된다.

[0149] 본 명세서에서 사용된 바와 같이, 용어 "알킬"은 직쇄 및 분지쇄 C_1-C_{40} 탄화수소(예를 들어, C_6-C_{20} 탄화수소) 둘 다를 지칭하고, 포화 및 불포화 탄화수소를 모두 포함한다. 특정 실시형태에서, 알킬은 하나 이상의 환식 알킬 및/또는 하나 이상의 헤테로원자, 예컨대, 산소, 질소 또는 황을 포함할 수 있고 선택적으로 치환체(예를 들어, 알킬, 할로, 알콕실, 하이드록시, 아미노, 아릴, 에터, 에스터 또는 아마이드 중 하나 이상)로 치환될 수 있다. 특정 실시형태에서, 고려되는 알킬은 (9Z,12Z)-옥타데카-9,12-다이엔을 포함한다. 예를 들어, "C_6-C_{20}"과 같은 명칭의 사용은 언급된 범위의 탄소 원자를 갖는 알킬(예를 들어, 직쇄형 또는 분지쇄형 및 알켄 및 알킬 포함)을 지칭하도록 의도된다. 일부 실시형태에서, 알킬기는 1 내지 10개의 탄소 원자를 갖는다("C_{1-10} 알킬"). 일부 실시형태에서, 알킬기는 1 내지 9개의 탄소 원자를 갖는다("C_{1-9} 알킬"). 일부 실시형태에서, 알킬기는 1 내지 8개의 탄소 원자를 갖는다("C_{1-8} 알킬"). 일부 실시형태에서, 알킬기는 1 내지 7개의 탄소 원자를 갖는다("C_{1-7} 알킬"). 일부 실시형태에서, 알킬기는 1 내지 6개의 탄소 원자를 갖는다("C_{1-6} 알킬"). 일부 실시형태에서, 알킬기는 1 내지 5개의 탄소 원자를 갖는다("C_{1-5} 알킬"). 일부 실시형태에서, 알킬기는 1 내지 4개의 탄소 원자를 갖는다("C_{1-4} 알킬"). 일부 실시형태에서, 알킬기는 1 내지 3개의 탄소 원자를 갖는다("C_{1-3} 알킬"). 일부 실시형태에서, 알킬기는 1 내지 2개의 탄소 원자를 갖는다("C_{1-2} 알킬"). 일부 실시형태에서, 알킬기는 1개의 탄소 원자를 갖는다("C_1 알킬"). C_{1-6} 알킬기의 예는 메틸, 에틸, 프로필, 아이소프로필, 부틸, 아이소부틸, 펜틸, 헥실 등을 포함한다.

[0150] 본 명세서에서 사용되는 바와 같이, "알켄일"은 2 내지 20개의 탄소 원자, 하나 이상의 탄소-탄소 이중 결합(예를 들어, 1, 2, 3 또는 4개의 탄소-탄소 이중 결합), 및 선택적으로 하나 이상의 탄소-탄소 삼중 결합(예를 들어, 1, 2, 3 또는 4개의 탄소-탄소 삼중 결합)을 갖는 직쇄형 또는 분지형 탄화수소기의 라디칼을 지칭한다("C_{2-20} 알켄일"). 특정 실시형태에서, 알켄일은 어떠한 삼중 결합도 함유하지 않는다. 일부 실시형태에서, 알켄일기는 2 내지 10개의 탄소 원자를 갖는다("C_{2-10} 알켄일"). 일부 실시형태에서, 알켄일기는 2 내지 9개의 탄소 원자를 갖는다("C_{2-9} 알켄일"). 일부 실시형태에서, 알켄일기는 2 내지 8개의 탄소 원자를 갖는다("C_{2-8} 알켄일"). 일부 실시형태에서, 알켄일기는 2 내지 7개의 탄소 원자를 갖는다("C_{2-7} 알켄일"). 일부 실시형태에서, 알켄일기는 2 내지 6개의 탄소 원자를 갖는다("C_{2-6} 알켄일"). 일부 실시형태에서, 알켄일기는 2 내지 5개의 탄소 원자를 갖는다("C_{2-5} 알켄일"). 일부 실시형태에서, 알켄일기는 2 내지 4개의 탄소 원자를 갖는다("C_{2-4} 알켄일"). 일부 실시형태에서, 알켄일기는 2 내지 3개의 탄소 원자를 갖는다("C_{2-3} 알켄일"). 일부 실시형태에서, 알켄일기는 2개의 탄소 원자를 갖는다("C_2 알켄일"). 하나 이상의 탄소-탄소 이중 결합이 내부(예컨대, 2-부텐일에서) 또는 말단(예컨대, 1-부텐일에서)에 존재할 수 있다. C_{2-4} 알켄일기의 예는 에텐일(C_2), 1-프로펜일(C_3), 2-프로펜일(C_3), 1-부텐일(C_4), 2-부텐일(C_4), 부타다이엔일(C_4) 등을 포함한다. C_{2-6} 알켄일기의 예는 상기에 언급된 C_{2-4} 알켄일기뿐만 아니라 펜틸 (C_5), 펜타다이엔일(C_5), 헥센일(C_6) 등을 포함한다. 알켄일의 추가 예는 헵텐일(C_7), 옥텐일(C_8), 옥타트라이엔일(C_8) 등을 포함한다.

[0151] 본 명세서에서 사용되는 바와 같이, "알킨일"은 2 내지 20개의 탄소 원자, 하나 이상의 탄소-탄소 삼중 결합(예를 들어, 1, 2, 3 또는 4개의 탄소-탄소 삼중 결합), 및 선택적으로 하나 이상의 탄소-탄소 이중 결합(예를 들어, 1, 2, 3 또는 4개의 탄소-탄소 이중 결합)을 갖는 직쇄형 또는 분지형 탄화수소기의 라디칼을 지칭한다("C_{2-20} 알킨일"). 특정 실시형태에서, 알킨일은 어떠한 이중 결합도 함유하지 않는다. 일부 실시형태에서, 알킨일기는 2 내지 10개의 탄소 원자를 갖는다("C_{2-10} 알킨일"). 일부 실시형태에서, 알킨일기는 2 내지 9개의 탄소 원자를 갖는다("C_{2-9} 알킨일"). 일부 실시형태에서, 알킨일기는 2 내지 8개의 탄소 원자를 갖는다("C_{2-8} 알킨일"). 일

부 실시형태에서, 알킨일기는 2 내지 7개의 탄소 원자를 갖는다("C_{2-7} 알킨일"). 일부 실시형태에서, 알킨일기는 2 내지 6개의 탄소 원자를 갖는다("C_{2-6} 알킨일"). 일부 실시형태에서, 알킨일기는 2 내지 5개의 탄소 원자를 갖는다("C_{2-5} 알킨일"). 일부 실시형태에서, 알킨일기는 2 내지 4개의 탄소 원자를 갖는다("C_{2-4} 알킨일"). 일부 실시형태에서, 알킨일기는 2 내지 3개의 탄소 원자를 갖는다("C_{2-3} 알킨일"). 일부 실시형태에서, 알킨일기는 2개의 탄소 원자를 갖는다("C_2 알킨일"). 하나 이상의 탄소-탄소 삼중 결합이 내부(예컨대, 2-부틴일에서) 또는 말단(예컨대, 1-부틴일에서)에 존재할 수 있다. C_{2-4} 알킨일기의 예는 비제한적으로 에틴일(C_2), 1-프로핀일(C_3), 2-프로핀일(C_3), 1-부틴일(C_4), 2-부틴일(C_4) 등을 포함한다. C_{2-6} 알켄일기의 예는 상기에 언급된 C_{2-4} 알킨일기뿐만 아니라 펜틴일(C_5), 헥신일(C_6) 등을 포함한다. 알킨일의 추가 예는 헵틴일(C_7), 옥틴일(C_8) 등을 포함한다.

[0152] 본 명세서에서 사용되는 바와 같이, "알킬렌", "알켄일렌" 및 "알킨일렌"은 각각 알킬, 알켄일 및 알킨일기의 2가 라디칼을 지칭한다. 탄소의 범위 또는 수가 특정 "알킬렌", "알켄일렌" 또는 "알킨일렌"기에 대해서 제공되는 경우, 그 범위 또는 수는 선형 탄소 2가 쇄 내의 탄소의 범위 또는 수를 지칭한다. "알킬렌", "알켄일렌" 및 "알킨일렌"기는 본 명세서에 기재된 바와 같은 하나 이상의 치환체로 치환되거나 치환되지 않을 수 있다.

[0153] 본 명세서에서 사용된 바와 같이, 용어 "아릴"은 고리 부분에 6 내지 10개의 탄소를 함유하는 방향족 기(예를 들어, 단환식, 이환식 및 삼환식 구조)를 지칭한다. 아릴기는 이용 가능한 탄소 원자를 통해 선택적으로 치환될 수 있고, 특정 실시형태에서 하나 이상의 헤테로원자, 예컨대, 산소, 질소 또는 황을 포함할 수 있다. 일부 실시형태에서, 아릴기는 6개의 고리 탄소 원자("C_6 아릴"; 예를 들어, 페닐)를 갖는다. 일부 실시형태에서, 아릴기는 10개의 탄소 원자("C_{10} 아릴"; 예를 들어, 나프틸 예컨대, 1-나프틸 및 2-나프틸)를 갖는다.

[0154] 본 명세서에서 사용되는 바와 같이, "헤테로아릴"은 방향족 고리 시스템에 제공된 고리 탄소 원자 및 1 내지 4개의 고리 헤테로원자를 갖는 5 내지 10원의 단환식 또는 이환식 4n+2 방향족 고리 시스템(예를 들어, 환식 배열에 공유된 6 또는 10개의 전자를 가짐)의 라디칼을 지칭하고, 여기서 각각의 헤테로원자는 질소, 산소 및 황으로부터 독립적으로 선택된다("5 내지 10원 헤테로아릴"). 하나 이상의 질소 원자를 함유하는 헤테로아릴기에서, 부착점은 원자가가 허용하는 한 탄소 또는 질소 원자일 수 있다. 헤테로아릴 이환식 고리 시스템은 하나 또는 양쪽 고리에 하나 이상의 헤테로원자를 포함할 수 있다. "헤테로아릴"은 상기에 정의된 바와 같은 헤테로아릴 고리가 하나 이상의 카보사이클릴 또는 헤테로사이클릴기와 융합되고, 부착점이 헤테로아릴 고리 상에 존재하는 고리 시스템을 포함하고, 이러한 예에서 고리 구성원의 수는 헤테로아릴 고리 시스템 내의 고리 구성원의 수를 계속 지정한다. "헤테로아릴"은 또한 상기에 정의된 바와 같은 헤테로아릴 고리가 하나 이상의 아릴기와 융합된 고리 시스템을 포함하되, 부착점은 아릴 또는 헤테로아릴 고리 상에 존재하고, 이러한 예에서 고리 구성원의 수는 융합된(아릴/헤테로아릴) 고리 시스템 내의 고리 구성원의 수를 지정한다. 하나의 고리가 헤테로원자를 함유하지 않는 이환식 헤테로아릴기(예를 들어, 인돌릴, 퀴놀리닐, 카바졸릴 등)에서 부착점은 고리, 즉 헤테로원자를 갖는 고리(예를 들어, 2-인돌릴) 또는 헤테로원자를 포함하지 않는 고리(예를 들어, 5-인돌릴) 상에 존재할 수 있다.

[0155] 용어 "사이클로알킬"은 본 명세서에 언급된 3 내지 12, 3 내지 8, 4 내지 8, 또는 4 내지 6개 탄소의 1가 포화 환식, 이환식 또는 가교된 환식(예를 들어, 아다만틸) 탄화수소 기, 예를 들어, 사이클로알칸으로부터 유래된 "C_{4-8}사이클로알킬"을 지칭한다. 예시적인 사이클로알킬기는 사이클로헥산, 사이클로펜탄, 사이클로부탄 및 시클로프로판을 포함하지만 이들로 제한되지 않는다.

[0156] 본 명세서에서 사용되는 바와 같이, "헤테로사이클릴" 또는 "헤테로사이클릭"은 고리 탄소 원자 및 1 내지 4개의 고리 헤테로원자를 갖는 3- 내지 10-원 비-방향족 고리 시스템의 라디칼을 지칭하고, 여기서 각각의 헤테로원자는 질소, 산소, 황, 붕소, 인 및 규소("3 내지 10원 헤테로사이클릴")로부터 독립적으로 선택된다. 하나 이상의 질소 원자를 함유하는 헤테로사이클릴기에서, 부착점은 원자가가 허용하는 한 탄소 또는 질소 원자일 수 있다. 헤테로사이클릴기는 단환식("단환식 헤테로사이클릴") 또는 융합, 가교 또는 스피로 고리 시스템, 예컨대, 이환식 시스템("이환식 헤테로사이클릴")일 수 있고, 포화될 수 있거나, 부분적으로 불포화될 수 있다. 헤테로사이클릴 이환식 고리 시스템은 하나 또는 양쪽 고리에 하나 이상의 헤테로원자를 포함할 수 있다. "헤테로사이클릴"은 또한 상기에 정의된 바와 같은 헤테로사이클릴 고리가 하나 이상의 카보사이클릴기와 융합되고, 부착점이 카보사이클릴 또는 헤테로사이클릴 고리 상에 존재하는 고리 시스템, 또는 상기에 정의된 바와 같은 헤테로사이클릴 고리가 하나 이상의 아릴 또는 헤테로아릴기와 융합되고, 부착점이 헤테로사이클릴 고리 상에

존재하는 고리 시스템을 포함하고, 이러한 예에서 고리 구성원의 수는 헤테로사이클릴 고리 시스템 내의 고리 구성원의 수를 계속 지정한다. 용어 "헤테로사이클", "헤테로사이클릴", "헤테로사이클릴 고리", "헤테로사이클릭기", "헤테로사이클릭 모이어티" 및 "헤테로사이클릭 라디칼"은 상호 교환 가능하게 사용될 수 있다.

[0157] 본 명세서에서 사용되는 바와 같이, "사이아노"는 -CN을 지칭한다.

[0158] 본 명세서에서 사용되는 바와 같은 용어 "할로" 및 "할로겐"은 플루오린(플루오로, F), 염소(클로로, Cl), 브로민(브로모, Br) 및 아이오딘(아이오도, I)로부터 선택된 원자를 지칭한다. 특정 실시형태에서, 할로기는 플루오로 또는 클로로이다.

[0159] 본 명세서에서 사용되는 바와 같은 용어 "알콕시"는 산소 원자(-O(알킬))를 통해 다른 부분에 부착된 알킬기를 지칭한다. 비제한적인 예는 예를 들어, 메톡시, 에톡시, 프로폭시 및 부톡시를 포함한다.

[0160] 본 명세서에서 사용되는 바와 같이, "옥소"는 -C=O를 지칭한다.

[0161] 일반적으로, "선택적으로"라는 용어가 앞에 있든 없든, 용어 "치환된"은 기(예를 들어, 탄소 또는 질소 원자) 상에 존재하는 적어도 하나의 수소가 허용 가능한 치환체, 예를 들어, 치환 시 안정적인 화합물, 예를 들어, 변형, 예컨대, 재배열, 고리화, 제거 또는 다른 반응을 자발적으로 겪지 않는 화합물을 생성하는 치환체로 대체된다는 것을 의미한다. 달리 제시되지 않는 한, "치환된" 기는 기의 하나 이상의 치환 가능한 위치에 치환체를 가지며, 임의의 주어진 구조에서 하나 초과의 위치가 치환되는 경우, 치환체는 각각의 위치에서 동일하거나 상이하다.

[0162] 본 명세서에서 사용되는 바와 같이, "약제학적으로 허용 가능한 염"은 건전한 의학적 판단의 범위 내에서 과도한 독성, 자극, 알레르기 반응 등이 없이 인간 및 하등 동물의 조직과 접촉하여 사용하기에 적합하고, 합리적인 유익/유해 비율에 상응하는 염을 지칭한다. 약제학적으로 허용 가능한 염은 당업계에 널리 공지되어 있다. 예를 들어, Berge 등은 문헌[*J. Pharmaceutical Sciences* (1977) 66:1-19]에서 약제학적으로 허용 가능한 염을 자세히 기재한다. 본 발명의 화합물의 약제학적으로 허용 가능한 염은 적합한 무기 및 유기 산 및 염기로부터 유래된 것을 포함한다. 약학적으로 허용 가능한 무독성 산 부가염의 예는 무기산, 예컨대, 염산, 브로민화수소산, 인산, 황산 및 과염소산 또는 유기산, 예컨대, 아세트산, 옥살산, 말레산, 타타르산, 시트르산, 석신산 또는 말론산 또는 당업계에서 사용되는 다른 방법, 예컨대, 이온 교환을 사용하여 형성된 아미노기의 염이다. 다른 약제학적으로 허용 가능한 염은 아디페이트, 알기네이트, 아스코르베이트, 아스파테이트, 벤젠설포네이트, 벤조에이트, 바이설페이트, 보레이트, 부티레이트, 캄포레이트, 캄포설포네이트, 시트레이트, 사이클로펜탄프로피오네이트, 다이글루코네이트, 도데실설페이트, 에탄설포네이트, 폼에이트, 푸마레이트, 글루코헵토네이트, 글리세로포스페이트, 글루코네이트, 헤미설페이트, 헵타노에이트, 헥사노에이트, 하이드로아이오다이드, 2-하이드록시-에탄설포네이트, 락토비오네이트, 락테이트, 라우레이트, 라우릴 설페이트, 말레이트, 말레에이트, 말로네이트, 메탄설포네이트, 2-나프탈렌설포네이트, 니트레이트, 올레에이트, 옥살레이트, 팔미테이트, 파모에이트, 펙티네이트, 퍼설페이트, 3-페닐프로피오네이트, 포스페이트, 피크레이트, 피발레이트, 프로피오네이트, 스테아레이트, 석시네이트, 설페이트, 타트레이트, 티오사이아네이트, p-톨루엔설포네이트, 운데카노에이트, 발레레이트 염 등을 포함한다. 적절한 염기로부터 유래되는 약제학적으로 허용 가능한 염은 알칼리 금속, 알칼리 토금속, 암모늄 및 $N^{+}(C_{1-4}$알킬$)_4$ 염을 포함한다. 대표적인 알칼리 또는 알칼리 토금속 염은 나트륨, 리튬, 칼륨, 칼슘, 마그네슘 등을 포함한다. 추가의 약제학적으로 허용 가능한 염은 적절한 경우 무독성 암모늄, 4차 암모늄, 및 반대이온, 예컨대, 할라이드, 하이드록사이드, 카복실레이트, 설페이트, 포스페이트, 니트레이트, 저급 알킬 설포네이트 및 아릴 설포네이트를 사용하여 형성된 아민 양이온을 포함한다.

[0163] 전형적인 실시형태에서, 본 발명은 본 명세서에 개시된 화합물, 및 이러한 화합물의 약제학적으로 허용 가능한 염, 약제학적으로 허용 가능한 에스터, 호변이성질체 형태, 다형체 및 전구약물을 포괄하도록 의도된다. 일부 실시형태에서, 본 발명은 본 명세서에 기재된 화합물의 약제학적으로 허용 가능한 부가염, 약제학적으로 허용 가능한 에스터, 부가염의 용매화물(예를 들어, 수화물), 호변이성질체 형태, 다형체, 거울상이성질체, 거울상이성질체의 혼합물, 입체이성질체 또는 입체이성질체의 혼합물(순수한 것 또는 라세미 또는 비-라세미 혼합물)을 포함한다.

[0164] 본 명세서에 기재된 화합물은 하나 이상의 비대칭 중심을 포함할 수 있고, 따라서 다양한 이성질체 형태, 예를 들어, 거울상이성질체 및/또는 부분입체이성질체로 존재할 수 있다. 예를 들어, 본 명세서에 기재된 화합물은 개별 거울상이성질체, 부분입체이성질체 또는 기하 이성질체의 형태로 존재할 수 있거나, 라세미 혼합물 및 하나 이상의 입체이성질체가 풍부한 혼합물을 비롯한 입체이성질체의 혼합물 형태로 존재할 수 있다. 이성질체는

카이럴 고압 액체 크로마토그래피(HPLC) 및 카이럴 염의 형성 및 결정화를 비롯한, 당업계에 공지된 방법에 의해서 혼합물로부터 단리될 수 있거나; 바람직한 이성질체는 비대칭 합성에 의해 제조될 수 있다. 예를 들어, 문헌[Jacques *et al.*, *Enantiomers, Racemates and Resolutions* (Wiley Interscience, New York, 1981); Wilen *et al.*, *Tetrahedron* 33:2725 (1977); Eliel, *Stereochemistry of Carbon Compounds* (McGraw-Hill, NY, 1962); 및 Wilen, *Tables of Resolving Agents and Optical Resolutions* p. 268 (E.L. Eliel, Ed., Univ. of Notre Dame Press, Notre Dame, IN 1972)] 참조. 본 발명은 다른 이성질체가 실질적으로 없는 개별 이성질체로서, 그리고 대안적으로 다양한 이성질체의 혼합물로서 본 명세서에 기재된 화합물을 추가로 포함한다.

[0165] 특정 실시형태에서, 화합물 및 이러한 화합물이 성분인 전달 비히클(예를 들어, 지질 나노입자)은 하나 이상의 표적 세포를 형질주입시키는 향상된(예를 들어, 증가된) 능력을 나타낸다. 따라서, 또한 본 명세서에는 하나 이상의 표적 세포를 형질주입시키는 방법이 제공된다. 이러한 방법은 일반적으로 하나 이상의 표적 세포가 내부에 캡슐화된 원형 RNA로 형질주입되도록 하나 이상의 표적 세포를 본 명세서에 개시된 화합물 및/또는 약제학적 조성물과 접촉시키는 단계를 포함한다. 본 명세서에서 사용된 바와 같이, 용어 "형질주입시키다" 또는 "형질주입"은 하나 이상의 캡슐화된 물질(예를 들어, 핵산 및/또는 폴리뉴클레오타이드)을 세포, 또는 바람직하게는 표적 세포로 세포내 도입하는 것을 지칭한다. 용어 "형질주입 효율"은 형질주입에 적용되는 표적 세포에 의해 흡수, 도입 및/또는 발현되는 이러한 캡슐화된 물질(예를 들어, 폴리뉴클레오타이드)의 상대적인 양을 지칭한다. 일부 실시형태에서, 형질주입 효율은 형질주입 후 표적 세포에 의해 생성된 리포터 폴리뉴클레오타이드 산물의 양에 의해 추정될 수 있다. 일부 실시형태에서, 전달 비히클은 높은 형질주입 효율을 갖는다. 일부 실시형태에서, 전달 비히클은 적어도 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 또는 90% 형질주입 효율을 갖는다.

[0166] 본 명세서에서 사용된 바와 같이, 용어 "리포솜"은 일반적으로 하나 이상의 구형 이중층 또는 이중층에 배열된 지질(예를 들어, 양친매성 지질)로 구성된 소포를 지칭한다. 특정 실시형태에서, 리포솜은 지질 나노입자(예를 들어, 본 명세서에 개시된 이온화 가능한 지질 화합물 중 하나 이상을 포함하는 지질 나노입자)이다. 이러한 리포솜은 하나 이상의 표적 세포, 조직 및 기관에 전달될 캡슐화된 circRNA를 함유하는 수성 내부 및 친유성 물질로부터 형성된 막을 갖는 단층 또는 다중층 소포일 수 있다. 특정 실시형태에서, 본 명세서에 기재된 조성물은 하나 이상의 지질 나노입자를 포함한다. 고려되는 리포솜 및 지질 나노입자를 형성하는데 사용될 수 있는 적합한 지질(예를 들어, 이온화 가능한 지질)의 예는 본 명세서에 개시된 화합물(예를 들어, HGT4001, HGT4002, HGT4003, HGT4004 및/또는 HGT4005) 중 하나 이상을 포함한다. 이러한 리포솜 및 지질 나노입자는 추가적인 이온화 가능한 지질 예컨대 C12-200, DLin-KC2-DMA 및/또는 HGT5001, 헬퍼 지질, 구조 지질, PEG-변형된 지질, MC3, DLinDMA, DLinkC2DMA, cKK-E12, ICE, HGT5000, DODAC, DDAB, DMRIE, DOSPA, DOGS, DODAP, DODMA, DMDMA, DODAC, DLenDMA, DMRIE, CLinDMA, CpLinDMA, DMOBA, DOcarbDAP, DLinDAP, DLincarbDAP, DLinCDAP, KLin-K-DMA, DLin-K-XTC2-DMA, HGT4003 및 이들의 조합물을 또한 포함할 수 있다.

[0167] 본 명세서에서 사용된 바와 같이, "비-양이온성 지질", "비-양이온성 헬퍼 지질" 및 "헬퍼 지질"이라는 어구는 상호 교환 가능하게 사용되며, 임의의 중성, 쯔비터이온성 또는 음이온성 지질을 지칭한다.

[0168] 본 명세서에서 사용된 바와 같이, "음이온성 지질"이라는 어구는 생리 pH와 같은 선택된 pH에서 순 음전하를 갖는 다수의 지질 종 중 임의의 것을 지칭한다.

[0169] 본 명세서에서 사용된 바와 같이, "생분해성 지질" 또는 "분해성 지질"이라는 어구는 숙주 환경에서 수분, 수시간 또는 수일의 정도에서 분해되어 이상적으로는 이것을 독성이 더 적게 만들고, 시간에 따라서 숙주에서 축적될 가능성을 적게 하는 다수의 지질 종 중 임의의 것을 지칭한다. 지질에 대한 일반적인 변형은 지질의 생분해성을 증가시키기 위한 에스터 결합 및 이황화 결합을 포함한다.

[0170] 본 명세서에서 사용된 바와 같이, "생분해성 PEG 지질" 또는 "분해성 PEG 지질"이라는 어구는 PEG 분자가 숙주 환경에서 수분, 수시간 또는 수일의 정도에서 지질로부터 절단되어 이상적으로는 이들을 면역원성을 더 적게 만드는 다수의 지질 종 중 임의의 것을 지칭한다. PEG 지질에 대한 일반적인 변형은 지질의 생분해성을 증가시키기 위한 에스터 결합 및 이황화 결합을 포함한다.

[0171] 본 발명의 특정 실시형태에서, 전달 비히클(예를 들어, 지질 나노입자)은 하나 이상의 물질 또는 치료제(예를 들어, circRNA)를 캡슐화하도록 제조된다. 목적하는 치료제(예를 들어, circRNA)를 전달 비히클 내에 혼입하는 과정은 본 명세서에서 "로딩" 또는 "캡슐화"로 지칭된다(문헌[Lasic, *et al.*, FEBS Lett., 312: 255-258, 1992]). 전달 비히클-로딩된 또는 -캡슐화된 물질(예를 들어, circRNA)은 전달 비히클의 내부 공간, 전달 비히클의 이중층 막인 내에 완전히 또는 부분적으로 위치되거나 또는 전달 비히클의 외부 표면과 완전히 또는 부분적으로 회합될 수 있다.

[0172] 본 명세서에서 사용된 바와 같이, 용어 "구조 지질"은 스테롤 및 또한 스테롤 모이어티를 함유하는 지질을 지칭한다.

[0173] 본 명세서에 정의된 바와 같이, "스테롤"은 스테로이드 알코올로 이루어진 스테로이드의 하위군이다.

[0174] 본 명세서에서 사용된 바와 같이, 용어 "구조 지질"은 스테롤 및 또한 스테롤 모이어티를 함유하는 지질을 지칭한다.

[0175] 본 명세서에서 사용된 바와 같이, 용어 "PEG"는 임의의 폴리에틸렌 글리콜 또는 기타 폴리알킬렌 에터 중합체를 의미한다.

[0176] 본 명세서에 일반적으로 정의된 바와 같이, "PEG-OH 지질"(또한 본 명세서에서 "하이드록시-PEG화 지질"로 지칭됨)은 지질 상에 하나 이상의 하이드록실(-OH)기를 갖는 PEG화 지질이다.

[0177] 본 명세서에서 사용된 바와 같이, "인지질"은 포스페이트 모이어티 및 불포화 지방산 쇄와 같은 하나 이상의 탄소 쇄를 포함하는 지질이다.

[0178] 본 명세서에 개시된 모든 뉴클레오타이드 서열은 RNA 서열 또는 상응하는 DNA 서열을 나타낼 수 있다. DNA에서 데옥시티미딘(dT 또는 T)은 RNA에서 우리딘(U)으로 전사된다. 따라서, "T" 및 "U"는 뉴클레오타이드 서열에서 본 명세서에서 상효 교환 가능하게 사용된다.

[0179] 본 발명에서 사용된 것으로서 설명 "서열 동일성" 또는 예를 들어, "와 50% 동일한 서열"을 포함하는 설명은 서열이 비교창에 걸쳐서 뉴클레오타이드-대-뉴클레오타이드 기준 또는 아미노산-대-아미노산 기준으로 동일한 정도를 지칭한다. 따라서, "서열 동일성의 백분율"은 2개의 최적으로 정렬된 열을 비교창에 걸쳐서 비교하고, 동일한 핵산 염기(예를 들어, A, T, C, G, I) 또는 동일한 아미노산 잔기(예를 들어, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys 및 Met)가 두 서열 모두 나타나는 위치의 수를 결정하여 매치된 위치의 수를 얻고, 매치된 위치의 수를 비교창 내의 위치의 총 수(즉, 창 크기)로 나누어 주고, 결과에 100을 곱하여 서열 동일성의 백분율을 수득함으로써 계산될 수 있다. 본 명세서에 기재된 참조 서열 중 임의의 것과 적어도 약 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 뉴클레오타이드 및 폴리펩타이드가 포함되고, 전형적으로 여기서 폴리펩타이드 변이체는 참조 폴리펩타이드의 적어도 하나의 생물학적 활성을 유지한다.

[0180] 용어 "항체"(Ab)는 항원에 특이적으로 결합하는 당단백질 면역글로불린을 비제한적으로 포함한다. 일반적으로, 항체는 이황화 결합, 또는 이의 항원 결합 분자에 의해 상호 연결된 적어도 2개의 중쇄(H) 및 2개의 경쇄(L)를 포함할 수 있다. 각각의 H쇄는 중쇄 가변 영역(본 명세서에서 VH로 약칭됨) 및 중쇄 불변 영역을 포함할 수 있다. 중쇄 불변 영역은 3개의 불변 도메인, CH1, CH2 및 CH3을 포함할 수 있다. 각각의 경쇄는 경쇄 가변 영역(본 명세서에서 VL로 약칭됨) 및 경쇄 불변 영역을 포함할 수 있다. 경쇄 불변 영역은 하나의 불변 도메인, CL을 포함할 수 있다. VH 및 VL 영역은 프레임워크 영역(FR)이라고 하는 보다 보존된 영역이 산재되어 있는 상보성 결정 영역(CDR)이라고 하는 초가변성 영역으로 더 세분화될 수 있다. 각각의 VH 및 VL은 아미노-말단에서 카복시-말단으로 FR1, CDR1, FR2, CDR2, FR3, CDR3 및 FR4의 순서로 배열된 3개의 CDR 및 4개의 FR을 포함할 수 있다. 중쇄 및 경쇄의 가변 영역은 항원과 상호 작용하는 결합 도메인을 함유한다. Ab의 불변 영역은 면역계의 다양한 세포(예를 들어, 효과기 세포) 및 고전적 보체 시스템의 제1 성분을 비롯한 숙주 조직 또는 인자에 대한 면역글로불린의 결합을 매개할 수 있다. 항체는 예를 들어, 단클론성 항체, 재조합 방식으로 생산된 항체, 단일특이적 항체, 다중특이적 항체(이중특이적 항체 포함), 인간 항체, 조작된 항체, 인간화 항체, 키메라 항체, 면역글로불린, 합성 항체, 2개의 중쇄 및 2개의 중쇄를 포함하는 사량체 항체, 항체 경쇄 단량체, 항체 중쇄 단량체, 항체 경쇄 이량체, 항체 중쇄 이량체, 항체 경쇄-항체 중쇄 쌍, 인트라바디, 항체 융합체(본 명세서에서 때때로 "항체 접합체"라고 지칭함), 이종접합체 항체, 단일 도메인 항체, 1가 항체, 단일쇄 항체 또는 단일쇄 가변 단편(scFv), 낙타화 항체, 어피바디, Fab 단편, $F(ab')_2$ 단편, 이황화-연결 가변 단편(sdFv), 항-이디오타입(항-id) 항체(예를 들어, 항-항-Id 항체 포함), 미니바디, 도메인 항체, 합성 항체(본 명세서에서 때때로 "항체 모방체"라고도 함) 및 상기 중 임의의 것의 항원-결합 단편을 포함할 수 있다. 일부 실시형태에서, 본 명세서에 기재된 항체는 다클론성 항체 집단을 지칭한다.

[0181] 면역글로불린은 IgA, 분비성 IgA, IgG 및 IgM을 포함하지만 이들로 제한되지 않는 일반적으로 알려진 임의의 아이소타입으로부터 유래할 수 있다. IgG 서브클래스는 또한 당업자에게 널리 공지되어 있으며 인간 IgG1, IgG2, IgG3 및 IgG4를 포함하지만 이들로 제한되지 않는다. "아이소타입"은 중쇄 불변 영역 유전자에 의해 암호화되는 Ab 클래스 또는 서브클래스(예를 들어, IgM 또는 IgG1)를 지칭한다. 용어 "항체"는 예를 들어, 자연 발생 및 비

-자연 발생 Ab 둘 다를 포함하고; 단클론성 및 다클론성 Ab; 키메라 및 인간화 Ab; 인간 또는 비인간 Ab; 완전 합성 Ab; 및 단일 쇄 Ab를 포함한다. 비인간 Ab는 인간에서 이의 면역원성을 감소시키기 위한 재조합 방법에 의해 인간화될 수 있다. 명시적으로 언급되지 않은 경우 그리고 문맥에서 달리 나타내지 않는 한, 용어 "항체"는 또한 상기에 언급된 면역글로불린 중 임의의 것의 항원-결합 단편 또는 항원-결합 부분을 포함하고, 1가 및 2가 단편 또는 부분 및 단일 쇄 Ab를 포함한다.

[0182] "항원 결합 분자", "항원 결합 부분" 또는 "항체 단편"은 분자가 유래된 항체의 항원 결합 부분(예를 들어, CDR)을 포함하는 임의의 분자를 지칭한다. 항원 결합 분자는 항원 상보성 결정 영역(CDR)을 포함할 수 있다. 항체 단편의 예는 Fab, Fab', $F(ab')_2$, Fv 단편, dAb, 선형 항체, scFv 항체 및 항원 결합 분자로부터 형성된 다중특이적 항체를 포함하지만 이들로 제한되지 않는다. 펩티바디(즉, 펩타이드 결합 도메인을 포함하는 Fc 융합 분자)는 적합한 항원 결합 분자의 또 다른 예이다. 일부 실시형태에서, 항원 결합 분자는 종양 세포 상의 항원에 결합한다. 일부 실시형태에서, 항원 결합 분자는 과증식성 질환에 관련된 세포 상의 항원 또는 바이러스 또는 박테리아 항원에 결합한다. 일부 실시형태에서, 항원 결합 분자는 BCMA에 결합한다. 추가 실시형태에서, 항원 결합 분자는 항원에 특이적으로 결합하는, 이의 상보성 결정 영역 (CDR) 중 하나 이상을 포함하는 항체 단편이다. 추가 실시형태에서, 항원 결합 분자는 단일 쇄 가변 단편(scFv)이다. 일부 실시형태에서, 항원 결합 분자는 아비머(avimer)를 포함하거나 이로 이루어진다.

[0183] 본 발명에서 사용된 바와 같이, 용어 "가변 영역" 또는 "가변 도메인"은 상호 교환 가능하게 사용되며, 당업계에서 일반적이다. 가변 영역은 전형적으로 항체의 일부, 일반적으로 경쇄 또는 중쇄의 일부, 전형적으로 성숙한 중쇄에서 약 아미노-말단 110 내지 120개 아미노산과 성숙한 경쇄에서 약 90 내지 115개 아미노산을 지칭하며, 이들은 항체 중에서 서열이 광범하게 상이하고, 특정 항원에 대한 특정 항체의 결합 및 특이성에 사용된다. 서열 변동성은 상보성 결정 영역(CDR)이라고 불리는 영역에 집중되며, 가변 도메인에서 더 고도로 보존된 영역은 프레임워크 영역(FR)이라고 불린다. 임의의 특정 기전 또는 이론에 얽매이고자 함은 아니지만, 경쇄 및 중쇄의 CDR이 항원과 항체의 상호작용 및 특이성을 주로 담당한다고 여겨진다. 일부 실시형태에서, 가변 영역은 인간 가변 영역이다. 일부 실시형태에서, 가변 영역은 설치류 또는 뮤린 CDR 및 인간 프레임워크 영역(FR)을 포함한다. 특정 실시형태에서, 가변 영역은 영장류(예를 들어, 비-인간 영장류) 가변 영역이다. 특정 실시형태에서, 가변 영역은 설치류 또는 뮤린 CDR 및 영장류(예를 들어, 비-인간 영장류) 프레임워크 영역(FR)을 포함한다.

[0184] 용어 "VL" 및 "VL 도메인"은 항체 또는 이의 항원 결합 분자의 경쇄 가변 영역을 지칭하기 위해 상호 교환 가능하게 사용된다.

[0185] 용어 "VH" 및 "VH 도메인"은 항체 또는 이의 항원 결합 분자의 중쇄 가변 영역을 지칭하기 위해 상호 교환 가능하게 사용된다.

[0186] CDR의 다수의 정의가 일반적으로 사용된다: 카바트 넘버링, 코티아 넘버링, AbM 넘버링, 또는 접촉 넘버링. AbM 정의는 옥스포드 몰레큘라사(Oxford Molecular)의 AbM 항체 모델링 소프트웨어에 의해 사용되는 2개 사이의 절충안이다. 접촉 정의는 이용 가능한 복합체 결정 구조의 분석을 기초로 한다. 용어 "카바트 넘버링" 등의 용어는 관련 기술 분야에서 인식되고, 항체 또는 이의 항원-결합 분자의 중쇄 및 경쇄 가변 영역 내의 아미노산 잔기를 넘버링하는 시스템을 지칭한다. 특정 양상에서, 항체의 CDR은 카바트 넘버링 시스템에 따라 결정될 수 있다(예를 들어, 문헌[Kabat EA & Wu TT (1971) Ann NY Acad Sci 190: 382-391 및 Kabat EA et al., (1991) Sequences of Proteins of Immunological Interest, Fifth Edition, U.S. Department of Health and Human Services, NIH Publication No. 91-3242] 참조). 카바트 넘버링 시스템을 사용하면, 항체 중쇄 분자 내의 CDR은 전형적으로 아미노산 위치 31 내지 35(이는 선택적으로 35 다음에 1 또는 2개의 추가의 아미노산(카바트 넘버링 체계에서 35A 및 35B로 지칭됨)을 포함할 수 있음)(CDR1), 아미노산 위치 50 내지 65(CDR2), 및 아미노산 위치 95 내지 102(CDR3)에 존재한다. 카바트 넘버링 시스템을 사용하면, 항체 경쇄 분자 내의 CDR은 전형적으로 아미노산 위치 24 내지 34 (CDR1), 아미노산 위치50 내지 56(CDR2), 및 아미노산 위치 89 내지 97(CDR3)에 존재한다. 구체적 실시형태에서, 본 명세서에 기재된 항체의 CDR은 카바트 넘버링 체계에 따라 결정되었다. 특정 양상에서, 항체의 CDR은 면역글로불린 구조 루프의 위치를 나타내는 코티아 넘버링 방식에 따라 결정될 수 있다(예를 들어, 문헌[Chothia C & Lesk AM, (1987), J Mol Biol 196: 901-917; Al-Lazikani B et al, (1997) J Mol Biol 273: 927-948; Chothia C et al., (1992) J Mol Biol 227: 799-817; Tramontano A et al, (1990) J Mol Biol 215(1): 175- 82]; 및 미국 특허 제7,709,226호 참조). 전형적으로, 카바트 넘버링 규정을 사용할 때, 코티아 CDR-H1 루프는 중쇄 아미노산 26 내지 32, 33, 또는 34에 존재하고, 코티아 CDR-H2 루프는 중쇄 아미노산 52 내지 56에 존재하고, 및 코티아 CDR-H3 루프는 중쇄 아미노산 95 내지 102에 존재하는 한편, 코티아

CDR-L1 루프는 경쇄 아미노산 24 내지 34에 존재하고, 코티아 CDR-L2 루프는 경쇄 아미노산 50 내지 56에 존재하고, 코티아 CDR-L3 루프는 경쇄 아미노산 89 내지 97에 존재한다. 코티아 CDR-HI 루프의 단부는 카바트 넘버링 규정을 사용하여 넘버링할 경우 루프의 길이에 따라 H32 및 H34 사이에서 달라진다(이는 카바트 넘버링 스킴이 H35A 및 H35B에 삽입을 두기 때문이고; 35A도 존재하지 않고 35B도 존재하지 않으면, 루프는 32에서 종결되고; 35A만 존재하면, 루프는 33에서 종결되고; 35A와 35B 둘 다가 존재하면, 루프는 34에서 종결된다). 구체적 실시형태에서, 본 명세서에 기재된 항체의 CDR은 코티아 넘버링 체계에 따라 결정되었다.

[0187] 본 명세서에 사용된 용어 "불변 영역" 및 "불변 도메인"은 상호 교환 가능하고, 관련 기술분야에서의 통상적인 의미를 갖는다. 불변 영역은 항체 부분, 예를 들어, 항체의 항원에 대한 결합에 직접 수반되지는 않지만 다양한 효과기 기능, 예컨대 Fc 수용체와의 상호작용을 나타낼 수 있는 경쇄 및/또는 중쇄의 카복실 말단 부분이다. 면역글로불린 분자의 불변 영역은 일반적으로 면역글로불린 가변 도메인에 비해 보다 보존된 아미노산 서열을 갖는다.

[0188] "결합 친화도"는 일반적으로, 분자(예를 들어, 항체)의 단일 결합 부위와 이의 결합 파트너(예를 들어, 항원) 사이의 비-공유 상호작용의 총 합계의 강도를 지칭한다. 달리 나타내지 않는 한, 본 명세서에 사용된 "결합 친화도"는 결합 쌍의 구성원(예를 들어, 항체 및 항원) 사이의 1:1 상호작용을 반영하는 고유의 결합 친화도를 지칭한다. 분자 X의 이의 파트너 Y에 대한 친화도는 일반적으로 해리 상수(KD 또는 Kd)로 표현될 수 있다. 친화도는 평형 해리 상수(KD) 및 평형 회합 상수(KA 또는 Ka)를 포함하지만 이들로 제한되지는 않는, 관련 기술 분야에 공지된 수많은 방식으로 측정되고/거나 표현될 수 있다. KD는 koff/kon의 비율로부터 계산되고, 반면에 KA는 kon/koff의 비율로부터 계산된다. kon은 예를 들어, 항체의 항원에 대한 회합 속도 상수를 지칭하고, koff는 예를 들어, 항체의 항원에 대한 해리 속도 상수를 지칭한다. kon 및 koff는 당업자에게 공지된 기술, 예컨대, BIACORE® 또는 KinExA에 의해 결정될 수 있다.

[0189] 본 명세서에서 사용된 바와 같이, "보존적 아미노산 치환"은 아미노산 잔기가 유사한 측쇄를 갖는 아미노산 잔기로 대체된 것이다. 유사한 측쇄를 갖는 아미노산 잔기의 패밀리는 관련 기술분야에 정의되어 있다. 이들 패밀리는 염기성 측쇄(예를 들어, 라이신, 아르기닌, 히스티딘), 산성 측쇄(예를 들어, 아스파르트산, 글루탐산), 비하전된 극성 측쇄(예를 들어, 글리신, 아스파라긴, 글루타민, 세린, 트레오닌, 티로신, 시스테인, 트립토판), 비극성 측쇄(예를 들어, 알라닌, 발린, 류신, 아이소류신, 프롤린, 페닐알라닌, 메티오닌), 베타-분지형 측쇄(예를 들어, 트레오닌, 발린, 아이소류신) 및 방향족 측쇄(예를 들어, 티로신, 페닐알라닌, 트립토판, 히스티딘)를 갖는 아미노산을 포함한다. 특정 실시형태에서, 항체 또는 이의 항원-결합 분자의 CDR(들) 내 또는 프레임워크 영역(들) 내의 1개 이상의 아미노산 잔기는 유사한 측쇄를 갖는 아미노산 잔기로 대체될 수 있다.

[0190] 본 명세서에서 사용된 바와 같이, 용어 "이종"은 자연 발생 서열 이외의 다른 임의의 공급원으로부터의 것을 의미한다.

[0191] 본 명세서에서 사용된 바와 같이, "에피토프"는 관련 기술분야의 용어이며, 항체가 특이적으로 결합할 수 있는 항원의 국지화된 영역을 지칭한다. 에피토프는 예를 들어, 폴리펩타이드의 인접 아미노산일 수 있거나(선형 또는 인접 에피토프), 또는 에피토프는 예를 들어, 폴리펩타이드 또는 폴리펩타이드들의 2개 이상의 비-인접 영역으로부터 함께 모일 수 있다 (입체배좌적, 비-선형, 불연속적, 또는 비-인접 에피토프). 특정 실시형태에서, 항체가 결합하는 에피토프는 예를 들어, NMR 분광분석법, X선 회절 결정학 연구, ELISA 검정, 질량 분광측정법과 커플링된 수소/중수소 교환 (예를 들어, 액체 크로마토그래피 전기분무 질량 분광측정법), 어레이-기반 올리고-펩타이드 스캐닝 검정 및/또는 돌연변이유발 맵핑(예를 들어, 부위-지정 돌연변이유발 맵핑)에 의해 결정될 수 있다. X선 결정학의 경우, 결정화는 관련 기술분야에 공지된 방법 중 임의의 것을 사용하여 달성될 수 있다(예를 들어, 문헌[Giege R *et al.*, (1994) Acta Crystallogr D Biol Crystallogr 50(Pt 4): 339-350; McPherson A (1990) Eur J Biochem 189: 1-23; Chayen NE (1997) Structure 5: 1269- 1274; McPherson A (1976) J Biol Chem 251: 6300-6303]). 항체:항원 결정은 널리 공지된 X선 회절 기술을 사용하여 연구될 수 있고, 컴퓨터 소프트웨어, 예컨대, X-PLOR(예일 대학교, 1992, 몰레큘라 시뮬레이션즈사(Molecular Simulations, Inc.)에 의해 배포됨; 예를 들어, 문헌[Meth Enzymol (1985) volumes 114 & 115, eds Wyckoff HW *et al.*]; 미국 특허 공개 제2004/0014194호), 및 BUSTER(문헌[Bricogne G (1993) Acta Crystallogr D Biol Crystallogr 49(Pt 1): 37-60; Bricogne G (1997) Meth Enzymol 276A: 361-423, ed Carter CW; Roversi P *et al.*, (2000) Acta Crystallogr D Biol Crystallogr 56(Pt 10): 1316-1323] 참조)를 사용하여 더 연구될 수 있다.

[0192] 본 명세서에서 사용된 바와 같이, 항원 결합 분자, 항체, 또는 이의 항원 결합 분자는, 항원 및 제1 결합 분자, 항체, 또는 이의 항원 결합 분자 사이의 상호작용이 참조 결합 분자, 참조 항체, 또는 이의 항원 결합 분자가

항원과 상호작용하는 능력을 차단, 제한, 저해 또는 달리 감소시키는 경우에, 참조 항체 또는 이의 항원 결합 분자와 "교차-경쟁"한다. 교차 경쟁은 완전할 수 있고, 예를 들어, 결합 분자의 항원에 대한 결합이 참조 결합 분자가 항원에 결합하는 능력을 완전히 차단하거나, 또는 이는 부분적일 수 있고, 예를 들어, 결합 분자의 항원에 대한 결합이 참조 결합 분자가 항원에 결합하는 능력을 감소시킨다. 특정 실시형태에서, 참조 항원 결합 분자와 교차-경쟁하는 항원 결합 분자는 참조 항원 결합 분자와 동일한 또는 중첩되는 에피토프에 결합한다. 다른 실시형태에서, 참조 항원 결합 분자와 교차-경쟁하는 항원 결합 분자는 참조 항원 결합 분자와 상이한 에피토프에 결합한다. 수많은 유형의 경쟁적 결합 검정, 예를 들어: 고체상 직접 또는 간접 방사선면역검정(RIA); 고체상 직접 또는 간접 효소 면역검정(EIA); 샌드위치 경쟁 검정(문헌[Stahli *et al.*, 1983, Methods in Enzymology 9:242-253]); 고체상 직접 바이오틴-아비딘 EIA(문헌[Kirkland et al., 1986, J. Immunol. 137:3614-3619]); 고체상 직접 표지된 검정, 고체상 직접 표지된 샌드위치 검정(문헌[Harlow and Lane, 1988, Antibodies, A Laboratory Manual, Cold Spring Harbor Press]); 1-125 표지를 사용하는 고체상 직접 표지 RIA(문헌[Morel et al., 1988, Molec. Immunol. 25:7-15]); 고체상 직접 바이오틴-아비딘 EIA(문헌[Cheung, *et al.*, 1990, Virology 176:546-552]); 및 직접 표지된 RIA(문헌[Moldenhauer *et al.*, 1990, Scand. J. Immunol. 32:77-82])를 사용하여 하나의 항원 결합 분자가 또 다른 것과 경쟁하는지 여부를 결정할 수 있다.

[0193] 본 명세서에서 사용된 바와 같이, 용어 "면역특이적으로 결합하다", "면역특이적으로 인식하다", "특이적으로 결합하다", 및 "특이적으로 인식하다"는 항체와 관련하여 유사한 용어이고, 항원(예를 들어, 에피토프 또는 면역 복합체)에 결합하는 분자를 지칭하며, 이러한 결합은 당업자에 의해 이해되는 바와 같다. 예를 들어, 항원에 특이적으로 결합하는 분자는 일반적으로, 예를 들어, 면역검정, BIACORE®, KinExA 3000 기기(사피디네 인스트루먼츠사(Sapidyne Instruments), 아이다호주 보이스 소재), 또는 관련 기술분야에 공지된 다른 검정에 의해 결정된 바와 같은 보다 낮은 친화도로 다른 펩타이드 또는 폴리펩타이드에 결합할 수 있다. 구체적 실시형태에서, 항원에 특이적으로 결합하는 분자는, 분자가 또 다른 항원에 결합하는 경우의 KA 보다 적어도 2 log, 2.5 log, 3 log, 4 log 또는 그 초과인 KA로 항원에 결합한다.

[0194] "항원"은 면역 반응을 촉발하거나 또는 항체 또는 항원 결합 분자가 결합할 수 있는 임의의 분자를 지칭한다. 면역 반응은 항체 생산, 또는 특이적 면역학적-적격 세포의 활성화, 또는 둘 다를 수반할 수 있다. 당업자는 실질적으로 모든 단백질 또는 펩타이드를 포함하는 임의의 매크로분자가 항원으로서의 역할을 할 수 있다는 것을 용이하게 이해할 것이다. 항원은 내인성으로 발현될 수 있고, 즉 게놈 DNA에 의해 발현될 수 있거나, 또는 재조합 방식으로 발현될 수 있다. 항원은 특정 조직, 예컨대, 암 세포에 특이적일 수 있거나, 또는 광범위하게 발현될 수 있다. 또한, 보다 큰 분자의 단편이 항원으로서 작용할 수 있다. 일부 실시형태에서, 항원은 종양 항원이다.

[0195] 용어 "자가"는 이후에 재-도입될 동일한 개체로부터 유래된 임의의 물질을 지칭한다. 예를 들어, 본 명세서에 기재된 조작된 자가 세포 요법(eACT™) 방법은 환자로부터 림프구를 수집한 후, 이를 예를 들어, CAR 작제물을 발현하도록 조작하고, 이어서 이를 다시 동일한 환자에게 투여하는 것을 수반한다.

[0196] 용어 "동종"은 한 개체로부터 유래되고 이어서 동일한 종의 또 다른 개체에 도입되는 임의의 물질, 예를 들어, 동종 T 세포 이식을 지칭한다.

[0197] "암"은 신체 내에서의 비정상 세포의 비제어된 성장을 특징으로 하는 다양한 질환의 광범위한 군을 지칭한다. 조절되지 않는 세포 분열 및 성장은 이웃하는 조직을 침습하는 악성 종양의 형성을 야기하고, 또한 림프계 또는 혈류를 통해 신체의 원위 부분으로 전이될 수 있다. "암" 또는 "암 조직"은 종양을 포함할 수 있다. 본 발명의 방법에 의해 치료될 수 있는 암의 예는 림프종, 백혈병, 골수종, 및 다른 백혈구 악성종양을 포함한 면역계의 암을 포함하나 이들로 제한되지 않는다. 일부 실시형태에서, 본 명세서에 개시된 방법은 예를 들어, 골암, 췌장암, 피부암, 두경부암, 피부 또는 안내 악성 흑색종, 자궁암, 난소암, 직장암, 항문부암, 위암, 고환암, 자궁암, 다발성 골수종, 호지킨병, 비-호지킨 림프종(NHL), 원발성 종격 대 B 세포 림프종(PMBC), 미만성 대 B 세포 림프종(DLBCL), 여포성 림프종(FL), 형질전환된 여포성 림프종, 비장 변연부 림프종(SMZL), 식도암, 소장암, 내분비계암, 갑상선암, 부갑상선암, 부신암, 요도암, 음경암, 만성 또는 급성 백혈병, 급성 골수성 백혈병, 만성 골수성 백혈병, 급성 림프모구성 백혈병(ALL)(비 T 세포 ALL 포함), 만성 림프구성 백혈병(CLL), 소아기 고형 종양, 림프구성 림프종, 방광암, 신장암 또는 요관암, 중추 신경계(CNS)의 신생물, 원발성 CNS 림프종, 종양 혈관신생, 척수축 종양, 뇌간 신경교종, 뇌하수체 선종, 표피양암, 편평 세포암, T 세포 림프종, 석면에 의해 유발된 것을 포함한 환경적으로 유발된 암, 다른 B 세포 악성종양 및 상기 암의 조합으로부터 유래된 종양의 종양 크기를 감소시키는데 사용될 수 있다. 일부 실시형태에서, 본 명세서에 개시된 방법은 예를 들어, 육종 및 암종, 섬유육종, 점액육종, 지방육종, 연골육종, 골육종, 카포시 육종, 연조직 육종 다른 육종, 활액종,

중피종, 유잉 종양, 평활근육종, 횡문근육종, 결장 암종, 췌장암, 유방암, 난소암, 전립선암, 간세포 암종, 폐암, 결장직장암, 편평 세포 암종, 기저 세포 암종, 선암종(예를 들어, 췌장, 결장, 난소, 폐, 유방, 위, 전립선, 자궁경부 또는 식도의 선암종), 땀샘 암종, 피지선 암종, 유두 암종, 유두 선암종, 수질 암종, 기관지원성 암종, 신장 세포 암종, 간세포암, 담관 암종, 융모막암종, 빌름스 종양, 자궁경부암, 고환 종양, 방광 암종, 나팔관 암종, 자궁내막 암종, 자궁경부 암종, 질 암종, 음문 암종, 신우 암종, CNS 종양(예컨대 신경교종, 성상세포종, 수모세포종, 두개인두종, 뇌실막종, 송과체종, 혈관모세포종, 청신경종, 희소돌기신경교종, 수막종, 흑색종, 신경모세포종 및 망막모세포종)으로부터 유래된 종양의 종양 크기를 감소시키는 데 사용될 수 있다. 특정 암은 화학요법 또는 방사선 요법에 대해 반응성일 수 있거나, 또는 암은 불응성일 수 있다. 불응성 암은 수술적 개입으로 변형될 수 없는 암을 지칭하고, 이러한 암은 처음부터 화학요법 또는 방사선 요법에 대해 비반응성이거나 또는 암이 시간의 경과에 따라 비반응성이 된다.

[0198] 본 명세서에서 사용된 바와 같은 "항종양 효과"는 종양 부피의 감소, 종양 세포 수의 감소, 종양 세포 증식의 감소, 전이 수의 감 소, 전체 또는 무진행 생존의 증가, 기대 수명의 증가, 또는 종양과 연관된 다양한 생리학적 증상의 호전으로서 나타날 수 있는 생물학적 효과를 지칭한다. 항종양 효과는 또한 종양 발생의 방지, 예를 들어, 백신을 지칭할 수 있다.

[0199] 본 명세서에서 사용된 바와 같은 "사이토카인"은 특이적 항원과의 접촉에 반응하여 한 세포에 의해서 방출되는 비-항체 단백질을 지칭하고, 여기서 사이토카인은 제2 세포와 상호작용하여 제2 세포에서의 반응을 매개한다. "사이토카인"은 본 명세서에서 사용된 바와 같이 세포간 매개자로서 또 다른 세포 상에서 작용하는 하나의 세포 집단에 의해서 방출되는 단백질을 지칭하는 의미이다. 사이토카인은 세포에 의해 내인성으로 발현될 수 있거나, 또는 대상체에게 투여될 수 있다. 사이토카인은 대식세포, B 세포, T 세포, 호중구, 수지상 세포, 호산구 및 비만 세포를 비롯한 면역 세포에 의해서 방출되어 면역 반응을 전파할 수 있다. 사이토카인은 수용자 세포에서 다양한 반응을 유도할 수 있다. 사이토카인은 항상성 사이토카인, 케모카인, 전염증성 사이토카인, 효과기 및 급성기 단백질을 포함할 수 있다. 예를 들어, 인터류킨(IL) 7 및 IL-15를 포함하는 항상성 사이토카인은 면역 세포 생존 및 증식을 촉진하고, 전염증성 사이토카인은 염증성 반응을 촉진할 수 있다. 항상성 사이토카인의 예는 IL-2, IL-4, IL-5, IL-7, IL-10, IL-12p40, IL-12p70, IL-15, 및 인터페론(IFN) 감마를 포함하지만 이들로 제한되지 않는다. 전염증성 사이토카인의 예는 IL-1a, IL-1b, IL-6, IL-13, IL-17a, IL-23, IL-27, 종양 괴사 인자(tumor necrosis factor; TNF)-알파, TNF-베타, 섬유모세포 성장 인자(fibroblast growth factor; FGF) 2, 과립구 대식세포 콜로니-자극 인자(granulocyte macrophage colony-stimulating factor; GM-CSF), 가용성 세포간 부착 분자 1 (soluble intercellular adhesion molecule 1; sICAM-1), 가용성 혈관 부착 분자 1(soluble vascular adhesion molecule 1; sVCAM-1), 혈관 내피 성장 인자(vascular endothelial growth factor; VEGF), VEGF-C, VEGF-D 및 태반 성장 인자(placental growth factor; PLGF)를 포함하지만 이들로 제한되지 않는다. 효과기의 예는 그랜자임 A, 그랜자임 B, 가용성 Fas 리간드(sFasL), TGF-베타, IL-35 및 퍼포린을 포함하지만 이들로 제한되지 않는다. 급성기-단백질의 예는 C-반응성 단백질(CRP) 및 혈청 아밀로이드 A(SAA)를 포함하지만 이들로 제한되지 않는다.

[0200] 본 명세서에서 사용된 바와 같은 용어 "림프구"는 자연 킬러(NK) 세포, T 세포, 또는 B 세포를 포함한다. NK 세포는 선천성 면역계의 주요 성분을 대표하는 세포독성(세포 독성) 림프구 유형이다. NK 세포는 종양 및 바이러스에 감염된 세포를 거부한다. 이는 아폽토시스 또는 세포 예정사의 과정을 통해 작용한다. 이들은 세포를 사멸시키기 위해 활성화를 필요로 하지 않기 때문에 "자연 킬러"로 명명되었다. T-세포는 세포-매개-면역(어떠한 항체도 수반되지 않음)에서 주요 역할을 한다. T 세포 수용체(TCR)는 다른 림프구 유형으로부터 T 세포로 분화한다. 면역계의 전문화된 기관인 흉선은 T 세포 성숙을 위한 주요 부위이다. 헬퍼 T-세포(예를 들어, CD4+ 세포), 세포독성 T-세포(또한 TC, 세포독성 T 림프구, CTL, T-킬러 세포, 세포용해 T 세포, CD8+ T-세포 또는 킬러 T 세포로도 알려짐), 기억 T-세포((i) 줄기 기억 세포(TSCM, 예컨대, 미경험 세포는 CD45RO-, CCR7+, CD45RA+, CD62L+(L-셀렉틴), CD27+, CD28+ 및 IL-7Ra+이지만, 또한 다량의 CD95, IL-2R, CXCR3, 및 LFA-1을 발현하고, 기억 세포 특유의 수많은 기능 속성을 나타냄; (ii) 중심 기억 TCM 세포(TCM)는 L-셀렉틴 및 CCR7을 발현하고, IL-2를 분비하지만 IFNγ 또는 IL-4는 분비하지 않음, (iii) 그러나, 효과기 기억 세포(TEM)는 L-셀렉틴 또는 CCR7을 발현하지 않지만, 효과기 사이토카인, 예컨대 IFNγ 및 IL-4를 생산함), 조절 T-세포 (Treg, 억제자 T 세포, 또는 CD4+CD25+ 또는 CD4+ FoxP3+ 조절 T 세포), 자연 킬러 T-세포(NKT) 및 감마 델타 T-세포를 비롯한 다수의 유형의 T 세포가 존재한다. 다른 한편으로는, B-세포는 체액성 면역(항체가 수반됨)에서 중요한 역할을 한다. B-세포는 항체를 생성하고, 항원-제시 세포(APC)로서 작용할 수 있고, 항원 상호작용에 의한 활성화 후에 둘 다 수명이 짧은 및 수명이 긴 B-세포 및 형질 세포로 변한다. 포유동물에서, 미성숙 B-세포는 골수에서 형성

된다.

[0201] 용어 "유전자 조작된" 또는 "조작된"은 암호 또는 비-암호 영역 또는 이의 부분을 결실시키거나 또는 암호 영역 또는 이의 부분을 삽입하는 것을 포함하지만 이에 제한되지는 않는, 세포의 게놈을 변형시키는 방법을 지칭한다. 일부 실시형태에서, 변형되는 세포는 림프구, 예를 들어, T 세포이며, 이는 환자 또는 공여자로부터 수득될 수 있다. 외인성 작제물, 예컨대, 키메라 항원 수용체(CAR) 또는 T 세포 수용체(TCR)를 발현하도록 세포가 변형될 수 있고, 이는 세포의 게놈 내로 혼입된다.

[0202] "면역 반응"은 침입 병원체, 병원체로 감염된 세포 또는 조직, 암성 또는 다른 비정상 세포, 또는 자가면역 또는 병리학적 염증의 경우에 정상 인간 세포 또는 조직을 선택적으로 표적화하고/하거나, 이에 결합하고/하거나, 이에 대해 손상을 입히고/입히거나, 이를 파괴시키고/시키거나, 이를 척추동물의 신체로부터 제거하는, 면역계의 세포(예를 들어, T 림프구, B 림프구, 자연 킬러(NK) 세포, 대식세포, 호산구, 비만 세포, 수지상 세포 및 호중구), 및 이들 세포 중 임의의 것 또는 간에 의해 생산된 가용성 거대분자(Ab, 사이토카인, 및 보체 포함)의 작용을 지칭한다.

[0203] 본 명세서에서 사용된 바와 같은 "공자극성 신호"는 1차 신호, 예컨대 TCR/CD3 결찰과 조합되어, T 세포 반응, 예컨대 비제한적으로, 증식 및/또는 핵심 분자의 상향조절 또는 하향 조절로 이어지는 신호를 지칭한다.

[0204] 본 명세서에서 사용된 바와 같은 "공자극성 리간드"는 T 세포 상의 동족 공동자극 분자에 특이적으로 결합하는 항원 제시 세포 상의 분자를 포함한다. 공자극성 리간드의 결합은 증식, 활성화, 분화 등을 포함하지만 이들로 제한되지는 않는 T 세포 반응을 매개하는 신호를 제공한다. 공자극성 리간드는 자극 분자에 의해, 예를 들어, T 세포 수용체(TCR)/CD3 복합체와 펩타이드가 로딩된 주요 조직적합성 복합체(MHC) 분자의 결합에 의해 제공되는 1차 신호에 더하여 신호를 유도한다. 공자극성 리간드는 3/TR6, 4-IBB 리간드, 톨-유사 수용체에 결합하는 효능제 또는 항체, B7-1(CD80), B7-2 (CD86), CD30 리간드, CD40, CD7, CD70, CD83, 포진 바이러스 진입 매개자(herpes virus entry mediator: HVEM), 인간 백혈구 항원 G(human leukocyte antigen G: HLA-G), ILT4, 면역글로불린-유사 전사체(immunoglobulin-like transcript: ILT) 3, 유도성 공자극성 리간드(ICOS-L), 세포간 부착 분자(ICAM), B7-H3과 특이적으로 결합하는 리간드, 림프독소 베타 수용체, MHC 부류 I 쇄-관련 단백질 A(MICA), MHC 부류 I 쇄-관련 단백질 B(MICB), OX40 리간드, PD-L2, 또는 예정사(PD) L1을 포함할 수 있으나, 이들로 제한되지 않는다. 공자극성 리간드는, 비제한적으로 T 세포 상에 존재하는 공자극성 분자, 예컨대, 비제한적으로 4-1BB, B7-H3, CD2, CD27, CD28, CD30, CD40, CD7, ICOS, CD83과 특이적으로 결합하는 리간드, 림프구 기능-연관 항원-1(LFA-1), 자연 킬러 세포 수용체 C(NKG2C), OX40, PD-1, 또는 종양 괴사 인자 슈퍼패밀리 구성원 14(TNFSF14 또는 LIGHT)와 특이적으로 결합하는 항체를 포함한다.

[0205] "공자극성 분자"는 공자극성 리간드와 특이적으로 결합함으로써 T 세포에 의한 공자극성 반응, 예컨대 비제한적으로 증식을 매개하는 T 세포 상의 동족 결합 파트너이다. 공자극성 분자는 4-1BB/CD137, B7-H3, BAFFR, BLAME(SLAMF8), BTLA, CD 33, CD 45, CD100(SEMA4D), CD103, CD134, CD137, CD154, CD16, CD160(BY55), CD18, CD19, CD19a, CD2, CD22, CD247, CD27, CD276(B7-H3), CD28, CD29, CD3(알파; 베타; 델타; 엡실론; 감마; 제타), CD30, CD37, CD4, CD4, CD40, CD49a, CD49D, CD49f, CD5, CD64, CD69, CD7, CD80, CD83 리간드, CD84, CD86, CD8알파, CD8베타, CD9, CD96(Tactile), CD1-la, CD1-lb, CD1-lc, CD1-ld, CDS, CEACAM1, CRT AM, DAP-10, DNAM1(CD226), Fc 감마 수용체, GADS, GITR, HVEM(LIGHTR), IA4, ICAM-1, ICAM-1, ICOS, Ig 알파(CD79a), IL2R 베타, IL2R 감마, IL7R 알파, 인테그린, ITGA4, ITGA4, ITGA6, IT GAD, ITGAE, ITGAL, ITGAM, ITGAX, ITGB2, ITGB7, ITGB1, KIRDS2, LAT, LFA-1, LFA-1, LIGHT, LIGHT(종양 괴사 인자 슈퍼패밀리 구성원 14; TNFSF14), LTBR, Ly9(CD229), 림프구 기능-연관 항원-1(LFA-1 (CD1 la/CD18), MHC 클래스 I 분자, NKG2C, NKG2D, NKp30, NKp44, NKp46, NKp80(KLRF1), OX40, PAG/Cbp, PD-1, PSGL1, SELPLG(CD162), 신호전달 림프구성 활성화 분자, SLAM(SLAMF1; CD150; IPO-3), SLAMF4(CD244; 2B4), SLAMF6(NTB-A; Ly108), SLAMF7, SLP-76, TNF, TNFr, TNFR2, 톨 리간드 수용체, TRANCE/RANKL, VLA1 또는 VLA-6 또는 이들의 단편, 절단부 또는 이들의 조합물을 포함하지만 이들로 제한되지 않는다.

[0206] 본 발명에서 사용된 것으로서 설명 "서열 동일성" 또는 예를 들어, "와 50% 동일한 서열"을 포함하는 설명은 서열이 비교창에 걸쳐서 뉴클레오타이드-대-뉴클레오타이드 기준 또는 아미노산-대-아미노산 기준으로 동일한 정도를 지칭한다. 따라서, "서열 동일성의 백분율"은 2개의 최적으로 정렬된 열을 비교창에 걸쳐서 비교하고, 동일한 핵산 염기(예를 들어, A, T, C, G, U) 또는 동일한 아미노산 잔기(예를 들어, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys 및 Met)가 두 서열 모두 나타나는 위치의 수를 결정하여 매치된 위치의 수를 얻고, 매치된 위치의 수를 비교창 내의 위치의 총 수(즉, 창 크기)로

나누어 주고, 결과에 100을 곱하여 서열 동일성의 백분율을 수득함으로써 계산될 수 있다. 본 명세서에 기재된 참조 서열 중 임의의 것과 적어도 약 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100% 서열 동일성을 갖는 뉴클레오타이드 및 폴리펩타이드가 포함되고, 전형적으로 여기서 폴리펩타이드의 경우에 폴리펩타이드 변이체는 참조 폴리펩타이드의 적어도 하나의 생물학적 활성을 유지한다.

[0207] 본 명세서에서 사용되는 바와 같이, "백신"은 질환의 예방 및/또는 치료를 위한 면역 생성을 위한 조성물을 지칭한다. 따라서, 백신은 항원을 포함하는 의약이며, 인간 또는 동물에게 투여 시 특정 방어 및 보호 물질을 생성하기 위해 인간 또는 동물에 사용되도록 의도된다.

[0208] 본 명세서에서 사용되는 바와 같이, "신생항원"은 발현된 단백질에서 종양-특이적 돌연변이로부터 발생하는 종양 항원의 부류를 지칭한다.

[0209] **2. 벡터, 전구체 RNA 및 원형 RNA**

[0210] 또한 본 명세서에는 원형 RNA, 원형 RNA로 원형화될 수 있는 전구체 RNA, 및 전구체 RNA 또는 원형 RNA로 전사될 수 있는 벡터(예를 들어, DNA 벡터)가 제공된다.

[0211] 특정 양상에서, 본 명세서에는 포스트 스플라이싱 3' 그룹 I 인트론 단편, 선택적으로 제1 스페이서, 내부 리보솜 유입 부위(IRES), 발현 서열, 선택적으로 제2 스페이서 및 포스트 스플라이싱 5' 그룹 I 인트론 단편을 포함하는 원형 RNA 폴리뉴클레오타이드가 제공된다. 일부 실시형태에서, 이들 영역은 그 순서대로 존재한다. 일부 실시형태에서, 원형 RNA는 본 명세서에 제공된 방법에 의해서 또는 본 명세서에 제공된 벡터로부터 제조된다.

[0212] 특정 실시형태에서, 본 명세서에 제공된(예를 들어, 5' 듀플렉스 형성 영역, 3' 그룹 I 인트론 단편, 선택적으로 제1 스페이서, 내부 리보솜 유입 부위(IRES), 제1 발현 서열, 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, 제2 발현 서열, 선택적으로 제2 스페이서, 5' 그룹 I 인트론 단편 및 3' 듀플렉스 형성 영역을 포함하는) 벡터는 원형화할 수 있는 전구체 선형 RNA 폴리뉴클레오타이드를 형성한다. 일부 실시형태에서, 이러한 전구체 선형 RNA 폴리뉴클레오타이드는 구아노신 뉴클레오타이드 또는 뉴클레오사이드(예를 들어, GTP) 및 2가 양이온(예를 들어, Mg2+)의 존재 하에서 인큐베이션되는 경우 원형화된다.

[0213] 일부 실시형태에서, 본 명세서에 제공된 벡터 및 전구체 RNA 폴리뉴클레오타이드는 제1(5') 듀플렉스 형성 영역 및 제2(3') 듀플렉스 형성 영역을 포함한다. 특정 실시형태에서, 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역은 완전한 또는 불완전한 듀플렉스를 형성할 수 있다. 따라서, 특정 실시형태에서 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역의 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%는 서로와 짝지워진 염기일 수 있다. 일부 실시형태에서, 듀플렉스 형성 영역은 RNA(예를 들어, 비-듀플렉스 형성 영역 서열) 내의 의도되지 않은 서열과 짝지워진 50% 미만(예를 들어, 45% 미만, 40% 미만, 35% 미만, 30% 미만, 25% 미만)의 염기를 갖는다고 예측된다. 일부 실시형태에서, 전구체 RNA 가닥 각각 상에 그리고 그룹 I 인트론 단편에 인접하거나 매우 가깝게 이러한 듀플렉스 형성 영역을 포함시키는 것은 그룹 I 인트론 단편을 서로에 매우 근접하게 하여 스플라이싱 효율을 증가시킨다. 일부 실시형태에서, 듀플렉스 형성 영역은 3 내지 100개 뉴클레오타이드 길이(예를 들어, 3 내지 75개 뉴클레오타이드 길이, 3 내지 50개 뉴클레오타이드 길이, 20 내지 50개 뉴클레오타이드 길이, 35 내지 50개 뉴클레오타이드 길이, 5 내지 25개 뉴클레오타이드 길이, 9 내지 19개 뉴클레오타이드 길이)이다. 일부 실시형태에서, 듀플렉스 형성 영역은 약 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 또는 50개 뉴클레오타이드 길이이다. 일부 실시형태에서, 듀플렉스 형성 영역은 약 9 내지 약 50개 뉴클레오타이드의 길이를 갖는다. 일 실시형태에서, 듀플렉스 형성 영역은 약 9 내지 약 19개 뉴클레오타이드의 길이를 갖는다. 일부 실시형태에서, 듀플렉스 형성 영역은 약 20 내지 약 40개 뉴클레오타이드의 길이를 갖는다. 특정 실시형태에서, 듀플렉스 형성 영역은 약 30개 뉴클레오타이드의 길이를 갖는다.

[0214] 전구체 RNA 원형화 및/또는 원형 RNA로부터의 유전자 발현을 개선하기 위해 2가지 유형의 스페이서가 설계되었다. 첫 번째 유형의 스페이서는 외부 스페이서인데, 즉, 이것은 전구체 RNA에 존재하지만 원형화 시 제거된다. 이론에 얽매이고자 함은 아니지만, 외부 스페이서는 리보자임 자체의 구조를 유지하고 다른 이웃하는 서열 요소가 이의 접힘 및 기능을 방해하는 것을 방지함으로써 리보자임-매개 원형화를 개선할 수 있다고 생각된다. 두 번째 유형의 스페이서는 내부 스페이서인데, 즉, 이것은 전구체 RNA에 존재하고 생성된 원형 RNA에 유지된다. 이론에 얽매이고자 함은 아니지만, 내부 스페이서는 리보자임 자체의 구조를 유지하고 다른 이웃하는 서열 요소, 특히 이웃하는 IRES 및 암호 영역이 이의 접힘 및 기능을 방해하는 것을 방지함으로써 리보자임-매개 원

형화를 개선할 수 있다고 생각된다. 내부 스페이서는 이웃하는 서열 요소, 특히, 인트론 요소가 IRES 내의 서열과 혼성화하는 것을 방지하고 가장 바람직하고 활성인 입체배좌로 접히는 능력을 저해함으로써 IRES로부터의 단백질 발현을 개선할 수 있다고 또한 생각된다.

[0215] 특정 실시형태에서, 본 명세서에 제공된 벡터, 전구체 RNA 및 원형 RNA는 제1(5') 및/또는 제2(3') 스페이서를 포함한다. 일부 실시형태에서, 3' 그룹 I 인트론 단편과 IRES 사이에 스페이서를 포함시키는 것은 이들 영역이 상호작용하는 것을 방지함으로써 그 영역에서 2차 구조를 보존하여, 스플라이싱 효율을 증가시킬 수 있다. 일부 실시형태에서, 제1(3' 그룹 I 인트론 단편과 IRES 사이) 및 제2(발현 서열과 5' 그룹 I 인트론 단편 사이) 스페이서는 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역이 아닌 서로 염기 짝지움을 형성할 것이라고 예측되는 추가 염기 짝지움 영역을 포함한다. 일부 실시형태에서, 이러한 스페이서 염기 짝지움은 그룹 I 인트론 단편을 서로 매우 근접하게 하여 스플라이싱 효율을 추가로 증가시킨다. 추가로, 일부 실시형태에서, 제1 듀플렉스 형성 영역과 제2 듀플렉스 형성 영역 사이의 염기 짝지움, 및 별도로 제1 스페이서와 제2 스페이서 사이의 염기 짝지움의 조합은 염기 짝지움의 인접한 영역이 측접된 그룹 I 인트론 단편을 함유하는 스플라이싱 버블의 형성을 촉진한다. 전형적인 스페이서는 다음 특성 중 하나 이상을 갖는 연속적인 서열이다: 1) 근위 구조, 예를 들어, IRES, 발현 서열 또는 인트론의 간섭을 피할 것으로 예측됨; 2) 적어도 7nt 길이 내지 100nt 이하임; 3) 3' 인트론 단편 뒤 그리고 이에 인접하여 그리고/또는 5' 인트론 단편 이전 그리고 이에 인접하여 위치함; 그리고 4) a) 적어도 5nt 길이인 비구조화 영역, b) 또 다른 스페이서를 포함하는 원위 서열에 대해 적어도 5nt 길이의 염기 짝지움 영역 및 c) 스페이서의 서열로 범주가 제한되는 적어도 7nt 길이의 구조화 영역 중 하나 이상을 함유함. 스페이서는 비구조화 영역, 염기 짝지움 영역, 헤어핀/구조화 영역 및 이들의 조합을 포함하는 몇몇 영역을 가질 수 있다. 일 실시형태에서, 스페이서는 높은 GC 함량을 갖는 구조화 영역을 갖는다. 일 실시형태에서, 스페이서 내의 영역은 동일한 스페이서 내의 또 다른 영역과 염기 짝지움을 형성한다. 일 실시형태에서, 스페이서 염기 내의 영역은 또 다른 스페이서 내의 영역과 염기 짝지움을 형성한다. 일 실시형태에서, 스페이서는 하나 이상의 헤어핀 구조를 포함한다. 일 실시형태에서, 스페이서는 4 내지 12개 뉴클레오타이드의 줄기 및 2 내지 10개 뉴클레오타이드의 루프를 갖는 하나 이상의 헤어핀 구조를 포함한다. 일 실시형태에서, 3' 그룹 I 인트론 단편과 IRES 사이에 추가 스페이서가 존재한다. 일 실시형태에서, 이러한 추가 스페이서는 IRES의 구조화 영역이 3' 그룹 I 인트론 단편의 접힘을 방해하는 것을 방지하거나 이것이 발생하는 정도를 감소시킨다. 일부 실시형태에서, 5' 스페이서 서열은 적어도 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25 또는 30개 뉴클레오타이드 길이이다. 일부 실시형태에서, 5' 스페이서 서열은 100, 90, 80, 70, 60, 50, 45, 40, 35 또는 30개 이하의 뉴클레오타이드 길이이다. 일부 실시형태에서 5' 스페이서 서열은 5 내지 50, 10 내지 50, 20 내지 50, 20 내지 40 및/또는 25 내지 35개 뉴클레오타이드 길이이다. 특정 실시형태에서, 5' 스페이서 서열은 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 또는 50개 뉴클레오타이드 길이이다. 한 실시형태에서, 5' 스페이서 서열은 polyA 서열이다. 다른 실시형태에서, 5' 스페이서 서열은 polyAC 서열이다. 일 실시형태에서, 스페이서는 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 또는 100%의 polyAC 함량을 포함한다. 일 실시형태에서, 스페이서는 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 100% 폴리피리미딘(C/T 또는 C/U) 함량을 포함한다.

[0216] 특정 실시형태에서, 3' 그룹 I 인트론 단편은 3' 스플라이스 부위 다이뉴클레오타이드 및 선택적으로 적어도 1nt 길이(예를 들어, 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 또는 30nt 길이) 및 최대 엑손의 길이의 인접한 엑손 서열을 포함하는 자연 그룹 I 인트론의 3' 근위 단편과 적어도 75% 상동성(예를 들어, 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 상동성)인 연속적인 서열이다. 전형적으로, 5' 그룹 I 인트론 단편은 5' 스플라이스 부위 다이뉴클레오타이드 및 선택적으로 적어도 1nt 길이(예를 들어, 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 또는 30nt 길이) 및 최대 엑손의 길이의 인접한 엑손 서열을 포함하는 자연 그룹 I 인트론의 5' 근위 단편과 적어도 75% 상동성(예를 들어, 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 상동성)인 연속적인 서열이다. 문헌[Umekage et al. (2012)]에 기재된 바와 같이, 3' 그룹 I 인트론 단편 및 5' 그룹 I 인트론 단편의 외부 부분이 원형화에서 제거되어, 본 명세서에 제공된 원형 RNA가 적어도 1nt 길이인 선택적인 엑손 서열에 의해 형성된 3' 그룹 I 인트론 단편 및 적어도 1nt 길이(이러한 서열이 원형화되지 않은 전구체 RNA에 존재하는 경우)인 선택적인 엑손 서열에 의해 형성된 5' 그룹 I 인트론 단편의 일부만을 포함하도록 한다. 원형 RNA에 의해 유지되는 3' 그룹 I 인트론 단편의 부분은 본 명세서에서 "포스트 스플라이싱 3' 그룹 I 인트론 단편"으로 지칭된다. 원형 RNA에 의해 유지되는 5' 그룹 I 인트론 단편의 부분은 본 명세서에서 "포스트 스플라이싱 5' 그룹 I 인트론 단편"으로 지칭된다.

[0217] 특정 실시형태에서, 본 명세서에 제공된 벡터, 전구체 RNA 및 원형 RNA는 내부 리보솜 유입 부위(IRES)를 포함한다. IRES를 포함하면 원형 RNA로부터 하나 이상의 오픈 리딩 프레임(예를 들어, 발현 서열을 형성하는 오픈 리딩 프레임)의 번역이 가능하다. IRES 요소는 진핵의 리보솜 번역 개시 복합체를 끌어당겨 번역 개시를 촉진한다. 예를 들어, 문헌[Kaufman *et al.*, Nuc. Acids Res. (1991) 19:4485-4490; Gurtu *et al.*, Biochem. Biophys. Res. Comm. (1996) 229:295-298; Rees *et al.*, BioTechniques (1996) 20: 102-110; Kobayashi *et al.*, BioTechniques (1996) 21 :399-402; 및 Mosser *et al.*, BioTechniques 1997 22 150-161] 참조.

[0218] 다수의 IRES 서열이 이용 가능하며, 광범위한 바이러스, 예컨대, 피코르나바이러스, 예컨대, 뇌심근염 바이러스(EMCV) UTR의 리더 서열(문헌[Jang *et al.* J. Virol. (1989) 63: 1651-1660]), 폴리오 리더 서열, A형 간염 바이러스 리더, C형 간염 바이러스 IRES, 인간 리노바이러스 타입 2 IRES(문헌[Dobrikova *et al.*, Proc. Natl. Acad. Sci. (2003) 100(25): 15125- 15130]), 구제역 바이러스로부터의 IRES 요소(문헌[Ramesh *et al.*, Nucl. Acid Res. (1996) 24:2697-2700]), 지아디아바이러스 IRES(문헌[Garlapati *et al.*, J. Biol. Chem. (2004) 279(5):3389-3397]) 등으로부터 유래된 서열을 포함한다.

[0219] 일부 실시형태에서, IRES는 타우라 증후군 바이러스, 트리아토마 바이러스, 타일러 뇌척수염 바이러스, 유인원 바이러스 40, 붉은 불개미 바이러스 1, 로팔로시품 파디 바이러스, 세망내피증 바이러스, 인간 폴리오바이러스 1, 갈색날개 노린재 장 바이러스, 카슈미르 벌 바이러스, 인간 리노바이러스 2, 호말로디스카 코아귤라타 바이러스-1, 인간 면역결핍 바이러스 타입 1, 히메토비 P 바이러스, C형 간염 바이러스, A형 간염 바이러스, 간염 GB 바이러스, 구제역 바이러스, 인간 엔테로바이러스 71, 말 비염 바이러스, 엑트로피스 오블리쿠아 피코르나-유사 바이러스, 뇌심근염 바이러스, 초파리 C 바이러스, 인간 콕사키바이러스 B3, 크루시퍼 토바모바이러스, 귀뚜라미 마비병 바이러스, 소 바이러스성 설사 바이러스 1, 블랙 퀸 세포 바이러스, 진딧물 치명 마비 바이러스, 조류 뇌척수염 바이러스, 급성 벌 마비 바이러스, 히비스커스 위황병 원형반점 바이러스, 돼지 열병 바이러스, 인간 FGF2, 인간 SFTPA1, 인간 AML1/RUNX1, 초파리 안테나페디아, 인간 AQP4, 인간 AT1R, 인간 BAG-1, 인간 BCL2, 인간 BiP, 인간 c-IAP1, 인간 c-myc, 인간 eIF4G, 마우스 NDST4L, 인간 LEF1, 마우스 HIF1 알파, 인간 n.myc, 마우스 Gtx, 인간 p27kipl, 인간 PDGF2/c-sis, 인간 p53, 인간 Pim-1, 마우스 Rbm3, 초파리 리퍼, 캐닌 스캠퍼, 초파리 Ubx, 인간 UNR, 마우스 UtrA, 인간 VEGF-A, 인간 XIAP, 털없는 초파리, 에스. 세레비지애 TFIID, 에스. 세레비지애 YAP1, 담배 식각 바이러스, 터닙 주름 바이러스, EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, 피코비르나바이러스, HCV QC64, 인간 코사바이러스 E/D, 인간 코사바이러스 F, 인간 코사바이러스 JMY, 리노바이러스 NAT001, HRV14, HRV89, HRVC-02, HRV-A21, 살리바이러스 A SH1, 살리바이러스 FHB, 살리바이러스 NG-J1, 인간 파레코바이러스 1, 크로히바이러스 B, Yc-3, 로사바이러스 M-7, 산바바이러스 A, 파시바이러스 A, 파시바이러스 A 2, 에코바이러스 14, 인간 파레코바이러스 5, 아이치 바이러스, A형 간염 바이러스 HA16, 포피바이러스, CVA10, 엔테로바이러스 C, 엔테로바이러스 D, 엔테로바이러스 J, 인간 페기바이러스 2, GBV-C GT110, GBV-C K1737, GBV-C 아이오와, 페기바이러스 A 1220, 파시바이러스 A 3, 사펠로바이러스, 로사바이러스 B, 바쿤사 바이러스, 트레모바이러스) A, 돼지 파시바이러스 1, PLV-CHN, 파시바이러스 A, 시시니바이러스, 헤파시바이러스 K, 헤파시바이러스 A, BVDV1, 보더병 바이러스, BVDV2, CSFV-PK15C, SF573 디시스트로바이러스, 후베이 피코르나-유사 바이러스, CRPV, 살리바이러스 A BN5, 살리바이러스 A BN2, 살리바이러스 A 02394, 살리바이러스 A GUT, 살리바이러스 A CH, 살리바이러스 A SZ1, 살리바이러스 FHB, CVB3, CVB1, 에코바이러스 7, CVB5, EVA71, CVA3, CVA12, EV24 또는 eIF4G에 대한 압타머의 IRES 서열이다.

[0220] 단백질 발현을 유도하기 위해서, 원형 RNA는 단백질 암호 서열에 작동 가능하게 연결된 IRES를 포함한다. 예시적인 IRES 서열은 하기 표 17에 제공되어 있다. 일부 실시형태에서, 본 명세서에 개시된 원형 RNA는 표 17의 IRES 서열과 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99% 동일한 IRES 서열을 포함한다. 일부 실시형태에서, 본 명세서에 개시된 원형 RNA는 표 17의 IRES 서열을 포함한다. 예를 들어, IRES의 5' 및/또는 3' 단부를 절단하고, 스페이서 5'를 IRES에 부가하고, 6개 뉴클레오타이드 5'를 번역 개시 부위(코작 서열)로 변형시키고, 대안적인 번역 개시 부위로 변형시키고, 키메라/혼성 IRES 서열을 생성시킴으로써 IRES를 증가시키거나 감소시키기 위해서 IRES 및 액세서리(accessory) 서열의 변형이 본 명세서에 개시되어 있다. 일부 실시형태에서, 본 명세서에 개시된 원형 RNA에서 IRES 서열은 천연 IRES(예를 들어, 표 17에 개시된 천연 IRES)에 비해서 이들 변형 중 하나 이상을 포함한다.

[0221] 다수의 IRES 서열이 이용 가능하며, 광범위한 바이러스, 예컨대, 피코르나바이러스, 예컨대, 뇌심근염 바이러스(EMCV) UTR의 리더 서열(문헌[Jang *et al.* J. Virol. (1989) 63: 1651-1660]), 폴리오 리더 서열, A형 간염 바이러스 리더, C형 간염 바이러스 IRES, 인간 리노바이러스 타입 2 IRES(문헌[Dobrikova *et al.*, Proc. Natl. Acad. Sci. (2003) 100(25): 15125- 15130]), 구제역 바이러스로부터의 IRES 요소(문헌[Ramesh *et al.*, Nucl.

Acid Res. (1996) 24:2697-2700]), 지아디아바이러스 IRES(문헌[Garlapati *et al*., J. Biol. Chem. (2004) 279(5):3389-3397]) 등으로부터 유래된 서열을 포함한다.

[0222] 일부 실시형태에서, IRES는 타우라 증후군 바이러스, 트리아토마 바이러스, 타일러 뇌척수염 바이러스, 유인원 바이러스 40, 붉은 불개미 바이러스 1, 로팔로시품 파디 바이러스, 세망내피증 바이러스, 인간 폴리오바이러스 1, 갈색날개 노린재 장 바이러스, 카슈미르 벌 바이러스, 인간 리노바이러스 2, 호말로디스카 코아귤라타 바이러스-1, 인간 면역결핍 바이러스 타입 1, 히메토비 P 바이러스, C형 간염 바이러스, A형 간염 바이러스, 간염 GB 바이러스, 구제역 바이러스, 인간 엔테로바이러스 71, 말 비염 바이러스, 엑트로피스 오블리쿠아 피코르나-유사 바이러스, 뇌심근염 바이러스, 초파리 C 바이러스, 인간 콕사키바이러스 B3, 크루시퍼 토바모바이러스, 귀뚜라미 마비병 바이러스, 소 바이러스성 설사 바이러스 1, 블랙 퀸 세포 바이러스, 진딧물 치명 마비 바이러스, 조류 뇌척수염 바이러스, 급성 벌 마비 바이러스, 히비스커스 위황병 원형반점 바이러스, 돼지 열병 바이러스, 인간 FGF2, 인간 SFTPA1, 인간 AML1/RUNX1, 초파리 안테나페디아, 인간 AQP4, 인간 AT1R, 인간 BAG-1, 인간 BCL2, 인간 BiP, 인간 c-IAP1, 인간 c-myc, 인간 eIF4G, 마우스 NDST4L, 인간 LEF1, 마우스 HIF1 알파, 인간 n.myc, 마우스 Gtx, 인간 p27kipl, 인간 PDGF2/c-sis, 인간 p53, 인간 Pim-1, 마우스 Rbm3, 초파리 리퍼, 캐닌 스캠퍼, 초파리 Ubx, 인간 UNR, 마우스 UtrA, 인간 VEGF-A, 인간 XIAP, 털없는 초파리, 에스. 세레비지애 TFIID, 에스. 세레비지애 YAP1, 담배 식각 바이러스, 터닙 주름 바이러스, EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, 피코비르나바이러스, HCV QC64, 인간 코사바이러스 E/D, 인간 코사바이러스 F, 인간 코사바이러스 JMY, 리노바이러스 NAT001, HRV14, HRV89, HRVC-02, HRV-A21, 살리바이러스 A SH1, 살리바이러스 FHB, 살리바이러스 NG-J1, 인간 파레코바이러스 1, 크로히바이러스 B, Yc-3, 로사바이러스 M-7, 산바바이러스 A, 파시바이러스 A, 파시바이러스 A 2, 에코바이러스 E14, 인간 파레코바이러스 5, 아이치 바이러스, A형 간염 바이러스 HA16, 포피바이러스, CVA10, 엔테로바이러스 C, 엔테로바이러스 D, 엔테로바이러스 J, 인간 페기바이러스 2, GBV-C GT110, GBV-C K1737, GBV-C 아이오와, 페기바이러스 A 1220, 파시바이러스 A 3, 사펠로바이러스, 로사바이러스 B, 바쿤사 바이러스, 트레모바이러스 A, 돼지 파시바이러스 1, PLV-CHN, 파시바이러스 A, 시시니바이러스, 헤파시바이러스 K, 헤파시바이러스 A, BVDV1, 보더병 바이러스, BVDV2, CSFV-PK15C, SF573 디시스트로바이러스, 후베이 피코르나-유사 바이러스, CRPV, 살리바이러스 A BN5, 살리바이러스 A BN2, 살리바이러스 A 02394, 살리바이러스 A GUT, 살리바이러스 A CH, 살리바이러스 A SZ1, 살리바이러스 FHB, CVB3, CVB1, 에코바이러스 7, CVB5, EVA71, CVA3, CVA12, EV24 또는 eIF4G에 대한 압타머의 IRES 서열이다.

[0223] 일부 실시형태에서, 본 명세서의 폴리뉴클레오타이드는 1개 초과의 발현 서열을 포함한다. 일부 실시형태에서, 원형 RNA는 이시스트론성 RNA이다. 2개 이상의 폴리펩타이드를 암호화하는 서열은 프로테아제 절단 부위를 암호화하는 뉴클레오타이드 서열 또는 리보솜 스키핑 요소에 의해서 분리될 수 있다. 특정 실시형태에서, 리보솜 스키핑 요소는 토세아 아사인아 바이러스(thosea-asigna virus) 2A 펩타이드(T2A), 돼지 테스코바이러스(porcine teschovirus)-12 A 펩타이드(P2A), 구제역 바이러스(foot-and-mouth disease virus) 2A 펩타이드(F2A), 말 비염(equine rhinitis) A vims 2A 펩타이드(E2A), 세포질 다각체병 vims 2A 펩타이드(BmCPV 2A) 또는 B. mori 2A 펩타이드의 무름병 vims(BmIFV 2A)를 암호화한다.

[0224] 특정 실시형태에서, 본 명세서에 제공된 벡터는 3' UTR을 포함한다. 일부 실시형태에서, 3' UTR은 인간 베타 글로빈, 인간 알파 글로빈 제노푸스 베타 글로빈, 제노푸스 알파 글로빈, 인간 프로락틴, 인간 GAP-43, 인간 eEFlal, 인간 타우, 인간 TNFα, 뎅기열 바이러스, 한타바이러스 작은 mRNA, 부냐바이러스 작은 mRNA, 터닙 황색 모자이크 바이러스, C형 간염 바이러스, 풍진 바이러스, 담배 모자이크 바이러스, 인간 IL-8, 인간 액틴, 인간 GAPDH, 인간 튜불린, 히비스커스 위황병 원형반점 바이러스, 우드척 간염 바이러스 번역후 조절 요소, 신드비스 바이러스, 터닙 주름 바이러스, 담배 식각 바이러스 또는 베네수엘라 말 뇌염 바이러스로부터 유래한다.

[0225] 일부 실시형태에서, 본 명세서에 제공된 벡터는 5' UTR을 포함한다. 일부 실시형태에서, 5' UTR은 인간 베타 글로빈, 제노푸스 래비스 베타 글로빈, 인간 알파 글로빈, 제노푸스 래비스 알파 글로빈, 풍진 바이러스, 담배 모자이크 바이러스, 마우스 Gtx, 뎅기열 바이러스, 열 충격 단백질 70kDa 단백질 1A, 담배 알코올 탈수소효소, 담배 식각 바이러스, 터닙 주름 바이러스 또는 아데노바이러스 3엽 리더로부터 유래한다.

[0226] 일부 실시형태에서, 본 명세서에 제공된 벡터는 3' 및/또는 5' 그룹 I 인트론 단편의 외부에 polyA 영역을 포함한다. 일부 실시형태에서 polyA 영역은 적어도 15, 30 또는 60개 뉴클레오타이드 길이이다. 일부 실시형태에서, 하나 또는 두 polyA 영역은 15 내지 50개 뉴클레오타이드 길이이다. 일부 실시형태에서, 하나 또는 두 polyA 영역은 20 내지 25개 뉴클레오타이드 길이이다. polyA 서열은 원형화 시에 제거된다. 따라서, 고체 표면(예를 들어, 수지)에 접합된 polyA 서열, 예컨대, 데옥시티민 올리고뉴클레오타이드(oligo(dT))와의 올리고뉴클레오타이드 혼성화를 사용하여 이의 전구체 RNA로부터 원형 RNA를 분리할 수 있다. 다른 서열이 또한 3' 그룹 I 인트론

단편에 대해서 5'에 5' 그룹 I 인트론 단편에 대해서 3'에 배치될 수 있고, 상보성 서열이 유사하게 원형 RNA 정제를 위해서 사용될 수 있다.

[0227] 일부 실시형태에서, 본 명세서에 제공된 DNA(예를 들어, 벡터), 선형 RNA(예를 들어, 전구체 RNA) 및/또는 원형 RNA 폴리뉴클레오타이드는 300 내지 15000, 300 내지 14000, 300 내지 13000, 300 내지 12000, 300 내지 11000, 300 내지 10000, 400 내지 9000, 500 내지 8000, 600 내지 7000, 700 내지 6000, 800 내지 5000, 900 내지 5000, 1000 내지 5000, 1100 내지 5000, 1200 내지 5000, 1300 내지 5000, 1400 내지 5000 및/또는 1500 내지 5000개 뉴클레오타이드 길이이다. 일부 실시형태에서, 폴리뉴클레오타이드는 적어도 300nt, 400nt, 500nt, 600nt, 700nt, 800nt, 900nt, 1000nt, 1100nt, 1200nt, 1300nt, 1400nt, 1500nt, 2000nt, 2500nt, 3000nt, 3500nt, 4000nt, 4500nt, 5000nt, 6000nt, 7000nt, 8000nt, 9000nt, 10000nt, 11000nt, 12000nt, 13000nt, 14000nt 또는 15000nt 길이이다. 일부 실시형태에서, 폴리뉴클레오타이드는 3000nt, 3500nt, 4000nt, 4500nt, 5000nt, 6000nt, 7000nt, 8000nt, 9000nt, 10000nt, 11000nt, 12000nt, 13000nt, 14000nt, 15000nt 또는 16000nt 이하의 길이이다. 일부 실시형태에서, 본 명세서에 제공된 DNA, 선형 RNA 및/또는 원형 RNA 폴리뉴클레오타이드의 길이는 약 300nt, 400nt, 500nt, 600nt, 700nt, 800nt, 900nt, 1000nt, 1100nt, 1200nt, 1300nt, 1400nt, 1500nt, 2000nt, 2500nt, 3000nt, 3500nt, 4000nt, 4500nt, 5000nt, 6000nt, 7000nt, 8000nt, 9000nt, 10000nt, 11000nt, 12000nt, 13000nt, 14000nt 또는 15000nt이다.

[0228] 일부 실시형태에서, 본 명세서에는 벡터가 제공된다. 특정 실시형태에서, 벡터는 다음 순서로, a) 5' 듀플렉스 형성 영역, b) 3' 그룹 I 인트론 단편, c) 선택적으로, 제1 스페이서 서열, d) IRES, e) 제1 발현 서열, f) a 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, g) 제2 발현 서열, h) 선택적으로, 제2 스페이서 서열, i) 5' 그룹 I 인트론 단편 및 g) 3' 듀플렉스 형성 영역을 포함한다. 일부 실시형태에서, 벡터는 5' 듀플렉스 형성 영역의 상류에 전사 프로모터를 포함한다.

[0229] 일부 실시형태에서, 본 명세서에는 벡터가 제공된다. 특정 실시형태에서, 벡터는, 다음 순서로, a) 5' 듀플렉스 형성 영역, b) 3' 그룹 I 인트론 단편, c) 선택적으로, 제1 스페이서 서열, d) 제1 IRES, e) 제1 발현 서열, f) 제2 IRES, g) 제2 발현 서열, h) 선택적으로, 제2 스페이서 서열, i) 5' 그룹 I 인트론 단편, 및 g) 3' 듀플렉스 형성 영역을 포함한다. 일부 실시형태에서, 벡터는 5' 듀플렉스 형성 영역의 상류에 전사 프로모터를 포함한다.

[0230] 일부 실시형태에서, 본 명세서에는 전구체 RNA가 제공된다. 특정 실시형태에서, 전구체 RNA는 본 명세서에 제공된 벡터의 시험관내 전사에 의해서 생산된 선형 RNA이다. 일부 실시형태에서, 전구체 RNA는 다음 순서로, a) 선택적으로, 5' 듀플렉스 형성 영역, b) 3' 그룹 I 인트론 단편, c) 선택적으로, 제1 스페이서 서열, d) IRES, e) 제1 발현 서열, f) 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, g) 제2 발현 서열, h) 선택적으로, 제2 스페이서 서열, i) 5' 그룹 I 인트론 단편 및 j) 선택적으로, 3' 듀플렉스 형성 영역을 포함한다. 일부 실시형태에서, 전구체 RNA는 다음 순서로, a) 5' 듀플렉스 형성 영역, b) 3' 그룹 I 인트론 단편, c) 선택적으로, 제1 스페이서 서열, d) 제1 IRES, e) 제1 발현 서열, f) 제2 IRES, g) 제2 발현 서열, h) 선택적으로, 제2 스페이서 서열, i) 5' 그룹 I 인트론 단편, 및 g) 3' 듀플렉스 형성 영역을 포함한다. 전구체 RNA는 비변형되거나, 부분적으로 변형되거나 완전히 변형될 수 있다.

[0231] 특정 실시형태에서, 본 명세서에는 원형 RNA가 제공된다. 특정 실시형태에서, 원형 RNA는 본 명세서에 제공된 벡터에 의해서 생산된 원형 RNA이다. 일부 실시형태에서, 원형 RNA는 본 명세서에 제공된 전구체 RNA의 원형화에 의해서 생산된 원형 RNA이다. 특정 실시형태에서, 본 명세서에 제공된 벡터는 원형화할 수 있는 전구체 선형 RNA를 형성한다. 일부 실시형태에서, 이러한 전구체 선형 RNA 폴리뉴클레오타이드는 구아노신 뉴클레오타이드 또는 뉴클레오사이드(예를 들어, GTP) 및 2가 양이온(예를 들어, Mg2+)의 존재 하에서 인큐베이션되는 경우 원형화된다. 일부 실시형태에서, 원형 RNA는 다음 순서로, a) 제1 스페이서 서열, b) IRES, c) 제1 발현 서열, d) 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, e) 제2 발현 서열 및 f) 제2 스페이서 서열을 포함한다. 일부 실시형태에서, 원형 RNA는 다음 순서로, 포스트 스플라이싱 3' 그룹 I 인트론 단편, b) 제1 스페이서 서열, c) IRES, d) 제1 발현 서열, e) 절단 부위를 암호화하는 폴리뉴클레오타이드 서열, f) 제2 발현 서열 및 g) 제2 스페이서 서열, h) 포스트 스플라이싱 5' 그룹 I 인트론 단편을 포함한다. 일부 실시형태에서, 원형 RNA는 다음 순서로, a) 제1 스페이서 서열, b) 제1 IRES, c) 제1 발현 서열, d) 제2 IRES, e) 제2 발현 서열 및 f) 제2 스페이서 서열을 포함한다. 일부 실시형태에서, 원형 RNA는 3' 스플라이스 부위의 3'인 3' 그룹 I 인트론 단편의 일부를 추가로 포함한다. 일부 실시형태에서, 원형 RNA는 5' 스플라이스 부위의 5'인 5' 그룹 I 인트론 단편의 일부를 추가로 포함한다. 일부 실시형태에서, 원형 RNA는 적어도500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13000, 14000 또는

15000개 뉴클레오타이드 크기이다. 원형 RNA는 비변형되거나, 부분적으로 변형되거나 완전히 변형될 수 있다.

[0232] 일부 실시형태에서, 본 명세서에 제공된 벡터 및 전구체 RNA 폴리뉴클레오타이드는 제1(5') 듀플렉스 형성 영역 및 제2(3') 듀플렉스 형성 영역을 포함한다. 특정 실시형태에서, 제1 상동성 영역 및 제2 상동성 영역은 완전한 또는 불완전한 듀플렉스를 형성할 수 있다. 따라서, 특정 실시형태에서 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역의 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%는 서로와 짝지워진 염기일 수 있다. 일부 실시형태에서, 듀플렉스 형성 영역은 RNA(예를 들어, 비-듀플렉스 형성 영역 서열) 내의 의도되지 않은 서열과 짝지워진 50% 미만(예를 들어, 45% 미만, 40% 미만, 35% 미만, 30% 미만, 25% 미만)의 염기를 갖는다고 예측된다. 일부 실시형태에서, 전구체 RNA 가닥 각각 상에 그리고 그룹 I 인트론 단편에 인접하거나 매우 가깝게 이러한 듀플렉스 형성 영역을 포함시키는 것은 그룹 I 인트론 단편을 서로에 매우 근접하게 하여 스플라이싱 효율을 증가시킨다. 일부 실시형태에서, 듀플렉스 형성 영역은 3 내지 100개 뉴클레오타이드 길이(예를 들어, 3 내지 75개 뉴클레오타이드 길이, 3 내지 50개 뉴클레오타이드 길이, 20 내지 50개 뉴클레오타이드 길이, 35 내지 50개 뉴클레오타이드 길이, 5 내지 25개 뉴클레오타이드 길이, 9 내지 19개 뉴클레오타이드 길이)이다. 일부 실시형태에서, 듀플렉스 형성 영역은 약 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 또는 50개 뉴클레오타이드 길이이다. 일부 실시형태에서, 듀플렉스 형성 영역은 약 9 내지 약 50개 뉴클레오타이드의 길이를 갖는다. 일 실시형태에서, 듀플렉스 형성 영역은 약 9 내지 약 19개 뉴클레오타이드의 길이를 갖는다. 일부 실시형태에서, 듀플렉스 형성 영역은 약 20 내지 약 40개 뉴클레오타이드의 길이를 갖는다. 특정 실시형태에서, 듀플렉스 형성 영역은 약 30개 뉴클레오타이드의 길이를 갖는다.

[0233] 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 동일한 발현 서열을 포함하는 mRNA보다 더 높은 기능적 안정성을 갖는다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 동일한 발현 서열, 5moU 변형, 최적화된 UTR, 캡 및/또는 polyA 테일을 포함하는 mRNA보다 더 높은 기능적 안정성을 갖는다.

[0234] 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 적어도 5시간, 10시간, 15시간, 20시간, 30시간, 40시간, 50시간, 60시간, 70시간 또는 80시간의 기능성 반감기를 갖는다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 5 내지 80, 10 내지 70, 15 내지 60 및/또는 20 내지 50시간의 기능성 반감기를 갖는다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 동일한 단백질을 암호화하는 등가의 선형 RNA 폴리뉴클레오타이드의 기능성 반감기보다 더 큰(예를 들어, 적어도 1.5배 더 큰, 적어도 2배 더 큰) 기능성 반감기를 갖는다. 일부 실시형태에서, 기능성 반감기는 기능성 단백질 합성의 검출을 통해서 평가될 수 있다.

[0235] 특정 실시형태에서, 본 명세서에 제공된 벡터, 전구체 RNA 및 원형 RNA는 제1(5') 및/또는 제2(3') 스페이서를 포함한다. 일부 실시형태에서, 3' 그룹 I 인트론 단편과 IRES 사이에 스페이서를 포함시키는 것은 이들 영역이 상호작용하는 것을 방지함으로써 그 영역에서 2차 구조를 보존하여, 스플라이싱 효율을 증가시킬 수 있다. 일부 실시형태에서, 제1(3' 그룹 I 인트론 단편과 IRES 사이) 및 제2(2개의 발현 서열과 5' 그룹 I 인트론 단편 사이) 스페이서는 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역이 아닌 서로 염기 짝지움을 형성할 것이라고 예측되는 추가 염기 짝지움 영역을 포함한다. 다른 실시형태에서, 제1(3' 그룹 I 인트론 단편과 IRES 사이) 및 제2(발현 서열 중 하나와 5' 그룹 I 인트론 단편 사이) 스페이서는 제1 듀플렉스 형성 영역 및 제2 듀플렉스 형성 영역이 아닌 서로 염기 짝지움을 형성할 것이라고 예측되는 추가 염기 짝지움 영역을 포함한다. 일부 실시형태에서, 이러한 스페이서 염기 짝지움은 그룹 I 인트론 단편을 서로 매우 근접하게 하여 스플라이싱 효율을 추가로 증가시킨다. 추가로, 일부 실시형태에서, 제1 듀플렉스 형성 영역과 제2 듀플렉스 형성 영역 사이의 염기 짝지움, 및 별도로 제1 스페이서와 제2 스페이서 사이의 염기 짝지움의 조합은 염기 짝지움의 인접한 영역이 측접된 그룹 I 인트론 단편을 함유하는 스플라이싱 버블의 형성을 촉진한다. 전형적인 스페이서는 다음 특성 중 하나 이상을 갖는 연속적인 서열이다: 1) 근위 구조, 예를 들어, IRES, 발현 서열 또는 인트론의 간섭을 피할 것으로 예측됨; 2) 적어도 7nt 길이 내지 100nt 이하임; 3) 3' 인트론 단편 뒤 그리고 이에 인접하여 그리고/또는 5' 인트론 단편 이전 그리고 이에 인접하여 위치함; 그리고 4) a) 적어도 5nt 길이인 비구조화 영역, b) 또 다른 스페이서를 포함하는 원위 서열에 대해 적어도 5nt 길이의 염기 짝지움 영역 및 c) 스페이서의 서열로 범주가 제한되는 적어도 7nt 길이의 구조화 영역 중 하나 이상을 함유함. 스페이서는 비구조화 영역, 염기 짝지움 영역, 헤어핀/구조화 영역 및 이들의 조합을 포함하는 몇몇 영역을 가질 수 있다. 일 실시형태에서, 스페이서는 높은 GC 함량을 갖는 구조화 영역을 갖는다. 일 실시형태에서, 스페이서 내의 영역은 동일한 스페이서 내의 또 다른 영역과 염기 짝지움을 형성한다. 일 실시형태에서, 스페이서 염기 내의 영역은 또 다른 스페이

서 내의 영역과 염기 짝지움을 형성한다. 일 실시형태에서, 스페이서는 하나 이상의 헤어핀 구조를 포함한다. 일 실시형태에서, 스페이서는 4 내지 12개 뉴클레오타이드의 줄기 및 2 내지 10개 뉴클레오타이드의 루프를 갖는 하나 이상의 헤어핀 구조를 포함한다. 일 실시형태에서, 3' 그룹 I 인트론 단편과 IRES 사이에 추가 스페이서가 존재한다. 일 실시형태에서, 이러한 추가 스페이서는 IRES의 구조화 영역이 3' 그룹 I 인트론 단편의 접힘을 방해하는 것을 방지하거나 이것이 발생하는 정도를 감소시킨다. 일부 실시형태에서, 5' 스페이서 서열은 적어도 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25 또는 30개 뉴클레오타이드 길이이다. 일부 실시형태에서, 5' 스페이서 서열은 100, 90, 80, 70, 60, 50, 45, 40, 35 또는 30개 이하의 뉴클레오타이드 길이이다. 일부 실시형태에서 5' 스페이서 서열은 5 내지 50, 10 내지 50, 20 내지 50, 20 내지 40 및/또는 25 내지 35개 뉴클레오타이드 길이이다. 특정 실시형태에서, 5' 스페이서 서열은 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 또는 50개 뉴클레오타이드 길이이다. 한 실시형태에서, 5' 스페이서 서열은 polyA 서열이다. 다른 실시형태에서, 5' 스페이서 서열은 polyAC 서열이다. 일 실시형태에서, 스페이서는 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 또는 100%의 polyAC 함량을 포함한다. 일 실시형태에서, 스페이서는 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 또는 100% 폴리피리미딘(C/T 또는 C/U) 함량을 포함한다.

[0236] 특정 실시형태에서, 3' 그룹 I 인트론 단편은 3' 스플라이스 부위 다이뉴클레오타이드 및 선택적으로 적어도 1nt 길이(예를 들어, 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 또는 30nt 길이) 및 최대 엑손의 길이의 인접한 엑손 서열을 포함하는 자연 그룹 I 인트론의 3' 근위 단편과 적어도 75% 동일한(예를 들어, 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일한) 연속적인 서열이다. 전형적으로, 5' 그룹 I 인트론 단편은 5' 스플라이스 부위 다이뉴클레오타이드 및 선택적으로 적어도 1nt 길이(예를 들어, 적어도 2, 5, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 또는 30nt 길이) 및 최대 엑손의 길이의 인접한 엑손 서열을 포함하는 자연 그룹 I 인트론의 5' 근위 단편과 적어도 75% 동일한(예를 들어, 적어도 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100% 동일한) 연속적인 서열이다. 문헌[Umekage et al. (2012)]에 기재된 바와 같이, 3' 그룹 I 인트론 단편 및 5' 그룹 I 인트론 단편의 외부 부분이 원형화에서 제거되어, 본 명세서에 제공된 원형 RNA가 적어도 1nt 길이인 선택적인 엑손 서열에 의해 형성된 3' 그룹 I 인트론 단편 및 적어도 1nt 길이(이러한 서열이 원형화되지 않은 전구체 RNA에 존재하는 경우)인 선택적인 엑손 서열에 의해 형성된 5' 그룹 I 인트론 단편의 일부만을 포함하도록 한다. 원형 RNA에 의해 유지되는 3' 그룹 I 인트론 단편의 부분은 본 명세서에서 포스트 스플라이싱 3' 그룹 I 인트론 단편으로 지칭된다. 원형 RNA에 의해 유지되는 5' 그룹 I 인트론 단편의 부분은 본 명세서에서 포스트 스플라이싱 5' 그룹 I 인트론 단편으로 지칭된다.

[0237] 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 등가의 선형 mRNA보다 더 높은 규모의 발현, 예를 들어, 세포에 RNA를 투여한지 24시간 후에 더 높은 규모의 발현을 가질 수 있다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 동일한 발현 서열, 5moU 변형, 최적화된 UTR, 캡 및/또는 polyA 테일을 포함하는 mRNA보다 더 큰 발현 규모를 갖는다.

[0238] 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 적어도 5시간, 10시간, 15시간, 20시간, 30시간, 40시간, 50시간, 60시간, 70시간 또는 80시간의 기능성 반감기를 갖는다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 5 내지 80, 10 내지 70, 15 내지 60 및/또는 20 내지 50시간의 기능성 반감기를 갖는다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 동일한 단백질을 암호화하는 등가의 선형 RNA 폴리뉴클레오타이드의 기능성 반감기보다 더 큰(예를 들어, 적어도 1.5배 더 큰, 적어도 2배 더 큰) 기능성 반감기를 갖는다. 일부 실시형태에서, 기능성 반감기는 기능성 단백질 합성의 검출을 통해서 평가될 수 있다.

[0239] 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 적어도 5시간, 10시간, 15시간, 20시간, 30시간, 40시간, 50시간, 60시간, 70시간 또는 80시간의 반감기를 갖는다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 5 내지 80, 10 내지 70, 15 내지 60 및/또는 20 내지 50시간의 반감기를 갖는다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 동일한 단백질을 암호화하는 등가의 선형 RNA 폴리뉴클레오타이드의 반감기보다 더 큰(예를 들어, 적어도 1.5배 더 큰, 적어도 2배 더 큰) 반감기를 갖는다. 일부 실시형태에서, 원형 RNA 폴리뉴클레오타이드 또는 이의 약제학적 조성물은 미리 결정된 역치 값의 것과 동일하거나 더 큰 인간 세포에서의 기능성 반감기를 갖는다. 일부 실시형태에서 기능성 반감기는 기능성 단백질 검정에 의해서 결정된다. 예를 들어, 일부 실시형태에서, 기능성 반감기는 시험관내 루시퍼라제 검정에 의해서 결정되며, 여기서 가우시아 루시퍼라제(GLuc)의 활성은 원형 RNA 폴리뉴클레오타이드를

발현하는 인간 세포(예를 들어, HepG2)의 배지에서 1, 2, 3, 4, 5, 6, 7 또는 14일에 걸쳐서 1, 2, 6, 12 또는 24시간마다 측정된다 다른 실시형태에서, 기능성 반감기는 생체내 검정에 의해서 결정되며, 여기서 원형 RNA 폴리뉴클레오타이드의 발현 서열에 의해서 암호화된 단백질의 수준은 1, 2, 3, 4, 5, 6, 7 또는 14일에 걸쳐서 1, 2, 6, 12 또는 24시간마다 환자 혈청 또는 조직 샘플에서 측정된다. 일부 실시형태에서, 미리 결정된 역치 값은 원형 RNA 폴리뉴클레오타이드와 동일한 발현 서열을 포함하는 참조 선형 RNA 폴리뉴클레오타이드의 기능성 반감기이다.

[0240] 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 등가의 선형 mRNA보다 더 높은 규모의 발현, 예를 들어, 세포에 RNA를 투여한지 24시간 후에 더 높은 규모의 발현을 가질 수 있다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 동일한 발현 서열, 5moU 변형, 최적화된 UTR, 캡 및/또는 polyA 테일을 포함하는 mRNA보다 더 큰 발현 규모를 갖는다.

[0241] 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 유기체의 면역계 또는 면역 세포의 특정 유형에 노출되는 경우 등가의 mRNA보다 덜 면역원성일 수 있다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 유기체의 면역계 또는 면역 세포의 특정 유형에 노출되는 경우 사이토카인의 조절된 생산과 연관된다. 예를 들어, 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 동일한 발현 서열을 포함하는 mRNA와 비교할 때 유기체의 면역계 또는 면역 세포의 특정 유형에 노출되는 경우 TNFα, RIG-I, IL-2, IL-6, IFNɣ 및/또는 타입 1 인터레폰, 예를 들어, IFN-β1의 감소된 생산과 연관된다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 동일한 발현 서열을 포함하는 mRNA와 비교할 때 유기체의 면역계 또는 면역 세포의 특정 유형에 노출되는 경우 더 적은 TNFα, RIG-I, IL-2, IL-6, IFNɣ 및/또는 타입 1 인터페론, 예를 들어, IFN-β1 전사 유도와 연관된다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 동일한 발현 서열을 포함하는 mRNA보다 더 낮은 면역원성이다. 일부 실시형태에서, 본 명세서에 제공된 원형 RNA는 동일한 발현 서열, 5moU 변형, 최적화된 UTR, 캡 및/또는 polyA 테일을 포함하는 mRNA보다 더 낮은 면역원성이다.

[0242] 일부 실시형태에서, 본 명세서에 기재된 조성물 및 방법은 뉴클레오사이드 변형이 필요하지 않은 등가의 선형 RNA보다 더 높은 안정성 또는 기능적 안정성을 갖는 RNA(예를 들어, circRNA)를 제공한다. 일부 실시형태에서, 뉴클레오사이드 변형이 결여된 RNA를 생산하는 방법은 감소된 중단 전사로 인해 뉴클레오사이드 변형을 함유하는 RNA를 생산하는 방법보다 더 높은 백분율의 전장 전사체를 생산한다. 일부 실시형태에서, 본 명세서에 기재된 조성물 및 방법은 RNA 함유 뉴클레오사이드 변형과 연관된 추가 중단 전사 없이 큰(예를 들어, 5kb, 6kb, 7kb, 8kb, 9kb, 10kb, 11kb, 12kb, 13kb, 14kb 또는 15kb) RNA 작제물을 생산할 수 있다.

[0243] 특정 실시형태에서, 본 명세서에 제공된 원형 RNA는 그 자체로 세포에 형질주입될 수 있거나 DNA 벡터 형태로 형질주입되고, 세포에서 발현될 수 있다. 형질주입된 DNA 벡터로부터의 원형 RNA의 전사는 첨가된 중합효소 또는 세포에 형질주입된 핵산에 의해서 암호화된 중합효소를 통해서, 바람직하게는 내인성 중합효소를 통해서 일어날 수 있다.

[0244] 특정 실시형태에서, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드는 변형된 RNA 뉴클레오타이드 및/또는 변형된 뉴클레오사이드를 포함한다. 일부 실시형태에서, 변형된 뉴클레오사이드는 m^5C(5-메틸시티딘)이다. 또 다른 실시형태에서, 변형된 뉴클레오사이드는 m^5U(5-메틸우리딘)이다. 또 다른 실시형태에서, 변형된 뉴클레오사이드는 m^6A(N^6-메틸아데노신)이다. 또 다른 실시형태에서, 변형된 뉴클레오사이드는 s^2U(2-티오우리딘)이다. 또 다른 실시형태에서, 변형된 뉴클레오사이드는 Ψ(슈도우리딘)이다. 또 다른 실시형태에서, 변형된 뉴클레오사이드는 Um(2'-O-메틸우리딘)이다. 다른 실시형태에서, 변형된 뉴클레오사이드는 m^1A(1-메틸아데노신); m^2A(2-메틸아데노신); Am(2'-O-메틸아데노신); ms^2m^6A(2-메틸티오-N^6-메틸아데노신); i^6A(N^6-아이소펜텐일아데노신); ms^2i6A(2-메틸티오-N^6 아이소펜텐일아데노신); io^6A(N^6-(시스-하이드록시아이소펜텐일)아데노신); ms^2io^6A(2-메틸티오-N^6-(시스-하이드록시아이소펜텐일)아데노신); g^6A(N^6-글리신일카바모일아데노신); t^6A(N^6-트레오닐카바모일아데노신); ms^2t^6A(2-메틸티오-N^6-트레오닐 카바모일아데노신); m^6t^6A(N^6-메틸-N^6-트레오닐카바모일아데노신); hn^6A(N^6-하이드록시노르발릴카바모일아데노신); ms^2hn^6A(2-메틸티오-N^6-하이드록시노르발릴 카바모일아데노신); Ar(p)(2'-O-리보실아데노신(포스페이트)); I(이노신); m^1I(1-메틸이노신); m^1Im(1,2'-O-다이메틸이노신); m^3C(3-메틸시티딘); Cm(2'-O-메틸시티딘); s^2C(2-티오시티딘); ac^4C(N^4-아세틸시티딘); f^5C(5-폼일시티딘); m^5Cm

(5,2'-O-다이메틸시티딘); ac^4Cm(N^4-아세틸-2'-O-메틸시티딘); k^2C(리시딘); m^1G(1-메틸구아노신); m^2G(N^2-메틸구아노신); m^7G(7-메틸구아노신); Gm(2'-O-메틸구아노신); $m^2{}_2G$(N^2,N^2-다이메틸구아노신); m^2Gm(N^2,2'-O-다이메틸구아노신); $m^2{}_2Gm$(N^2,N^2,2'-O-트라이메틸구아노신); Gr(p)(2'-O-리보실구아노신(포스페이트)); yW(위부토신); o_2yW(퍼옥시위부토신); OHyW(하이드록시위부토신); OHyW*(언더변형된(undermodified) 하이드록시위부토신); imG(위오신); mimG(메틸위오신); Q(케오신); oQ(에폭시케오신); galQ(갈락토실-케오신); manQ(만노실-케오신); $preQ_0$(7-사이아노-7-데아자구아노신); $preQ_1$(7-아미노메틸-7-데아자구아노신); G^+(아르카에오신); D(다이하이드로우리딘); m^5Um(5,2'-O-다이메틸우리딘); s^4U(4-티오우리딘); m^5s^2U(5-메틸-2-티오우리딘); s^2Um(2-티오-2'-O-메틸우리딘); acp^3U(3-(3-아미노-3-카복시프로필)우리딘); ho^5U(5-하이드록시우리딘); mo^5U(5-메톡시우리딘); cmo^5U(우리딘 5-옥시아세트산); $mcmo^5U$(우리딘 5-옥시아세트산 메틸 에스터); chm^5U(5-(카복시하이드록시메틸)우리딘)); $mchm^5U$(5-(카복시하이드록시메틸)우리딘 메틸 에스터); mcm^5U(5-메톡시카보닐메틸우리딘); mcm^5Um(5-메톡시카보닐메틸-2'-O-메틸우리딘); mcm^5s^2U(5-메톡시카보닐메틸-2-티오우리딘); nm^5S^2U(5-아미노메틸-2-티오우리딘); mnm^5U(5-메틸아미노메틸우리딘); mnm^5s^2U(5-메틸아미노메틸-2-티오우리딘); mnm^5se^2U(5-메틸아미노메틸-2-셀레노우리딘); ncm^5U(5-카바모일메틸우리딘); ncm^5Um(5-카바모일메틸-2'-O-메틸우리딘); $cmnm^5U$(5-카복시메틸아미노메틸우리딘); $cmnm^5Um$(5-카복시메틸아미노메틸-2'-O-메틸우리딘); $cmnm^5s^2U$(5-카복시메틸아미노메틸-2-티오우리딘); $m^6{}_2A$(N^6,N^6-다이메틸아데노신); Im(2'-O-메틸이노신); m^4C(N^4-메틸시티딘); m^4Cm(N^4,2'-O-다이메틸시티딘); hm^5C(5-하이드록시메틸시티딘); m^3U(3-메틸우리딘); cm^5U(5-카복시메틸우리딘); m^6Am(N^6,2'-O-다이메틸아데노신); $m^6{}_2Am$(N^6,N^6,O-2'-트라이메틸아데노신); $m^{2,7}G$(N^2,7-다이메틸구아노신); $m^{2,2,7}G$(N^2,N^2,7-트라이메틸구아노신); m^3Um(3,2'-O-다이메틸우리딘); m^5D(5-메틸다이하이드로우리딘); f^5Cm(5-폼일-2'-O-메틸시티딘); m^1Gm(1,2'-O-다이메틸구아노신); m^1Am(1,2'-O-다이메틸아데노신); τm^5U(5-타우리노메틸우리딘); τm^5s^2U(5-타우리노메틸-2-티오우리딘)); imG-14(4-데메틸위오신); imG2(아이소위오신); 또는 ac^6A (N^6-아세틸아데노신)이다.

[0245] 일부 실시형태에서, 변형된 뉴클레오사이드는 하기 군으로부터 선택된 화합물을 포함할 수 있다: 피리딘-4-온 리보뉴클레오사이드, 5-아자-우리딘, 2-티오-5-아자-우리딘, 2-티오우리딘, 4-티오-슈도우리딘, 2-티오-슈도우리딘, 5-하이드록시우리딘, 3-메틸우리딘, 5-카복시메틸-우리딘, 1-카복시메틸-슈도우리딘, 5-프로핀일-우리딘, 1-프로핀일-슈도우리딘, 5-타우리노메틸우리딘, 1-타우리노메틸-슈도우리딘, 5-타우리노메틸-2-티오-우리딘, 1-타우리노메틸-4-티오-우리딘, 5-메틸-우리딘, 1-메틸-슈도우리딘, 4-티오-1-메틸-슈도우리딘, 2-티오-1-메틸-슈도우리딘, 1-메틸-1-데아자-슈도우리딘, 2-티오-1-메틸-1-데아자-슈도우리딘, 다이하이드로우리딘, 다이하이드로슈도우리딘, 2-티오-다이하이드로우리딘, 2-티오-다이하이드로슈도우리딘, 2-메톡시우리딘, 2-메톡시-4-티오-우리딘, 4-메톡시-슈도우리딘, 4-메톡시-2-티오-슈도우리딘, 5-아자-시티딘, 슈도아이소시티딘, 3-메틸-시티딘, N4-아세틸시티딘, 5-폼일시티딘, N4-메틸시티딘, 5-하이드록시메틸시티딘, 1-메틸-슈도아이소시티딘, 피롤로-시티딘, 피롤로-슈도아이소시티딘, 2-티오-시티딘, 2-티오-5-메틸-시티딘, 4-티오-슈도아이소시티딘, 4-티오-1-메틸-슈도아이소시티딘, 4-티오-1-메틸-1-데아자-슈도아이소시티딘, 1-메틸-1-데아자-슈도아이소시티딘, 제불라린, 5-아자-제불라린, 5-메틸-제불라린, 5-아자-2-티오-제불라린, 2-티오-제불라린, 2-메톡시-시티딘, 2-메톡시-5-메틸-시티딘, 4-메톡시-슈도아이소시티딘, 4-메톡시-1-메틸-슈도아이소시티딘, 2-아미노퓨린, 2, 6-다이아미노퓨린, 7-데아자-아데닌, 7-데아자-8-아자-아데닌, 7-데아자-2-아미노퓨린, 7-데아자-8-아자-2-아미노퓨린, 7-데아자-2,6-다이아미노퓨린, 7-데아자-8-아자-2,6-다이아미노퓨린, 1-메틸아데노신, N6-메틸아데노신, N6-아이소펜텐일아데노신, N6-(시스-하이드록시아이소펜텐일)아데노신, 2-메틸티오-N6-(시스-하이드록시아이소펜텐일) 아데노신, N6-글리신일카바모일아데노신, N6-트레오닐카바모일아데노신, 2-메틸티오-N6-트레오닐 카바모일아데노신, N6,N6-다이메틸아데노신, 7-메틸아데닌, 2-메틸티오-아데닌, 2-메톡시-아데닌, 이노신, 1-메틸-이노신, 위오신, 위부토신, 7-데아자-구아노신, 7-데아자-8-아자-구아노신, 6-티오-구아노신, 6-티오-7-데아자-구아노신, 6-티오-7-데아자-8-아자-구아노신, 7-메틸-구아노신, 6-티오-7-메틸-구아노신, 7-메틸이노신, 6-메톡시-구아노신, 1-메틸구아노신, N2-메틸구아노신, N2,N2-다이메틸구아노신, 8-옥소-구아노신, 7-메틸-8-옥소-구아노신, 1-메틸-6-티오-구아노신, N2-메틸-6-티오-구아노신 및 N2,N2-다이메틸-6-티오-구아노신. 또 다른 실시

형태에서, 변형은 5-메틸사이토신, 슈도우리딘 및 1-메틸슈도우리딘으로 이루어진 군으로부터 독립적으로 선택된다.

[0246] 일부 실시형태에서, 변형된 리보뉴클레오사이드는 5-메틸시티딘, 5-메톡시우리딘, 1-메틸-슈도우리딘, N6-메틸아데노신 및/또는 슈도우리딘을 포함한다. 일부 실시형태에서, 이러한 변형된 뉴클레오사이드는 추가 안정성 및 면역 활성화에 대한 내성을 제공한다.

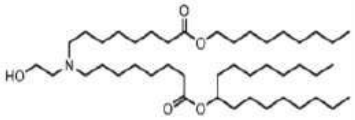
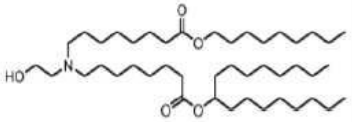
[0247] 특정 실시형태에서, 폴리뉴클레오타이드는 코돈-최적화될 수 있다. 코돈 최적화된 서열은 폴리펩타이드를 암호화하는 폴리뉴클레오타이드 내의 코돈이 폴리펩타이드의 발현, 안정성 및/또는 활성을 증가시키기 위해서 치환된 것일 수 있다. 코돈 최적화에 영향을 미치는 인자는 하기 중 하나 이상을 포함하지만 이들로 제한되지 않는다: (i) 2개 이상의 유기체 또는 유전자 또는 합성에 의해서 작제된 편향 표 사이의 코돈 편향의 변화, (ii) 유기체, 유전자 또는 유전자 세트 내의 코돈 편향 정도의 변화, (iii) 컨텍스트를 포함하는 코돈의 체계적 변화, (iv) 해독 tRNA에 따른 코돈의 변화, (v) 전체 또는 삼중 중 하나의 위치에서의 GC %에 따른 코돈의 변화, (vi) 참조 서열, 예를 들어, 자연 발생 서열과의 유사성 정도의 변화, (vii) 코돈 빈도 컷오프의 변화, (viii) DNA 서열로부터 전사된 mRNA의 구조적 특성, (ix) 코돈 치환 세트의 설계가 기반으로 할 DNA 서열의 기능에 대한 사전 지식 및/또는 (x) 각각의 아미노산에 대한 코돈 세트의 체계적 변화. 일부 실시형태에서, 코돈 최적화된 폴리뉴클레오타이드는 리보자임 충돌을 최소화할 수 있고/있거나 발현 서열과 IRES 사이의 구조적 개입을 제한할 수 있다.

[0248] 특정 실시형태에서 본 명세서에 제공된 원형 RNA는 세포 내에서 생산된다. 일부 실시형태에서, 전구체 RNA는 박테리오파지 RNA 중합효소에 의해서 세포질 내에서 또는 숙주 RNA 중합효소 II에 의해서 핵 내에서 DNA 주형을 사용하여(예를 들어, 일부 실시형태에서, 본 명세서에 제공된 벡터를 사용하여) 전사되고, 그 다음 원형화된다.

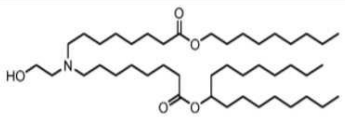
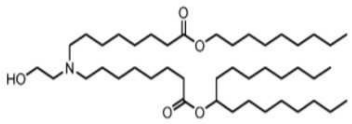
[0249] 특정 실시형태에서, 본 명세서에 제공된 원형 RNA는 동물(예를 들어, 인간)에게 주입되어, 원형 RNA 분자에 의해서 암호화된 폴리펩타이드가 동물에서 발현된다.

[0250] **3. 페이로드**

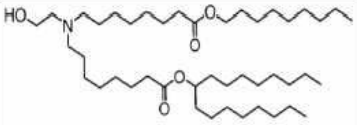
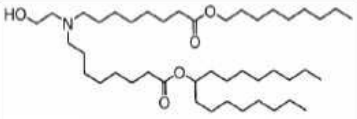
[0251] 일부 실시형태에서, 제1 발현 서열은 치료용 단백질을 암호화한다. 일부 실시형태에서, 제2 발현 서열은 치료용 단백질을 암호화한다. 일부 실시형태에서, 치료용 단백질 중 하나 또는 둘 다는 하기 표에 열거된 단백질로부터 선택된다.

페이로드	서열	표적 세포 / 기관	바람직한 전달 제형
CD19 CAR	서열 309 내지 314 중 임의의 것	T 세포	 (50 mol %) DSPC (10 mol %) 베타-시토스테롤(28.5% mol %) 콜레스테롤 (10 mol %) PEG DMG (1.5 mol %)
BCMA CAR	MALPVTALLLPLALLL HAARPDIVLTQSPASLA VSLGERATINCRASESV SVIGAHLIHWYQQKPG QPPKLLIYLASNLETGV PARFSGSGSGTDFTLTIS SLQAEDAAIYYCLQSRI FPRTFGQGTKLEIKGST SGSGKPGSGEGSTKGQ VQLVQSGSELKKPGAS VKVSCKASGYTFTDYSI NWVRQAPGQGLEWMG WINTETREPAYAYDFR GRFVFSLDTSVSTAYLQ ISSLKAEDTAVYYCAR DYSYAMDYWGQGTLV TVSSAAATTTPAPRPPT PAPTIASQPLSLRPEACR PAAGGAVHTRGLDFAC DIYIWAPLAGTCGVLLL SLVITLYCKRGRKKLLY IFKQPFMRPVQTTQEED GCSCRFPEEEEGGCELR VKFSRSADAPAYQQGQ NQLYNELNLGRREEYD VLDKRRGRDPEMGGKP RRKNPQEGLYNELQKD KMAEAYSEIGMKGERR RGKGHDGLYQGLSTAT KDTYDALHMQALPPR	T 세포	 (50 mol %) DSPC (10 mol %) 베타-시토스테롤(28.5% mol %) 콜레스테롤 (10 mol %) PEG DMG (1.5 mol %)

[0252]

MAGE-A4 TCR	TCR 알파쇄: KNQVEQSPQSLIILEGK NCTLQCNYTVSPFSNLR WYKQDTGRGPVSLTIM TFSENTKSNGRYTATLD ADTKQSSLHITASQLSD SASYICVVNHSGGSYIP TFGRGTSLIVHPYIQKP DPAVYQLRDSKSSDKS VCLFTDFDSQTNVSQSK DSDVYITDKTVLDMRS MDFKSNSAVAWSNKS DFACANAFNNSIIPEDT FFPSPESS TCR 베타쇄: DVKVTQSSRYLVKRTG EKVFLECVQDMDHEN MFWYRQDPGLGLRLIY FSYDVKMKEKGDIPEG YSVSREKKERFSLILES ASTNQTSMYLCASSFL MTSGDPYEQYFGPGTR LTVTEDLKNVFPPEVA VFEPSEAEISHTQKATL VCLATGFYPDHVELSW WVNGKEVHSGVSTDPQ PLKEQPALNDSRYCLSS RLRVSATFWQNPRNHF RCQVQFYGLSENDEWT QDRAKPVTQIVSAEAW GRAD	T 세포	 (50 mol %) DSPC (10 mol %) 베타-시토스테롤(28.5% mol %) 콜레스테롤 (10 mol %) PEG DMG (1.5 mol %)
NY-ESO TCR	TCR 알파 세포외 서열 MQEVTQIPAALSVPEGE NLVLNCSFTDSAIYNLQ WFRQDPGKGLTSLLLIQ SSQREQTSGRLNASLDK SSGRSTLYIAASQPGDS ATYLCAVRPTSGGSYIP TFGRGTSLIVHPY TCR 베타 세포외 서열 MGVTQTPKFQVLKTGQ SMTLQCAQDMNHEYM SWYRQDPGMGLRLIHY SVGAGITDQGEVPNGY NVSRSTTEDFPLRLLSA APSQTSVYFCASSYVG NTGELFFGEGSRLTVL	T 세포	 (50 mol %) DSPC (10 mol %) 베타-시토스테롤(28.5% mol %) 콜레스테롤 (10 mol %) PEG DMG (1.5 mol %)

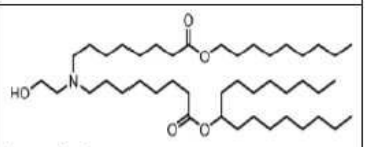
[0253]

EPO	APPRLICDSRVLERYLL EAKEAENITTGCAEHCS LNENITVPDTKVNFYA WKRMEVGQQAVEVW QGLALLSEAVLRGQAL LVNSSQPWEPLQLHVD KAVSGLRSLTTLLRALG AQKEAISPPDAASAAPL RTITADTFRKLFRVYSN FLRGKLKLYTGEACRT GDR	신장 또는 골수	
PAH	MSTAVLENPGLGRKLS DFGQETSYIEDNCNQN GAISLIFSLKEEVGALA KVLRLFEENDVNLTHIE SRPSRLKKDEYEFFTHL DKRSLPALTNIIKILRHD IGATVHELSRDKKKDT VPWFPRTIQELDRFANQ ILSYGAELDADHPGFKD PVYRARRKQFADIAYN YRHGQPIPRVEYMEEE KKTWGTVFKTLKSLYK THACYEYNHIFPLLEKY CGFHEDNIPQLEDVSQF LQTCTGFRLRPVAGLLS SRDFLGGLAFRVFHCT QYIRHGSKPMYTPEPDI CHELLGHVPLFSDRSFA QFSQEIGLASLGAPDEY IEKLATIYWFTVEFGLC KQGDSIKAYGAGLLSSF GELQYCLSEKPKLLPLE LEKTAIQNYTVTEFQPL YYVAESFNDAKEKVRN FAATIPRPFSVRYDPYT QRIEVLDNTQQLKILAD SINSEIGILCSALQKIK	간 세포	 (50 mol %) DSPC (10 mol %) 콜레스테롤 (38.5% mol %) PEG-DMG (1.5%) 또는 MC3 (50 mol %) DSPC (10 mol %) 콜레스테롤 (38.5% mol %) PEG-DMG (1.5%)
CPS1	LSVKAQTAHIVLEDGT KMKGYSFGHPSSVAGE VVFNTGLGGYPEAITDP AYKGQILTMANPIIGNG GAPDTTALDELGLSKY LESNGIKVSGLLVLDYS KDYNHWLATKSLGQW LQEEKVPAIYGVDTRM LTKIIRDKGTMLGKIEF EGQPVDFVDPNKQNLI AEVSTKDVKVYGKGNP TKVVAVDCGIKNNVIR	간 세포	 (50 mol %) DSPC (10 mol %) 콜레스테롤 (38.5% mol %) PEG-DMG (1.5%)

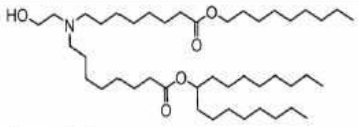
[0254]

	LLVKRGAEVHLVPWN HDFTKMEYDGILIAGGP GNPALAEPLIQNVRKIL ESDRKEPLFGISTGNLIT GLAAGAKTYKMSMAN RGQNQPVLNITNKQAFI TAQNHGYALDNTLPAG WKPLFVNVNDQTNEGI MHESKPFFAVQFHPEV TPGPIDTEYLFDSFFSLI KKGKATTITSVLPKPAL VASRVEVSKVLILGSGG LSIGQAGEFDYSGSQAV KAMKEENVKTVLMNP NIASVQTNEVGLKQAD TVYFLPITPQFVTEVIKA EQPDGLILGMGGQTAL NCGVELFKRGVLKEYG VKVLGTSVESIMATED RQLFSDKLNEINEKIAPS FAVESIEDALKAADTIG YPVMIRSAYALGGLGS GICPNRETLMDLSTKAF AMTNQILVEKSVTGWK EIEYEVVRDADDNCVT VCNMENVDAMGVHTG DSVVVAPAQTLSNAEF QMLRRTSINVVRHLGIV GECNIQFALHPTSMEYC IIEVNARLSRSSALASK ATGYPLAFIAAKIALGIP LPEIKNVVSGKTSACFE PSLDYMVTKIPRWDLD RFHGTSSRIGSSMKSVG EVMAIGRTFEESFQKAL RMCHPSIEGFTPRLPMN KEWPSNLDLRKELSEPS STRIYAIAKAIDDNMSL DEIEKLTYIDKWFLYK MRDILNMEKTLKGLNS ESMTEETLKRAKEIGFS DKQISKCLGLTEAQTRE LRLKKNIHPWVKQIDTL AAEYPSVTNYLYVTYN GQEHDVNFDDHGMMV LGCGPYHIGSSVEFDW CAVSSIRTLRQLGKKTV VVNCNPETVSTDFDEC DKLYFEELSLERILDIYH QEACGGCIISVGGQIPN		또는 MC3 (50 mol %) DSPC (10 mol %) 콜레스테롤 (38.5% mol %) PEG-DMG (1.5%)

[0255]

	NLAVPLYKNGVKIMGT SPLQIDRAEDRSIFSAVL DELKVAQAPWKAVNT LNEALEFAKSVDYPCLL RPSYVLSGSAMNVVFS EDEMKKFLEEATRVSQ EHPVVLTKFVEGAREV EMDAVGKDGRVISHAI SEHVEDAGVHSGDATL MLPTQTISQGAIEKVKD ATRKIAKAFAISGPFNV QFLVKGNDVLVIECNL RASRSFPFVSKTLGVDF IDVATKVMIGENVDEK HLPTLDHPIIPADYVAIK APMFSWPRLRDADPILR CEMASTGEVACFGEGI HTAFLKAMLSTGFKIPQ KGILIGIQQSFRPRFLGV AEQLHNEGFKLFATEA TSDWLNANNVPATPVA WPSQEGQNPSLSSIRKLI RDGSIDLVINLPNNNTK FVHDNYVIRRTAVDSGI PLLTNFQVTKLFAEAV QKSRKVDSKSLFHYRQ YSAGKAA		
Cas9	MKRNYILGLDIGITSVG YGIIDYETRDVIDAGVR LFKEANVENNEGRRSK RGARRLKRRRRHRIQR VKKLLFDYNLLTDHSE LSGINPYEARVKGLSQK LSEEEFSAALLHLAKRR GVHNVNEVEEDTGNEL STKEQISRNSKALEEKY VAELQLERLKKDGEVR GSINRFKTSDYVKEAK QLLKVQKAYHQLDQSF IDTYIDLLETRRTYYEG PGEGSPFGWKDIKEWY EMLMGHCTYFPEELRS VKYAYNADLYNALND LNNLVITRDENEKLEYY EKFQIIENVFKQKKKPT LKQIAKEILVNEEDIKG YRVTSTGKPEFTNLKV YHDIKDITARKEIIENAE LLDQIAKILTIYQSSEDI QEELTNLNSELTQEEIE	면역 세포	 (50 mol %) DSPC (10 mol %) 베타-시토스테롤(28.5% mol %) 콜레스테롤 (10 mol %) PEG DMG (1.5 mol %)

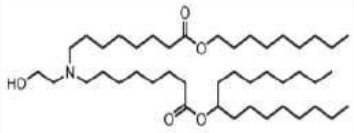
[0256]

	QISNLKGYTGTHNLSLK AINLILDELWHTNDNQI AIFNRLKLVPKKVDLSQ QKEIPTTLVDDFILSPVV KRSFIQSIKVINAIIKKY GLPNDIIIELAREKNSKD AQKMINEMQKRNRQT NERIEEIIRTTGKENAKY LIEKIKLHDMQEGKCLY SLEAIPLEDLLNNPFNY EVDHIIPRSVSFDNSFNN KVLVKQEENSKKGNRT PFQYLSSSDSKISYETFK KHILNLAKGKGRISKTK KEYLLEERDINRFSVQK DFINRNLVDTRYATRG LMNLLRSYFRVNNLDV KVKSINGGFTSFLRRK WKFKKERNKGYKHHA EDALIIANADFIFKEWK KLDKAKKVMENQMFE EKQAESMPEIETEQEYK EIFITPHQIKHIKDFKDY KYSHRVDKKPNRELIN DTLYSTRKDDKGNTLI VNNLNGLYDKDNDKL KKLINKSPEKLLMYHH DPQTYQKLKLIMEQYG DEKNPLYKYYEETGNY LTKYSKKDNGPVIKKIK YYGNKLNAHLDITDDY PNSRNKVVKLSLKPYR FDVYLDNGVYKFVTVK NLDVIKKENYYEVNSK CYEEAKKLKKISNQAEF IASFYNNDLIKINGELY RVIGVNNDLLNRIEVN MIDITYREYLENMNDK RPPRIIKTIASKTQSIKK YSTDILGNLYEVKSKK HPQIIKKG		
ADAM TS13	AAGGILHLELLVAVGP DVFQAHQEDTERYVLT NLNIGAELLRDPSLGAQ FRVHLVKMVILTEPEG APNITANLTSSLLSVCG WSQTINPEDDTDPGHA DLVLYITRFDLELPDGN RQVRGVTQLGGACSPT WSCLITEDTGFDLGVTI	간세포	 (50 mol %) DSPC (10 mol %) 콜레스테롤 (38.5% mol %)

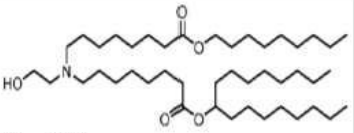
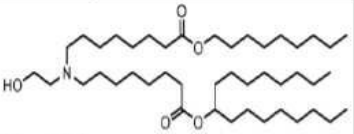
[0257]

	AHEIGHSFGLEHDGAPG SGCGPSGHVMASDGAA PRAGLAWSPCSRRQLL SLLSAGRARCVWDPPR PQPGSAGHPPDAQPGL YYSANEQCRVAFGPKA VACTFAREHLDMCQAL SCHTDPLDQSSCSRLLV PLLDGTECGVEKWCSK GRCRSLVELTPIAAVHG RWSSWGPRSPCSRSCG GGVVTRRRQCNNPRPA FGGRACVGADLQAEM CNTQACEKTQLEFMSQ QCARTDGQPLRSSPGG ASFYHWGAAVPHSQG DALCRHMCRAIGESFIM KRGDSFLDGTRCMPSG PREDGTLSLCVSGSCRT FGCDGRMDSQQVWDR CQVCGGDNSTCSPRKG SFTAGRAREYVTFLTVT PNLTSVYIANHRPLFTH LAVRIGGRYVVAGKMS ISPNTTYPSLLEDGRVE YRVALTEDRLPRLEEIRI WGPLQEDADIQVYRRY GEEYGNLTRPDITFTYF QPKPRQAWVWAAVRG PCSVSCGAGLRWVNYS CLDQARKELVETVQCQ GSQQPPAWPEACVLEP CPPYWAVGDFGPCSAS CGGGLRERPVRCVEAQ GSLLKTLPPARCRAGA QQPAVALETCNPQPCP ARWEVSEPSSCTSAGG AGLALENETCVPGADG LEAPVTEGPGSVDEKLP APEPCVGMSCPPGWGH LDATSAGEKAPSPWGSI RTGAQAAHVWTPAAG SCSVSCGRGLMELRFLC MDSALRVPVQEELCGL ASKPGSRREVCQAVPC PARWQYKLAACSVSCG RGVVRRILYCARAHGE DDGEEILLDTQCQGLPR PEPQEACSLEPCPPRWK VMSLGPCSASCGLGTA		PEG-DMG (1.5%) 또는 MC3 (50 mol %) DSPC (10 mol %) 콜레스테롤 (38.5% mol %) PEG-DMG (1.5%)

[0258]

	RRSVACVQLDQGQDVE VDEAACAALVRPEASV PCLIADCTYRWHVGTW MECSVSCGDGIQRRRD TCLGPQAQAPVPADFC QHLPKPVTVRGCWAGP CVGQGTPSLVPHEEAA APGRTTATPAGASLEW SQARGLLFSPAPQPRRL LPGPQENSVQSSACGR QHLEPTGTIDMRGPGQ ADCAVAIGRPLGEVVT LRVLESSLNCSAGDML LLWGRLTWRKMCRKL LDMTFSSKTNTLVVRQ RCGRPGGGVLLRYGSQ LAPETFYRECDMQLFG PWGEIVSPSLSPATSNA GGCRLFINVAPHARIAI HALATNMGAGTEGAN ASYILIRDTHSLRTTAFH GQQVLYWESESSQAEM EFSEGFLKAQASLRGQ YWTLQSWVPEMQDPQ SWKGKEGT		
FOXP3	MPNPRPGKPSAPSLALG PSPGASPSWRAAPKAS DLLGARGPGGTFQGRD LRGGAHASSSSLNPMPP SQLQLPTLPLVMVAPSG ARLGPLPHLQALLQDR PHFMHQLSTVDAHART PVLQVHPLESPAMISLT PPTTATGVFSLKARPGL PPGINVASLEWVSREPA LLCTFPNPSAPRKDSTL SAVPQSSYPLLANGVC KWPGCEKVFEEPEDFL KHCQADHLLDEKGRA QCLLQREMVQSLEQQL VLEKEKLSAMQAHLAG KMALTKASSVASSDKG SCCIVAAGSQGPVVPA WSGPREAPDSLFAVRR HLWGSHGNSTFPEFLH NMDYFKFHNMRPPFTY ATLIRWAILEAPEKQRT LNEIYHWFTRMFAFFR NHPATWKNAIRHNLSL HKCFVRVESEKGAVWT	면역 세포	 (50 mol %) DSPC (10 mol %) 베타-시토스테롤(28.5% mol %) 콜레스테롤 (10 mol %) PEG DMG (1.5 mol %)

[0259]

	VDELEFRKKRSQRPSRC SNPTPGP		
IL-10	SPGQGTQSENSCTHFPG NLPNMLRDLRDAFSRV KTFFQMKDQLDNLLLK ESLLEDFKGYLGCQALS EMIQFYLEEVMPQAEN QDPDIKAHVNSLGENL KTLRLRLRRCHRFLPCE NKSKAVEQVKNAFNKL QEKGIYKAMSEFDIFIN YIEAYMTMKIRN	면역 세포	 (50 mol %) DSPC (10 mol %) 베타-시토스테롤(28.5% mol %) 콜레스테롤 (10 mol %) PEG DMG (1.5 mol %)
IL-2	APTSSSTKKTQLQLEHL LLDLQMILNGINNYKNP KLTRMLTFKFYMPKKA TELKHLQCLEEELKPLE EVLNLAQSKNFHLRPR DLISNINVIVLELKGSET TFMCEYADETATIVEFL NRWITFCQSIISTLT	면역 세포	 (50 mol %) DSPC (10 mol %) 베타-시토스테롤(28.5% mol %) 콜레스테롤 (10 mol %) PEG DMG (1.5 mol %)

[0260]

[0261] 일부 실시형태에서, 발현 서열 중 적어도 하나는 치료용 단백질을 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 사이토카인, 예를 들어, IL-12p70, IL-15, IL-2, IL-18, IL-21, IFN-α, IFN- β, IL-10, TGF-

베타, IL-4 또는 IL-35, 또는 이의 기능성 단편을 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 면역 관문 저해제를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 효능제(예를 들어, TNFR 패밀리 구성원, 예컨대, CD137L, OX40L, ICOSL, LIGHT 또는 CD70)를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 키메라 항원 수용체를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 저해성 수용체 효능제(예를 들어, PDL1, PDL2, 갈렉틴-9, VISTA, B7H4 또는 MHCII) 또는 저해성 수용체(예를 들어, PD1, CTLA4, TIGIT, LAG3 또는 TIM3)를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 저해성 수용체 길항제를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 하나 이상의 TCR 쇄(알파 및 베타쇄 또는 감마 및 델타쇄)를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 분비된 T 세포 또는 면역 세포 인게이저(engager)(예를 들어, 이중특이적 항체, 예컨대, BiTE, 표적화, 예를 들어, CD3, CD137 또는 CD28 및 종양-발현된 단백질, 예를 들어, CD19, CD20 또는 BCMA 등)를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 전사 인자(예를 들어, *FOXP3*, *HELIOS*, *TOX1* 또는 *TOX2*)를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 면역억제 효소(예를 들어, IDO 또는 CD39/CD73)를 암호화한다. 일부 실시형태에서, 제1 또는 제2 발현 서열은 GvHD (예를 들어, 항-HLA-A2 CAR-Tregs)를 암호화한다.

[0262] 일부 실시형태에서, 제1 스페이서 및 제2 스페이서발현 서열은 T 세포 수용체(TCR)의 알파 및 베타쇄를 암호화한다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서발현 서열은 TCR의 감마 및 델타쇄를 암호화한다. 본 발명은 치료적 유효량의, TCR 알파쇄 및 TCR 베타쇄 또는 TCR 감마쇄 및 TCR 델타쇄를 암호화하는 원형 RNA 폴리뉴클레오타이드를 포함하는 조성물을 투여하는 단계를 포함하는 암을 앓고 있는 대상체를 치료하는 방법을 포함한다.

[0263] 일부 실시형태에서, 제1 스페이서 및 제2 스페이서발현 서열은 키메라 항원 수용체(CAR) 및 PD1 또는 PDL1의 길항제를 암호화한다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서발현 서열은 키메라 항원 수용체(CAR) 및 사이토카인을 암호화한다. 일부 실시형태에서, 사이토카인은 IL-12p70, IL-15, IL-2, IL-18, IL-21, IFN-α, IFN- β, IL-10, TGF-베타, IL-4 또는 IL-35, 또는 이의 기능성 단편이다. 본 발명은 치료적 유효량의, CAR 및 PD1 또는 PDL1의 길항제를 암호화하는 원형 RNA 폴리뉴클레오타이드를 포함하는 조성물을 투여하는 단계를 포함하는, 암을 앓고 있는 대상체를 치료하는 방법을 포함한다. 본 발명은 치료적 유효량의, CAR 및 사이토카인을 암호화하는 원형 RNA 폴리뉴클레오타이드를 포함하는 조성물을 투여하는 단계를 포함하는, 암을 앓고 있는 대상체를 치료하는 방법을 포함한다.

[0264] 일부 실시형태에서, 제1 스페이서 및 제2 스페이서발현 서열은 전사 인자 및 사이토카인을 암호화한다. 일부 실시형태에서, 전사 인자는 FOXP3, STAT5B 또는 HELIOS이고, 사이토카인은 IL10, IL12 또는 TGF 베타이다. 본 발명은 치료적 유효량의, 전사 인자, 예를 들어, FOXP3 및 사이토카인을 암호화하는 원형 RNA 폴리뉴클레오타이드를 포함하는 조성물을 투여하는 단계를 포함하는, 자가면역 장애를 앓고 있는 대상체를 치료하는 방법을 포함한다.

[0265] 일부 실시형태에서, 제1 스페이서 및 제2 스페이서발현 서열은 전사 인자 및 CAR을 암호화한다. 일부 실시형태에서, 전사 인자는 FOXP3, STAT5B 또는 HELIOS이다. 본 발명은 치료적 유효량의, 전사 인자, 예를 들어, FOXP3 및 CAR을 암호화하는 원형 RNA 폴리뉴클레오타이드를 포함하는 조성물을 투여하는 단계를 포함하는, 자가면역 장애를 앓고 있는 대상체를 치료하는 방법을 포함한다.

[0266] 일부 실시형태에서, 제1 스페이서 및 제2 스페이서발현 서열은 사이토카인 및 항원을 암호화한다. 일부 실시형태에서, 사이토카인은 IFNɣ다. 일부 실시형태에서, 항원은 신생항원이다. 본 발명은 치료적 유효량의, 사이토카인, 예를 들어, IFNɣ 및 종양 항원 또는 이의 단편을 암호화하는 원형 RNA 폴리뉴클레오타이드를 포함하는 조성물을 투여하는 단계를 포함하는, 암을 앓고 있는 대상체를 치료하는 방법을 포함한다.

[0267] 일부 실시형태에서, 제1 발현 서열은 제1 키메라 항원 수용체(CAR)를 암호화하고, 제2 발현 서열은 제2 CAR을 암호화한다. 일부 실시형태에서, 제1 CAR은 제1 항원에 특이적이고, 공자극성 도메인 및 세포내 신호전달 도메인을 함유하고, 제2 CAR은 제2 항원에 특이적이고, 공자극성 도메인 및 세포내 신호전달 도메인을 함유한다. 일부 실시형태에서, 다중 종양 항원을 표적화하는 CAR을 발현시키는 것은 이종 항원 발현을 갖는 종양에 대해 보다 효과적인 요법을 제공한다. 본 발명은 치료적 유효량의, 제1 CAR 및 제2 CAR을 암호화하는 원형 RNA 폴리뉴클레오타이드를 포함하는 조성물을 투여하는 단계를 포함하는, 암을 앓고 있는 대상체를 치료하는 방법을 포함한다.

[0268] 일부 실시형태에서, 제1 발현 서열은 제1 사이토카인을 암호화하고, 제2 발현 서열은 제2 사이토카인을 암호화한다. 일부 실시형태에서, 제1 스페이서 및 제2 스페이서사이토카인은 그룹 IL-10, TGFβ 및 IL-35이다. 일부

실시형태에서, 제1 스페이서 및 제2 스페이서사이토카인은 그룹 IFNɣ, IL-2, IL-7, IL-15 및 IL-18이다.

[0269] 일부 실시형태에서, 폴리뉴클레오타이드는 1종 초과의 유전자에 의해서 암호화되는 소단위로 구성된 단백질을 암호화한다. 예를 들어, 단백질은 이종이량체일 수 있고, 여기서 단백질의 각각의 쇄 또는 소단위는 별개의 유전자에 의해서 암호화된다. 1종 초과의 circRNA 분자가 전달 비히클로 전달되고, 각각의 circRNA가 단백질의 개별 소단위를 암호화한다. 대안적으로, 단일 circRNA는 1종 초과의 소단위를 암호화하도록 조작될 수 있다. 특정 실시형태에서, 각각의 소단위를 암호화하는 개별 circRNA 분자는 개별 전달 비히클로 투여될 수 있다.

[0270] **3.1 사이토카인**

[0271] IL-2, IL-7, IL-10, IL-12, IL-15, IL-18, IL-27β, IFNɣ 및/또는 TGFβ1의 설명 및/또는 아미노산 서열이 본 명세서 및 www.uniprot.org 데이터에 수탁 번호: P60568(IL-2), P29459(IL-12A), P29460(IL-12B), P13232(IL-7), P22301(IL-10), P40933(IL-15), Q14116(IL-18), Q14213(IL-27β), P01579(IFNɣ) 및/또는 P01137(TGFβ1)에 제공된다.

[0272] **3.2 PD-1 및 PD-L1 길항제**

[0273] 일부 실시형태에서, PD-1 저해제는 펨브롤리주맙, 피딜리주맙 또는 니볼루맙이다. 일부 실시형태에서, 니볼루맙은 국제 특허 공개 제WO2006/121168호에 기재되어 있다. 일부 실시형태에서, 펨브롤리주맙은 국제 특허 공개 제WO2009/114335호에 기재되어 있다. 일부 실시형태에서, 피딜리주맙은 국제 특허 공개 제WO2009/101611호에 기재되어 있다. 추가적인 항-PD1 항체는 미국 특허 제8,609,089호, 미국 특허 공개 제US 2010028330호 및 제US 20120114649호 및 국제 특허 공개 제WO2010/027827호 및 제WO2011/066342호에 기재되어 있다.

[0274] 일부 실시형태에서, PD-L1 저해제는 아테졸리주맙, 아벨루맙, 더발루맙, BMS-936559 또는 CK-301이다.

[0275] PD-1 및/또는 PD-L1 항체의 중쇄 및 경쇄의 설명 및/또는 아미노산 서열은 본 명세서 및 www.drugbank.ca 데이터베이스에 수탁 번호: DB09037(펨브롤리주맙), DB09035(니볼루맙), DB15383(피딜리주맙), DB11595(아테졸리주맙), DB11945(아벨루맙) 및 DB11714(더발루맙)에 제공된다.

[0276] **3.3 키메라 항원 수용체**

[0277] 키메라 항원 수용체(CAR 또는 CAR-T)는 유전자 조작된 수용체이다. 이들 조작된 수용체는 본 명세서에 기재된 바와 같은 원형 RNA를 통해서 T 세포를 포함하는 면역 세포로 용이하게 삽입되고 이에 의해 발현될 수 있다. CAR로, 단일 수용체는 특정 항원을 인식하고, 그 항원에 결합될 때, 면역 세포를 활성화시켜 그 항원을 보유하는 세포를 공격하고 파괴하도록 프로그래밍될 수 있다. 이들 항원이 종양 세포 상에 존재하는 경우, CAR을 발현하는 면역 세포는 종양 세포를 표적하고 사멸시킬 수 있다. 일부 실시형태에서, 폴리뉴클레오타이드에 의해서 암호화된 CAR은 (i) 표적 항원에 특이적으로 결합하는 항원-결합 분자, (ii) 힌지 도메인, 막관통 도메인 및 세포내 도메인 및 (iii) 활성화 도메인을 포함한다.

[0278] 일부 실시형태에서, 본 개시내용에 따른 CAR의 배향은 공자극성 도메인 및 활성화 도메인과 나란하게 항원 결합 도메인(예컨대, scFv)을 포함한다. 공자극성 도메인은 세포외 부분, 막관통 부분 및 세포내 부분 중 하나 이상을 포함할 수 있다. 다른 실시형태에서, 다수의 공자극성 도메인이 나란하게 사용될 수 있다.

[0279] ***항원 결합 도메인***

[0280] CAR은 표적화된 항원과 상호작용하는 항원 결합 분자를 혼입함으로써 항원(예컨대, 세포-표면 항원)에 결합하도록 조작될 수 있다. 일부 실시형태에서, 항원 결합 분자는 이의 항체 단편, 예를 들어, 1개 이상의 단일 쇄 항체 단편(scFv)이다. scFv는 함께 연결된 항체의 중쇄 및 경쇄의 가변 영역을 갖는 단일 쇄 항체 단편이다. 예를 들어, 미국 특허 제7,741,465호 및 제6,319,494호뿐만 아니라 문헌[Eshhar *et al.*, Cancer Immunol Immunotherapy (1997) 45: 131-136] 참조. scFv는 표적 항원과 특이적으로 상호작용하는 모 항체의 능력을 유지한다. scFv는 키메라 항원 수용체에 유용한데, 그 이유는 이들이 다른 CAR 성분과 함께 단일 쇄의 부분으로서 발현되도록 조작될 수 있기 때문이다. 또한 문헌[Krause *et al.*, J. Exp. Med., Volume 188, No. 4, 1998 (619-626); Finney *et al.*, Journal of Immunology, 1998, 161 : 2791-2797] 참조. 항원 결합 분자는 전형적으로 관심 항원을 인식하게 이에 결합할 수 있도록 CAR의 세포외 부분 내에 전형적으로 함유된다는 것을 인식할 것이다. 이중특이적 및 다중특이적 CAR이, 1종 초과의 관심 표적에 대한 특이성과 함께 본 발명의 범주에 포함되는 것으로 고려된다.

[0281] 일부 실시형태에서, 항원 결합 분자는 단일 쇄를 포함하고, 여기서 중쇄 가변 영역 및 경쇄 가변 영역은 링커에

의해서 연결된다. 일부 실시형태에서, VH는 링커의 N 말단에 위치하고, VL은 링커의 C 말단에 위치한다. 다른 실시형태에서, VL는 링커의 N 말단에 위치하고, VH는 링커의 C 말단에 위치한다. 일부 실시형태에서, 링커는 적어도 약 5, 적어도 약 8, 적어도 약 10, 적어도 약 13, 적어도 약 15, 적어도 약 18, 적어도 약 20, 적어도 약 25, 적어도 약 30, 적어도 약 35, 적어도 약 40, 적어도 약 45, 적어도 약 50, 적어도 약 60, 적어도 약 70, 적어도 약 80, 적어도 약 90 또는 적어도 약 100개의 아미노산을 포함한다.

[0282] 일부 실시형태에서, 항원 결합 분자는 나노바디를 포함한다. 일부 실시형태에서, 항원 결합 분자는 DARPin을 포함한다. 일부 실시형태에서, 항원 결합 분자는 표적 단백질에 특이적으로 결합할 수 있는 안티칼린 또는 다른 합성 단백질을 포함한다.

[0283] 일부 실시형태에서, CAR은 하기 군으로부터 선택된 항원에 특이적인 항원 결합 도메인을 포함한다: CD19, CD123, CD22, CD30, CD171, CS-1, C-타입 렉틴-유사 분자-1, CD33, 표피 성장 인자 수용체 변이체 III(epidermal growth factor receptor variant III: EGFRvIII), 강글리오사이드 G2(GD2), 글리오사이드 GD3, TNF 수용체 패밀리 구성원, B 세포 성숙 항원(B cell maturation antigen: BCMA), Tn 항원((Tn Ag) 또는 (GaINAca-Ser/Thr)), 전립선-특이적 막 항원(prostate-specific membrane antigen: PSMA), 수용체 티로신 키나제-유사 희귀 수용체 1(Receptor tyrosine kinase-like orphan receptor 1: ROR1), Fms-유사 티로신 키나제 3(Fms-Like Tyrosine Kinase 3: FLT3), 종양-연관 당단백질 72(Tumor-associated glycoprotein 72: TAG72), CD38, CD44v6, 암배아 항원(Carcinoembryonic antigen: CEA), 상피 세포 접착 분자(Epithelial cell adhesion molecule: EPCAM), B7H3(CD276), KIT(CD117), 인터류킨-13 수용체 소단위 알파-2, 메소텔린, 인터류킨 11 수용체 알파(Interleukin 11 receptor alpha: IL-11Ra), 전립선 줄기세포 항원(prostate stem cell antigen: PSCA), 프로테아제 세린 21, 혈관 내피 성장 인자 수용체 2(vascular endothelial growth factor receptor 2: VEGFR2), 루이스(Y) 항원, CD24, 혈소판-유래 성장 인자 수용체 베타(Platelet-derived growth factor receptor beta: PDGFR-베타), 단계-특이적 배아 항원-4(Stage-specific embryonic antigen-4: SSEA-4), CD20, 엽산염 수용체 알파, HER2, HER3, 뮤신 1, 세포 표면 연관(MUC1), 표피 성장 인자 수용체(epidermal growth factor receptor: EGFR), 신경 세포 접착 분자(neural cell adhesion molecule: NCAM), 전립선, 전립선 산 포스파타제(fibroblast activation protein alpha: PAP), 돌연변이된 연장 인자 2(elongation factor 2 mutated: ELF2M), 에프린 B2, 섬유모세포 활성화 단백질 알파(FAP), 인슐린-유사 성장 인자 1 수용체(I insulin-like growth factor 1 receptor: GF-I 수용체), 탄산탈수효소 IX(carbonic anhydrase IX: CAIX), 프로테아솜(Prosome, Macropain) 소단위, 베타 타입, 9(LMP2), 당단백질 100(glycoprotein 100: gp100), 중단점 클러스터 영역(BCR)과 아벨슨 뮤린 백혈병 바이러스 암유전자 동족체 1(Abl)로 이루어진 암유전자 융합 단백질(bcr-abl), 티로시나제, 에프린 타입-A 수용체 2(EphA2), 푸코실 GM1, 시알릴 루이스 접착 분자(sialyl Lewis adhesion molecule: sLe), 강글리오사이드 GM3, 트랜스글루타미나제 5(transglutaminase 5: TGS5), 중합체량-흑색종-연관 항원(high molecular weight-melanoma-associated antigen: HMWMAA), o-아세틸-GD2 강글리오사이드(o-acetyl-GD2 ganglioside: OAcGD2), 엽산염 수용체 베타, 종양 내피 마커 1(tumor endothelial marker 1: TEM1/CD248), 종양 내피 마커 7-관련(tumor endothelial marker 7-related: TEM7R), 클라우딘 6(claudin 6: CLDN6), 갑상선 자극 호르몬 수용체(thyroid stimulating hormone receptor: TSHR), G 단백질-커플드 수용체 클래스 C 그룹 5, 구성원 D(G protein-coupled receptor class C group 5, member D: GPRC5D), 염색체 X 오픈 리딩 프레임 61(chromosome X open reading frame 61: CXORF61), CD97, CD179a, 역형성 림프종 키나제(anaplastic lymphoma kinase: ALK), 폴리시알산, 태반-특이적 1(PLAC1), globoH 당세라마이드의 육당류 부분(GloboH), 유선 분화 항원(mammary gland differentiation antigen: NY-BR-1), 우로플라킨 2(UPK2), A형 간염 바이러스 세포 수용체 1(Hepatitis A virus cellular receptor 1: HAVCR1), 아드레날린수용체 베타 3(adrenoceptor beta 3: ADRB3), 파넥신 3(pannexin 3: PANX3), G 단백질-커플링된 수용체 20(G protein-coupled receptor 20: GPR20), 림프구 항원 6 복합체, 유전자좌 K 9(lymphocyte antigen 6 complex, locus K 9: LY6K), 후각 수용체 51E2(Olfactory receptor 51E2: OR51E2), TCR 감마 대체 리딩 프레임 단백질(TCR Gamma Alternate Reading Frame Protein: TARP), 빌름스 종양 단백질(WT1), 암/고환 항원 1(Cancer/testis antigen 1: NY-ESO-1), 암/고환 항원 2(LAGE-1a), MAGE 패밀리 구성원(MAGE-A1, MAGE-A3 및 MAGE-A4 포함), 염색체 12p에 위치된 ETS 전좌-변이체 유전자 6(ETS translocation-variant gene 6, located on chromosome 12p: ETV6-AML), 정자 단백질 17(sperm protein 17: SPA17), X 항원 패밀리, 구성원 1A(X Antigen Family, Member 1A: XAGE1), 앤지오포이에틴-결합 세포 표면 수용체 2(Tie 2), 흑색종 암 고환 항원-1(MAD-CT-1), 흑색종 암 고환 항원-2(MAD-CT-2), Fos-관련 항원 1, 종양 단백질 p53(p53), p53 돌연변이체, 프로스테인, 생존, 텔로머라제, 전립선 암종 종양 항원-1, T 세포 1에 의해서 인식되는 흑색종 항원, 래트 육종(Ras) 돌연변이체, 인간 텔로머라제 역전사효소(hTERT), 육종 전좌 파단점, 아폽토시스의 흑색종 저해제(melanoma inhibitor of apoptosis:

ML-IAP), ERG(막관통 프로테아제, 세린 2(TMPRSS2) ETS 융합 유전자), N-아세틸 글루코사민일-트랜스퍼라제 V(NA17), 짝지워진 박스 단백질 Pax-3(PAX3), 안드로겐 수용체, 사이클린 B1, v-myc 조류 골수세포종 바이러스 암유전자 신경모세포종 유래 동족체(v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog: MYCN), Ras 동족체 패밀리 구성원 C(RhoC), 티로시나제-관련 단백질 2(TRP-2), 사이토크롬 P450 1B1(Cytochrome P450 1B1: CYP1B1), CCCTC-결합 인자(아연 핑거 단백질)-유사, T 세포 3에 의해서 인식되는 편평 세포 암종 항원(Squamous Cell Carcinoma Antigen Recognized By T Cells 3: SART3), 짝지워진 박스 단백질 Pax-5(PAX5), 프로아크로신 결합 단백질 sp32(OY-TES1), 림프구-특이적 단백질 티로신 키나제(lymphocyte-specific protein tyrosine kinase: LCK), A 키나제 고정 단백질 4 (A kinase anchor protein 4: AKAP-4), 활막 육종, X 파단점 2(SSX2), 최종 당산화물에 대한 수용체(Receptor for Advanced Glycation Endproducts: RAGE-1), 신장 유비쿼터스 1(renal ubiquitous 1: RU1), 신장 유비쿼터스 2(RU2), 레구마인(legumain), 인간 유두종 바이러스 E6(HPV E6), 인간 유두종 바이러스 E7(HPV E7), 장 카복실 에스터라제, 돌연변이된 열 충격 단백질 70-2(mut hsp70-2), CD79a, CD79b, CD72, 백혈구-연관 면역글로불린-유사 수용체 1(LAIR1), IgA 수용체의 Fc 단편(FCAR 또는 CD89), 백혈구 면역글로불린-유사 수용체 서브패밀리 A 구성원 2(Leukocyte immunoglobulin-like receptor subfamily A member 2: LILRA2), CD300 분자-유사 패밀리 구성원 f(CD300LF), C-타입 렉틴 도메인 패밀리 12 구성원 A(C-type lectin domain family 12 member A: CLEC12A), 골수 기질 세포 항원 2(bone marrow stromal cell antigen 2: BST2), EGF-유사 모듈-함유 뮤신-유사 호르몬 수용체-유사 2(EMR2), 림프구 항원 75(LY75), 글리피칸-3(GPC3), Fc 수용체-유사 5(FCRL5), MUC16, 5T4, 8H9, αvβΘ 인테그린, αvβ6 인테그린, 알파태아단백질(AFP), B7-H6, ca-125, CA9, CD44, CD44v7/8, CD52, E-카드헤린, EMA(상피 막 항원), 상피 당단백질-2(EGP-2), 상피 당단백질-40(EGP-40), ErbB4, 상피 종양 항원(ETA), 엽산염 결합 단백질(FBP), 키나제 삽입부 도메인 수용체(KDR), k-경쇄, L1 세포 접착 분자, MUC18, NKG2D, 암배아 항원(h5T4), 종양/고환-항원 1B, GAGE, GAGE-1, BAGE, SCP-1, CTZ9, SAGE, CAGE, CT10, MART-1, 면역글로불린 람다-유사 폴리펩타이드 1(IGLL1), 간염 B 표면 항원 결합 단백질(HBsAg), 바이러스 캡시드 항원(VCA), 초기 항원(EA), EBV 핵 항원(EBNA), HHV-6 p41 초기 항원, HHV-6B U94 잠복항원, HHV-6B p98 후기 항원, 사이토메갈로바이러스(CMV) 항원, 큰 T 항원, 작은 T 항원, 아데노바이러스 항원, 호흡기 세포융합 바이러스(respiratory syncytial virus: RSV) 항원, 해마토글루티닌(HA), 뉴라미니다제(NA), 파라인플루엔자 타입 1 항원, 파리인플루엔자 타입 2 항원, 파라인플루엔자 타입 3 항원, 파라인플루엔자 타입 4 항원, 인간 메타뉴모바이러스(HMPV) 항원, C형 간염 바이러스(HCV) 코어 항원, HIV p24 항원, 인간 T-세포 림프친화성 바이러스(HTLV-1) 항원, 머켈 세포 폴리오마 바이러스 작은 T 항원, 머켈 세포 폴리오마 바이러스 큰 T 항원 및 카포시 육종-연관 헤르페스 바이러스(KSHV) 용해성 핵 항원 및 KSHV 잠복 핵 항원. 일부 실시형태에서, 항원 결합 도메인은 서열번호 321 및/또는 322를 포함한다.

[0284] ***힌지/스페이서 도메인***

[0285] 일부 실시형태에서, 본 개시내용의 CAR은 힌지 또는 스페이서 도메인을 포함한다. 일부 실시형태에서, 힌지/스페이서 도메인은 절단된 힌지/스페이서 도메인(THD)을 포함할 수 있다. THD 도메인은 완전한 힌지/스페이서 도메인("CHD")의 절단된 버전이다. 일부 실시형태에서, 세포외 도메인은 ErbB2, 글리코포린 A(GpA), CD2, CD3 델타, CD3 엡실론, CD3 감마, CD4, CD7, CD8a, CD8[T CD1 1a(IT GAL), CD1 1b(IT GAM), CD1 1c(ITGAX), CD1 1d(IT GAD), CD18(ITGB2), CD19(B4), CD27(TNFRSF7), CD28, CD28T, CD29(ITGB1), CD30(TNFRSF8), CD40(TNFRSF5), CD48(SLAMF2), CD49a(ITGA1), CD49d(ITGA4), CD49f(ITGA6), CD66a(CEACAM1), CD66b(CEACAM8), CD66c(CEACAM6), CD66d(CEACAM3), CD66e(CEACAM5), CD69(CLEC2), CD79A(B-세포 항원 수용체 복합체-연관 알파쇄), CD79B(B-세포 항원 수용체 복합체-연관 베타쇄), CD84(SLAMF5), CD96(Tactile), CD100(SEMA4D), CD103(ITGAE), CD134(OX40), CD137(4-1BB), CD150(SLAMF1), CD158A(KIR2DL1), CD158B1(KIR2DL2), CD158B2(KIR2DL3), CD158C(KIR3DP1), CD158D(KIRDL4), CD158F1(KIR2DL5A), CD158F2(KIR2DL5B), CD158K(KIR3DL2), CD160(BY55), CD162(SELPLG), CD226(DNAM1), CD229(SLAMF3), CD244(SLAMF4), CD247(CD3-제타), CD258(LIGHT), CD268(BAFFR), CD270(TNFSF14), CD272(BTLA), CD276(B7-H3), CD279(PD-1), CD314(NKG2D), CD319(SLAMF7), CD335(NK-p46), CD336(NK-p44), CD337(NK-p30), CD352(SLAMF6), CD353(SLAMF8), CD355(CRTAM), CD357(TNFRSF18), 유도성 T 세포 공자극인자(ICOS), LFA-1(CD1 1a/CD18), NKG2C, DAP-10, ICAM-1, NKp80(KLRF1), IL-2R 베타, IL-2R 감마, IL-7R 알파, LFA-1, SLAMF9, LAT, GADS(GrpL), SLP-76(LCP2), PAG1/CBP, CD83 리간드, Fc 감마 수용체, MHC 클래스 1 분자, MHC 클래스 2 분자, TNF 수용체 단백질, 면역글로불린 단백질, 사이토카인 수용체, 인테그린, 활성화 NK 세포 수용체, 톨 리간드 수용체 및 이들의 단편 또는 조합물로부터 기원되거나 유래된다(예를 들어, 이들 모두 또는 이들의 단편을 포함한다). 힌지 또는 스페이서 도메인은 자연 공급원으로부터 또는 합성 공급원으로부터 유래될 수 있다.

[0286] 일부 실시형태에서, 힌지 또는 스페이서 도메인은 항원 결합 분자(예를 들어, scFv)와 막관통 도메인 사이에 위치된다. 이러한 배향에서, 힌지/스페이서 도메인은 항원 결합 분자와 CAR이 발현되는 세포 막의 표면 사이의 거리를 제공한다. 일부 실시형태에서, 힌지 또는 스페이서 도메인은 면역글로불린으로부터 기원되거나 이로부터 유래된다. 일부 실시형태에서, 힌지 또는 스페이서 도메인은 IgG1, IgG2, IgG3, IgG4, IgA, IgD, IgE 및 IgM 또는 이들의 단편의 힌지/스페이서 영역으로부터 선택된다. 일부 실시형태에서, 힌지 또는 스페이서 도메인은 CD8 알파의 힌지/스페이서 영역을 포함하거나 이로부터 기원되거나 이로부터 유래된다. 일부 실시형태에서, 힌지 또는 스페이서 도메인은 CD28의 힌지/스페이서 영역을 포함하거나 이로부터 기원되거나 이로부터 유래된다. 일부 실시형태에서, 힌지 또는 스페이서 도메인은 CD8 알파의 힌지/스페이서 영역의 단편 또는 CD28의 힌지/스페이서 영역의 단편을 포함하고, 여기서 단편은 전체 힌지/스페이서 영역보다 더 작은 임의의 것이다. 일부 실시형태에서, CD8 알파 힌지/스페이서 영역의 단편 또는 CD28 힌지/스페이서 영역의 단편은 CD8 알파 힌지/스페이서 영역 또는 CD28 힌지/스페이서 영역의 N-말단 또는 C-말단 또는 둘 다에서 적어도 1, 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19 또는 적어도 20개의 아미노산을 제외한 아미노산 서열을 포함한다.

[0287] ***막관통 도메인***

[0288] 본 개시내용의 CAR은 막관통 도메인 및/또는 세포내 신호전달 도메인을 추가로 포함할 수 있다. 막관통 도메인은 CAR의 세포외 도메인에 융합되도록 설계될 수 있다. 그것은 CAR의 세포내 도메인에 유사하게 융합될 수 있다. 일부 실시형태에서, CAR 내의 도메인 중 하나와 자연적으로 연관된 막관통 도메인이 사용된다. 일부 예에서, 막관통 도메인은 수용체 복합체의 다른 구성원과의 상호작용을 최소화하기 위해 동일하거나 상이한 표면 막 단백질의 막관통 도메인에 대한 이러한 도메인의 결합을 회피하도록 (예를 들어, 아미노산 치환에 의해) 선택되거나 변형될 수 있다. 막관통 도메인은 자연 공급원으로부터 또는 합성 공급원으로부터 유래될 수 있다. 공급원이 자연적인 경우, 도메인은 임의의 막-결합된 또는 막관통 단백질로부터 유래될 수 있다.

[0289] 막관통 영역은 수용체 티로신 키나제(예를 들어, ErbB2), 글리코포린 A(GpA), 4-1BB/CD137, 활성화 NK 세포 수용체, 면역글로불린 단백질, B7-H3, BAFFR, BFAME(SEAMF8), BTEA, CD100(SEMA4D), CD103, CD160(BY55), CD 18, CD 19, CD 19a, CD2, CD247, CD27, CD276(B7-H3), CD28, CD29, CD3 델타, CD3 엡실론, CD3 감마, CD30, CD4, CD40, CD49a, CD49D, CD49f, CD69, CD7, CD84, CD8알파, CD8베타, CD96(Tactile), CD1 la, CD1 lb, CD1 lc, CD1 Id, CDS, CEACAM1, CRT AM, 사이토카인 수용체, DAP-10, DNAM1(CD226), Fc 감마 수용체, GADS, GITR, HVEM(EIGHTR), IA4, ICAM-1, ICAM-1, Ig 알파(CD79a), IE-2R 베타, IE-2R 감마, IE-7R 알파, 유도성 T 세포 공자극인자(ICOS), 인테그린, ITGA4, ITGA4, ITGA6, IT GAD, ITGAE, ITGAE, IT GAM, ITGAX, ITGB2, ITGB7, ITGB1, KIRDS2, EAT, LFA-1, LFA-1, CD83과 특이적으로 결합하는 리간드, LIGHT, LIGHT, LTBR, Ly9(CD229), 림프구 기능-연관 항원-1(LFA-1; CD1-la/CD18), MHC 클래스 1 분자, NKG2C, NKG2D, NKp30, NKp44, NKp46, NKp80(KLRF1), OX-40, PAG/Cbp, 예정사-1(PD-1), PSGL1, SELPLG(CD162), 신호전달 림프구성 활성화 분자(SLAM 단백질), SLAM(SLAMF1; CD150; IPO-3), SLAMF4(CD244; 2B4), SLAMF6(NTB-A; Ly108), SLAMF7, SLP-76, TNF 수용체 단백질, TNFR2, TNFSF14, 톨 리간드 수용체, TRANCE/RANKL, VLA1 또는 VLA-6 또는 단편, 이들의 절단부 또는 조합물로부터 유래될 수 있다(즉, 이들을 포함할 수 있다).

[0290] 일부 실시형태에서, 적합한 세포내 신호전달 도메인은 활성화 대식세포/골수 세포 수용체 CSFR1, MYD88, CD14, TIE2, TLR4, CR3, CD64, TREM2, DAP10, DAP12, CD169, DECTIN1, CD206, CD47, CD163, CD36, MARCO, TIM4, MERTK, F4/80, CD91, C1QR, LOX-1, CD68, SRA, BAI-1, ABCA7, CD36, CD31, 락토페린 또는 이들의 단편, 절단부 또는 조합물을 포함하지만 이들로 제한되지 않는다.

[0291] 일부 실시형태에서, 수용체 티로신 키나제는 인슐린 수용체(InsR), 인슐린-유사 성장 인자 I 수용체(IGF1R), 인슐린 수용체-관련 수용체(IRR), 혈소판 유래 성장 인자 수용체 알파(PDGFRa), 혈소판 유래 성장 인자 수용체 베타(PDGFRfi). KIT 원종양 유전자 수용체 티로신 키나제(Kit), 집락 자극 인자 1 수용체(CSFR), fms 관련 티로신 키나제 3(FLT3), fms 관련 티로신 키나제 1(VEGFR-1), 키나제 삽입부 도메인 수용체(VEGFR-2), fms 관련 티로신 키나제 4(VEGFR-3), 섬유모세포 성장 인자 수용체 1(FGFR1), 섬유모세포 성장 인자 수용체 2(FGFR2), 섬유모세포 성장 인자 수용체 3(FGFR3), 섬유모세포 성장 인자 수용체 4(FGFR4), 단백질 티로신 키나제 7(CCK4), 신경영양성 수용체 티로신 키나제 1(trkA), 신경영양성 수용체 티로신 키나제 2(trkB), 신경영양성 수용체 티로신 키나제 3(trkC), 수용체 티로신 키나제 유사 희귀 수용체 1(ROR1), 수용체 티로신 키나제 유사 희귀 수용체 2(ROR2), 근육 연관 수용체 티로신 키나제(MuSK), MET 원종양 유전자, 수용체 티로신 키나제(MET), 대식세포 자극 1 수용체(Ron), AXL 수용체 티로신 키나제(Ax1), TYRO3 단백질 티로신 키나제(Tyro3), MER 원종양 유전자,

티로신 키나제(Mer), 면역글로불린 유사 및 EGF 유사 도메인 1을 갖는 티로신 키나제(TIE1), TEK 수용체 티로신 키나제(TIE2), EPH 수용체 A1(EphA1), EPH 수용체 A2(EphA2),(EPH 수용체 A3) EphA3, EPH 수용체 A4(EphA4), EPH 수용체 A5(EphA5), EPH 수용체 A6(EphA6), EPH 수용체 A7(EphA7), EPH 수용체 A8(EphA8), EPH 수용체 A10(EphA10), EPH 수용체 B1(EphB1), EPH 수용체 B2(EphB2), EPH 수용체 B3(EphB3), EPH 수용체 B4(EphB4), EPH 수용체 B6(EphB6), ret 원종양 유전자(Ret), 수용체-유사 티로신 키나제(RYK), 디스코이딘 도메인 수용체 티로신 키나제 1(DDR1), 디스코이딘 도메인 수용체 티로신 키나제 2(DDR2), c-ros 암유전자 1, 수용체 티로신 키나제(ROS), 아폽토시스 연관 티로신 키나제(Lmr1), lemur 티로신 키나제 2(Lmr2), 레무(lemur) 티로신 키나제 3(Lmr3), 백혈구 수용체 티로신 키나제(LTK), ALK 수용체 티로신 키나제(ALK) 또는 세린/트레오닌/티로신 키나제 1(STYK1)로부터 유래될 수 있다(이들을 포함할 수 있다).

[0292] *공자극성 도메인*

[0293] 특정 실시형태에서, CAR은 공자극성 도메인을 포함한다. 일부 실시형태에서, 공자극성 도메인은 4-1BB(CD137), CD28 또는 둘 다 및/또는 세포내 T 세포 신호전달 도메인을 포함한다. 바람직한 실시형태에서, 공자극성 도메인은 인간 CD28, 인간 4-1BB 또는 둘 다이고, 세포내 T 세포 신호전달 도메인은 인간 CD3 제타(ζ)이다. 4-1BB, CD28, CD3 제타 또는 이들 중 임의의 것은 각각 전체 4-1BB, CD28 또는 CD3 제타보다 작은 것을 포함할 수 있다. 키메라 항원 수용체는 이의 효력을 증가시키기 위해서 공자극성(신호전달) 도메인을 혼입할 수 있다. 미국 특허 제7,741,465호 및 제6,319,494호뿐만 아니라 상기 문헌[Krause *et al.* and Finney *et al.*], [Song *et al.*, Blood 119:696-706 (2012)]; [Kalos *et al.*, Sci Transl. Med. 3:95 (2011)]; [Porter *et al.*, N. Engl. J. Med. 365:725-33 (2011)] 및 [Gross *et al.*, Amur. Rev. Pharmacol. Toxicol. 56:59-83 (2016)] 참조.

[0294] 일부 실시형태에서, 공자극성 도메인은 서열번호 318 또는 320의 아미노산 서열을 포함한다.

[0295] *세포내 신호전달 도메인*

[0296] 본 명세서에 개시된 조작된 T 세포의 세포내(신호전달) 도메인은 활성화 도메인에 신호전달을 제공할 수 있고, 이것은 그 다음 면역 세포의 정상 효과기 기능 중 적어도 하나를 활성화시킨다. T 세포의 효과기 기능은 예를 들어, 사이토카인의 분비를 포함하는 세포용해 활성 또는 헬퍼 활성일 수 있다.

[0297] 일부 실시형태에서, 적합한 세포내 신호전달 도메인은 4-1BB/CD137, 활성화 NK 세포 수용체, 면역글로불린 단백질, B7-H3, BAFFR, BLAME(SLAMF8), BTLA, CD100(SEMA4D), CD103, CD160(BY55), CD18, CD19, CD 19a, CD2, CD247, CD27, CD276(B7-H3), CD28, CD29, CD3 델타, CD3 엡실론, CD3 감마, CD30, CD4, CD40, CD49a, CD49D, CD49f, CD69, CD7, CD84, CD8알파, CD8베타, CD96(Tactile), CD1 la, CD1 lb, CD1 lc, CD1 Id, CDS, CEACAM1, CRT AM, 사이토카인 수용체, DAP-10, DNAM1(CD226), Fc 감마 수용체, GADS, GITR, HVEM(LIGHTR), IA4, ICAM-1, ICAM-1, Ig 알파(CD79a), IL-2R 베타, IL-2R 감마, IL-7R 알파, 유도성 T 세포 공자극인자(ICOS), 인테그린, ITGA4, ITGA4, ITGA6, IT GAD, ITGAE, ITGAL, IT GAM, ITGAX, ITGB2, ITGB7, ITGB1, KIRDS2, LAT, LFA-1, LFA-1, CD83와 특이적으로 결합하는 리간드, LIGHT, LIGHT, LTBR, Ly9(CD229), Ly108), 림프구 기능-연관 항원-1(LFA-1; CDl-la/CD18), MHC 클래스 1 분자, NKG2C, NKG2D, NKp30, NKp44, NKp46, NKp80(KLRF1), OX-40, PAG/Cbp, 예정사-1(PD-1), PSGL1, SELPLG(CD162), 신호전달 림프구성 활성화 분자(SLAM 단백질), SLAM(SLAMF1; CD150; IPO-3), SLAMF4(CD244; 2B4), SLAMF6(NTB-A, SLAMF7, SLP-76, TNF 수용체 단백질, TNFR2, TNFSF14, 톨 리간드 수용체, TRANCE/RANKL, VLA1 또는 VLA-6 또는 이들의 단편, 절단부 또는 조합물을 포함하지만 이들로 제한되지 않는다(예를 들어, 이들을 포함한다).

[0298] CD3은 천연 T 세포 상의 T 세포 수용체의 요소이고, CAR 내의 중요한 세포내 활성화 요소인 것으로 밝혀져 있다. 일부 실시형태에서, CD3은 CD3 제타이다. 일부 실시형태에서, 활성화 도메인은 서열번호 319의 폴리펩타이드 서열과 적어도 약 60%, 적어도 약 65%, 적어도 약 70%, 적어도 약 75%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99% 또는 약 100% 동일한 아미노산 서열을 포함한다.

[0299] **3.4 T 세포 수용체**

[0300] TCR은 IMGT(International Immunogenetics) TCR 명명법을 사용하여 설명되고, TCR 서열의 IMGT 공개 데이터베이스와 연관된다. 천연 알파-베타 이종이량체 TCR은 알파쇄 및 베타쇄를 갖는다. 광범위하게, 각각의 쇄는 가변, 연결 및 불변 영역을 포함할 수 있으며, 베타쇄는 또한 보통 가변 영역과 연결 영역 사이에 짧은 다양성 영역을 포함하지만, 이러한 다양성 영역은 종종 연결 영역의 일부로 간주된다. 각각의 가변 영역은 프레임워크 서열에 내장된 3개의 CDR(상보성 결정 영역)을 포함할 수 있으며, 하나는 CDR3으로 명명된 초가변 영역이다. 프

레임워크, CDR1 및 CDR2 서열에 의해서 그리고 부분적으로 정의된 CDR3 서열에 의해 구별되는 여러 유형의 알파쇄 가변(Vα) 영역 및 여러 유형의 베타쇄 가변(Vβ) 영역이 존재한다. Vα 유형은 IMGT 명명법에서 고유한 TRAV 번호로 지칭된다. 따라서 "TRAV21"은 고유한 프레임워크 및 CDR1 및 CDR2 서열을 갖는 TCR Vα 영역, 및 TCR에서 TCR까지 보존되지만 TCR에서 TCR까지 달라지는 아미노산 서열을 또한 포함하는 아미노산 서열에 의해 부분적으로 정의되는 CDR3 서열을 정의한다. 동일한 방식으로 "TRBV5-1"은 고유한 프레임워크 및 CDR1 및 CDR2 서열을 갖지만 부분적으로만 정의된 CDR3 서열을 갖는 TCR Vβ 영역을 정의한다.

[0301] TCR의 결합 영역은 고유한 IMGT TRAJ 및 TRBJ 명명법에 의해 유사하게 정의되고, 불변 영역은 IMGT TRAC 및 TRBC 명명법에 의해 정의된다.

[0302] 베타쇄 다양성 영역은 IMGT 명명법에서 약어 TRBD로 지칭되며, 언급된 바와 같이 연결된 TRBD/TRBJ 영역은 종종 결합 영역으로 함께 간주된다.

[0303] IMGT 명명법에 의해 정의된 고유한 서열은 널리 알려져 있고, TCR 분야에서 일하는 사람들에게 접근 가능하다. 예를 들어, 이들은 IMGT 공개 데이터베이스에서 찾을 수 있다. 문헌[The "T cell Receptor Factsbook", (2001) LeFranc and LeFranc, Academic Press, ISBN 0-12-441352-8]은 또한 IMGT 명명법에 의해 정의된 서열을 개시하지만, 이의 출판일 및 이에 따른 시간 지연으로 인해, 그 내의 정보는 때때로 IMGT 데이터베이스를 참조하여 확인할 필요가 있다.

[0304] 천연 TCR은 이종이량체 αβ 또는 γδ 형태로 존재한다. 그러나, αα 또는 ββ 동종이량체로 이루어진 재조합 TCR은 이전에 펩타이드 MHC 분자에 결합하는 것으로 나타났다. 따라서, 본 발명의 TCR은 이종이량체 αβ TCR일 수 있거나, αα 또는 ββ 동종이량체 TCR일 수 있다.

[0305] 입양 요법에 사용하기 위해, αβ 이종이량체 TCR은 예를 들어, 세포질 및 막관통 도메인 둘 다를 갖는 전장 쇄로서 형질주입될 수 있다. 특정 실시형태에서, 본 발명의 TCR은 예를 들어, 국제 공개 제WO 2006/000830호에 기재된 바와 같이 각각의 불변 도메인의 잔기 사이에 도입된 이황화 결합을 가질 수 있다.

[0306] 본 발명의 TCR, 특히 알파-베타 이종이량체 TCR은 알파쇄 TRAC 불변 도메인 서열 및/또는 베타쇄 TRBC1 또는 TRBC2 불변 도메인 서열을 포함할 수 있다. 알파 및 베타쇄 불변 도메인 서열은 TRAC의 엑손 2의 Cys4와 TRBC1 또는 TRBC2의 엑손 2의 Cys2 사이의 천연 이황화 결합을 결실시키기 위해 절단 또는 치환에 의해 변형될 수 있다. 알파 및/또는 베타쇄 불변 도메인 서열(들)은 또한 TRAC의 Thr 48 및 TRBC1 또는 TRBC2의 Ser 57에 대한 시스테인 잔기의 치환에 의해 변형될 수 있으며, 상기 시스테인은 TCR의 알파 불변 도메인과 베타 불변 도메인 사이에 이황화 결합을 형성한다.

[0307] 결합 친화도(평형 상수 K_D에 반비례함) 및 결합 반감기(T½로 표현됨)는 임의의 적절한 방법에 의해 결정될 수 있다. TCR의 친화도가 두 배가 되면 K_D가 반으로 줄어든다는 것이 인지될 것이다. T½은 ln 2를 오프-레이트(koff)로 나눈 값으로 계산된다. 따라서 T½이 두 배가 되면 koff가 반으로 줄어든다. TCR에 대한 K_D 및 koff 값은 일반적으로 TCR의 가용성 형태, 즉, 세포질 및 막관통 도메인 잔기를 제거하기 위해 절단된 형태에 대해 측정된다. 따라서, 주어진 TCR은 모 TCR의 가용성 형태가 상기 특성을 갖는다면 모 TCR에 대한 개선된 결합 친화도 및/또는 결합 반감기를 갖는 것으로 이해되어야 한다. 바람직하게는, 주어진 TCR의 결합 친화도 또는 결합 반감기를 동일한 검정 프로토콜을 사용하여 수회, 예를 들어, 3회 이상 측정하고, 결과의 평균을 취한다.

[0308] 본 발명의 TCR은 입양 요법에서 유용성을 갖기 때문에, 본 발명은 본 발명의 TCR을 제시하는 비-천연 발생 및/또는 정제 및/또는 조작된 세포, 특히 T-세포를 포함한다. 본 발명의 TCR을 암호화하는 핵산(예컨대, DNA, cDNA 또는 RNA)을 T 세포에 형질주입시키는 데 적합한 여러 방법이 존재한다(예를 들어, 문헌[Robbins *et al.*, (2008) J Immunol. 180: 6116-6131] 참조). 본 발명의 TCR을 발현하는 T 세포는 암, 예컨대, 췌장암 및 간암의 입양 요법-기반 치료에 사용하기에 적합할 것이다. 당업자에게 널리 공지된 바와 같이, 입양 요법이 수행될 수 있는 다수의 적합한 방법이 존재한다(예를 들어, 문헌[Rosenberg *et al.*, (2008) Nat Rev Cancer 8(4): 299-308] 참조).

[0309] 당업계에 널리 공지된 바와 같이, 본 발명의 TCR은 형질주입된 세포에 의해 발현될 때 번역 후 변형될 수 있다. 글리코실화는 TCR 쇄에서 정의된 아미노산에 대한 올리고당 모이어티의 공유 부착을 포함할 수 있는 그러한 하나의 변형이다. 예를 들어, 아스파라긴 잔기 또는 세린/트레오닌 잔기는 올리고당 부착을 위한 널리 공지된 위치이다. 특정 단백질의 글리코실화 상태는 단백질 서열, 단백질 입체배좌 및 특정 효소의 가용성을 비롯한 여러 요인에 좌우된다. 또한, 글리코실화 상태(즉, 올리고당 유형, 공유 결합 및 총 부착 수)는 단백질 기능에 영향

을 미칠 수 있다. 따라서 재조합 단백질을 생산할 때 글리코실화를 조절하는 것이 종종 바람직하다. 형질주입된 TCR의 글리코실화는 형질주입된 유전자의 돌연변이에 의해 제어될 수 있다(문헌[Kuball J *et al.* (2009), J Exp Med 206(2):463-475]). 이러한 돌연변이도 본 발명에 포함된다.

[0310] TCR은 군 MAGE-A1, MAGE-A2, MAGE-A3, MAGE-A4, MAGE-A5, MAGE-A6, MAGE-A7, MAGE-A8, MAGE-A9, MAGE-A10, MAGE-A11, MAGE-A12, MAGE-A13, GAGE-1, GAGE-2, GAGE-3, GAGE-4, GAGE-5, GAGE-6, GAGE-7, GAGE-8, BAGE-1, RAGE-1, LB33/MUM-1, PRAME, NAG, MAGE-Xp2(MAGE-B2), MAGE-Xp3(MAGE-B3), MAGE-Xp4(AGE-B4), 티로시나제, 뇌글리코겐 포스포릴라제, 멜란-A, MAGE-C1, MAGE-C2, NY-ESO-1, LAGE-1, SSX-1, SSX-2(HOM-MEL-40), SSX-1, SSX-4, SSX-5, SCP-1, CT-7, 알파-액티닌-4, Bcr-Abl 융합 단백질, Casp-8, 베타-카테닌, cdc27, cdk4, cdkn2a, coa-1, dek-can 융합 단백질, EF2, ETV6-AML1 융합 단백질, LDLR-푸코실트랜스퍼라제AS 융합 단백질, HLA-A2, HLA-A11, hsp70-2, KIAA0205, Mart2, Mum-2 및 3, 네오-PAP, 미오신 클래스 I, OS-9, pml-RARa 융합 단백질, PTPRK, K-ras, N-ras, 트라이오세포스페이트 아이소머라제, GnTV, Herv-K-mel, Lage-1, Mage-C2, NA-88, Lage-2, SP17 및 TRP2-Int2, (MART-I), gp100(Pmel 17), TRP-1, TRP-2, MAGE-1, MAGE-3, p15(58), CEA, NY-ESO (LAGE), SCP-1, Hom/Mel-40, p53, H-Ras, HER-2/neu, BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR, 엡스타인 바 바이러스 항원, EBNA, 인간 파필로마바이러스(HPV) 항원 E6 및 E7, TSP-180, MAGE-4, MAGE-5, MAGE-6, p185erbB2, p180erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CA 19-9, CA 72-4, CAM 17.1, NuMa, K-ras, 베타-카테닌, CDK4, Mum-1, p16, TAGE, PSMA, PSCA, CT7, 텔로머라제, 43-9F, 5T4, 791Tgp72, α-태아단백질, 13HCG, BCA225, BTAA, CA 125, CA 15-3(CA 27.29\BCAA), CA 195, CA 242, CA-50, CAM43, CD68\KP1, CO-029, FGF-5, G250, Ga733(EpCAM), HTgp-175, M344, MA-50, MG7-Ag, MOV18, NB\170K, NY-CO-1, RCAS1, SDCCAG16, TA-90(Mac-2 결합 단백질\사이클로필린 C-연관 단백질), TAAL6, TAG72, TLP 및 TPS 내의 항원에 특이적일 수 있다.

[0311] **3.5 전사 인자**

[0312] 조절 T 세포(Treg)는 항상성을 유지하고, 염증 반응의 크기 및 기간을 제어하고, 자가면역 및 알레르기 반응을 예방하는 데 중요하다.

[0313] 일반적으로, Treg는 주로 면역 반응을 억제하는 데 관여하는 것으로 생각되며, 부분적으로는 면역계가 과도한 반응을 방지하는 "자가 점검"으로 기능한다. 특히, Treg는 자가-항원, 무해한 물질, 예컨대, 꽃가루 또는 음식에 대한 내성을 유지하고, 자가면역 질환을 퇴치하는 데 관여한다.

[0314] Treg는 비제한적으로 내장, 피부, 폐 및 간을 포함하는 신체 전체에서 발견된다. 추가로, Treg 세포는 비장, 림프절 및 심지어는 지방 조직과 같은 외부 환경에 직접 노출되지 않는 신체의 특정 구획에서도 발견될 수 있다. 이러한 Treg 세포 집단 각각은 하나 이상의 고유한 특징을 갖고 있는 것으로 알려져 있거나 의심되며, 추가 정보는 개시내용 전문이 본 명세서에 포함된 문헌[Lehtimaki and Lahesmaa, Regulatory T cells control immune responses through their non-redundant tissue specific features, 2013, FRONTIERS IN IMMUNOL., 4(294): 1-10]에서 찾아볼 수 있다.

[0315] 전형적으로, Treg는 적절한 활성화 및 발달을 위해 TGF-β 및 IL-2를 필요로 하는 것으로 알려져 있다. 풍부한 양의 IL-2 수용체(IL-2R)를 발현하는 Treg는 활성화된 T 세포에 의해 생산되는 IL-2에 의존한다. Treg는 강력한 면역 억제 사이토카인인 IL-10과 TGF-β를 둘 다를 생산하는 것으로 알려져 있다. 또한, Treg는 T 세포를 자극하는 항원 제시 세포(APC)의 능력을 저해하는 것으로 알려져 있다. APC 저해에 대해 제안된 기전 중 하나는 Foxp3+ Treg에 의해 발현되는 CTLA-4를 통한 것이다. CTLA-4는 APC 상의 B7 분자에 결합하고 이러한 분자를 차단하거나 내재화를 일으켜 이들을 제거하여 B7의 사용 가능성을 감소시키고 면역 반응에 적절한 공동 자극을 제공할 수 없게 한다고 생각된다. Treg의 기원, 분화 및 기능에 대한 추가 논의는 개시내용 전문이 본 명세서에 포함된 문헌[Dhamne *et al.*, Peripheral and thymic Foxp3+ regulatory T cells in search of origin, distinction, and function, 2013, Frontiers in Immunol., 4 (253): 1-11]에서 찾아볼 수 있다.

[0316] FOXP3, STAT5B 및/또는 HELIOS의 설명 및/또는 아미노산 서열이 본 명세서 및 www.uniprot.org 데이터에 수탁번호: Q9BZS1(FOXP3), P51692(STAT5b) 및/또는 Q9UKS7(HELIOS)에 제공되어 있다.

[0317] ***Foxp3***

[0318] 일부 실시형태에서, 전사 인자는 포크헤드 박스(Forkhead box) P3 전사 인자(Foxp3)이다. Foxp3는 Treg의 분화 및 활성에 있어 핵심 조절자인 것으로 밝혀져 있다. 사실, Foxp3 유전자의 기능 상실 돌연변이는 치명적인 IPEX 증후군(면역 조절장애, 다분비병증, 장병증, X-연관)을 유발하는 것으로 밝혀져 있다. IPEX를 갖는 환자는 심각

한 자가면역 반응, 지속적인 습진 및 대장염을 앓고 있다. Foxp3를 발현하는 Treg 세포는 장의 염증 반응을 제한하는 데 중요한 역할을 한다(Josefowicz, S. Z. *et al.* Nature, 2012, 482, 395-U1510).

[0319] *STAT*

[0320] 전사의 신호 변환인자 및 전사 활성제(signal transducer and activator of transcription: STAT) 단백질 패밀리의 구성원은 세포 면역, 증식, 아폽토시스 및 분화의 많은 측면을 매개하는 세포내 전사 인자이다. 이들은 주로 막 수용체-연관 야누스 키나제(JAK)에 의해 활성화된다. 이 경로의 조절장애는 원발성 종양에서 흔히 관찰되며 증가된 혈관신생, 향상된 종양의 생존 및 면역억제로 이어진다. 유전자 넉아웃 연구는 STAT 단백질이 면역계의 발달과 기능에 관여하고, 면역 관용 및 종양 감시를 유지하는 역할을 한다는 증거를 제공하였다.

[0321] 확인된 7개의 포유동물 STAT 패밀리 구성원이 존재한다: STAT1, STAT2, STAT3, STAT4, STAT5(STAT5A 및 STAT5B 포함) 및 STATE.

[0322] 사이토카인 또는 성장 인자의 세포외 결합은 SH2 도메인을 통한 이량체화를 촉진하는 STAT 단백질 내의 특정 티로신 잔기를 인산화하는 수용체-연관 야누스 키나제의 활성화를 유도한다. 그런 다음 인산화된 이량체는 중요한 α/β 삼원 복합체를 통해 핵으로 활발히 수송된다. 본래, STAT 단백질은 인산화가 핵 보유에 필요한 것으로 생각되었기 때문에 잠복 세포질 전사 인자로 기술되었다. 그러나 인산화되지 않은 STAT 단백질은 사이토솔과 핵 사이를 왕복하며, 유전자 발현에 역할을 한다. STAT가 핵에 도달하면, 그것은 사이토카인-유도성 유전자의 프로모터 영역 내의 감마-활성화 부위(GAS)라고 불리는 공통 DNA-인식 모티프에 결합하여 전사를 활성화한다. STAT 단백질은 핵 포스파타제에 의해 탈인산화될 수 있는데, 이는 STAT의 불활성화 및 엑스포틴-RanGTP 복합체에 의한 핵 밖으로의 후속 수송으로 이어진다.

[0323] 일부 실시형태에서, 본 개시내용의 STAT 단백질은 이의 발현 수준 또는 활성을 조절하는 변형을 포함하는 STAT 단백질일 수 있다. 일부 실시형태에서, 이러한 변형은 특히 STAT 이량체화, 신호전달 파트너에 대한 STAT 단백질 결합, STAT 단백질 국지화 또는 STAT 단백질 분해에 영향을 미치는 돌연변이를 포함한다. 일부 실시형태에서, 본 개시내용의 STAT 단백질은 구성적으로 활성이다. 일부 실시형태에서, 본 개시내용의 STAT 단백질은 구성적 이량체화로 인해 구성적으로 활성이다. 일부 실시형태에서, 본 개시내용의 STAT 단백질은 전문이 참조에 의해 본 명세서에 포함된 문헌[Onishi, M. *et al.*, Mol. Cell. Biol. July 1998 vol. 18 no. 7 3871-3879]에 기재되어 있는 바와 같이 구성적 인산화로 인해서 구성적으로 활성이다.

[0324] **3.6 백신**

[0325] 실시형태에서, 하나 이상의 발현 서열은 항원, 예를 들어, 종양 항원, 또는 이의 단편을 암호화한다. 일부 실시형태에서, 이러한 서열의 발현은 면역원성 조성물, 예를 들어, 특이적 T-세포 반응을 일으킬 수 있는 백신 조성물을 생산한다. 일부 실시형태에서, 항원은 신생항원이다.

[0326] **4. 절단 부위**

[0327] 일부 실시형태에서, 폴리뉴클레오타이드 작제물에서 2개 이상의 발현 서열은 하나 이상의 절단 부위 서열에 의해 분리될 수 있다.

[0328] 절단 부위는 2개 이상의 폴리펩타이드가 분리될 수 있게 하는 임의의 서열일 수 있다. 절단 부위는 자가-절단될 수 있어서, 폴리펩타이드가 생산될 때 어떠한 외부 절단 활성에 대한 필요도 없이 개별 폴리펩타이드로 즉시 절단된다.

[0329] 절단 부위는 푸린 절단 부위일 수 있다.

[0330] 푸린은 서브틸리신-유사 프로단백질 전환효소 패밀리에 속하는 효소이다. 이 패밀리의 구성원은 잠복 전구체 단백질을 생물학적 활성 생성물로 처리하는 프로단백질 전환효소이다. 푸린은 쌍을 이루는 염기성 아미노산 처리 부위에서 전구체 단백질을 효율적으로 절단할 수 있는 칼슘-의존성 세린 엔도프로테아제이다. 푸린 기질의 예는 부갑상선 호르몬, 변형 성장 인자 베타 1 전구체, 프로알부민, 프로-베타-분비효소, 막 유형-1 기질 메탈로프로테이나제, 프로-신경 성장 인자의 베타 소단위 및 폰 빌레브란트 인자를 포함한다. 푸린은 염기성 아미노산 표적 서열(표준적으로 Arg-X-(Arg/Lys)-Arg)의 바로 하류에 있는 단백질을 절단하고, 골지체에 풍부하다.

[0331] 절단 부위는 자가-절단 펩타이드를 암호화할 수 있다.

[0332] 절단 부위는 2A 자가-절단 펩타이드의 C-말단에서 글리실-프로필 결합의 스키핑과 같은 리보솜 스키핑에 의해 작동될 수 있다. 일부 실시형태에서, 입체 장애는 리보솜 스키핑을 유발한다. 일부 실시형태에서, 2A 자가-절단

펩타이드는 서열 GDVEXNPGP(서열번호 324)를 함유하고, 여기서 X는 E 또는 S이다. 일부 실시형태에서, 2A 자가-절단 펩타이드의 상류에 암호화된 단백질은 번역 후 C-말단 프롤린을 제외한 2A 자가-절단 펩타이드에 부착된다. 일부 실시형태에서, 2A 자가-절단 펩타이드의 하류에 암호화된 단백질은 번역 후 이의 N-말단에서 프롤린에 부착된다.

[0333] 자가-절단 펩타이드는 압토- 또는 카디오바이러스로부터의 2A 자가-절단 펩타이드일 수 있다. 압토- 및 카디오바이러스의 1차 2A/2B 절단은 자체 C-말단에서 2A 절단에 의해 매개된다. 압토바이러스, 예컨대, 구제역 바이러스(FMDV) 및 말 비염 A 바이러스에서, 2A 영역은 단백질 2B의 N-말단 잔기(보존된 프롤린 잔기)와 함께 자체 C-말단에서 절단을 매개할 수 있는 자율적 요소를 나타내는 약 18개 아미노산의 짧은 부분이다(문헌[Donelly *et al.*(2001)]).

[0334] 2A-유사 서열은 압토- 또는 카디오바이러스 이외의 피코르나바이러스, '피코르나바이러스-유사' 곤충 바이러스, C형 로타바이러스 및 트리파노소마 종 내의 반복 서열 및 박테리아 서열에서 발견되었다(문헌[Donnelly *et al.*(2001)]). 절단 부위는 표 8에 열거된 것과 같은 이러한 2A-유사 서열 중 하나를 포함할 수 있다.

[0335] 일부 실시형태에서, 자가-절단 펩타이드는 F2A이다. 일부 실시형태에서, 자가-절단 펩타이드는 구제역 바이러스로부터 유래된다. 일부 실시형태에서, 자가-절단 펩타이드는 E2A이다. 일부 실시형태에서, 자가 절단 펩타이드는 말 비염 A 바이러스로부터 유래된다. 일부 실시형태에서, 자가-절단 펩타이드는 P2A이다. 일부 실시형태에서, 자가-절단 펩타이드는 돼지 테스코바이러스-1로부터 유래된다. 일부 실시형태에서, 자가-절단 펩타이드는 T2A이다. 일부 실시형태에서, 자가-절단 펩타이드는 아시그나 바이러스(asigna virus)로부터 유래된다. 일부 실시형태에서, 자가-절단 펩타이드는 표 8에 열거된 서열을 갖는다.

[0336] 실시형태에서, 절단 부위에 의해 분리된 치료용 단백질을 암호화하는 발현 서열은 동일한 수준의 단백질 발현을 갖는다.

[0337] 일부 실시형태에서, 자가-절단 펩타이드는 문헌[Liu, Z., Chen, O., Wall, J.B.J. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Sci Rep 7, 2193 (2017)]에 기재된 바와 같다.

[0338] **5. 제2 IRES를 함유하는 폴리뉴클레오타이드**

[0339] 일부 실시형태에서, 제1 스페이서 및 제2 스페이서발현 서열에 의해서 암호화된 치료용 단백질의 발현 비율은 circRNA에 사용된 IRES 및 절단 부위 또는 제2 IRES가 제1 스페이서 및 제2 스페이서발현 서열을 분리하는지 여부에 의해 제어되거나 영향을 받을 수 있다. 제1 스페이서 및 제2 스페이서발현 서열에 의해 암호화되는 단백질의 동일한 발현이 요구되는 경우, circRNA는 제1 발현 서열과 제2 발현 서열 사이의 절단 부위, 예를 들어, 2A 자가-절단 펩타이드를 암호화할 수 있다. 제1 발현 서열에 의해 암호화된 단백질의 더 큰 발현이 요구되는 경우, circRNA는 제1 IRES 및 제2 IRES를 암호화할 수 있고, 여기서 제1 IRES는 제2 IRES보다 더 큰 발현과 연관되거나, 또는 제2 IRES는 유전자간 영역(IGR) IRES이다. 제2 발현 서열에 의해 암호화된 단백질의 더 큰 발현이 요구되는 경우, circRNA는 제1 IRES 및 제2 IRES를 암호화할 수 있고, 여기서 제2 IRES는 제1 IRES보다 더 큰 발현과 연관되어 있다.

[0340] 일부 실시형태에서, RNA 폴리뉴클레오타이드는 본 명세서에 기재된 바와 같은 제1 IRES 및 제2 IRES를 함유한다. 일부 실시형태에서, DNA 벡터는 본 명세서에 기재된 바와 같은 제1 IRES 및 제2 IRES를 암호화한다.

[0341] 실시형태에서, 제1 IRES 및 제2 IRES는 동일한 서열을 갖는다. 실시형태에서, 제1 IRES 및 제2 IRES는 상이한 서열을 갖는다. 실시형태에서, 제1 IRES는 표 1에 열거된 서열(서열번호 1 내지 72)을 갖는 IRES이다. 일부 실시형태에서, 제1 IRES는 살리바이러스 IRES이다. 일부 실시형태에서, 제1 IRES는 살리바이러스 SZ1 IRES이다. 실시형태에서, 제2 IRES는 표 1에 열거된 서열(서열번호 1 내지 72)을 갖는 IRES이다. 일부 실시형태에서, 제2 IRES는 살리바이러스 IRES이다. 일부 실시형태에서, 제1 IRES는 살리바이러스 SZ1 IRES이다.

[0342] 일부 실시형태에서, 제1 IRES는 제2 IRES보다 더 큰 발현과 연관된다(예를 들어, 단일 IRES를 함유하는 작제물을 사용하는 것과 비교하는 경우 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 또는 100% 더 큰 발현). 일부 실시형태에서, 제2 IRES는 제1 IRES보다 더 큰 발현과 연관된다(예를 들어, 단일 IRES를 함유하는 작제물을 사용하는 것과 비교하는 경우 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 또는 100% 더 큰 발현). 일부 실시형태에서, 제2 IRES는 유전자간 영역(IGR) IRES이다.

[0343] 단일 circRNA 폴리뉴클레오타이드를 갖는 2개의 단백질을 발현시키는 것은 다중 폴리뉴클레오타이드를 사용한

발현보다 이점이 있다. 일부 실시형태에서, 본 발명의 circRNA 폴리뉴클레오타이드로부터의 2개의 단백질의 발현은 다중 폴리뉴클레오타이드로부터의 발현보다 더 일관된 발현 비율을 생성한다. 일부 실시형태에서, 본 발명의 circRNA 폴리뉴클레오타이드로부터의 2개의 단백질의 발현은 DNA의 지속적인 발현보다 바람직할 수 있는 일시적인 발현을 생성한다.

[0344] **6. 폴리뉴클레오타이드의 생산**

[0345] 본 명세서에 제공된 벡터는 표준 분자 생물학 기술을 사용하여 제조될 수 있다. 예를 들어, 본 명세서에서 제공되는 벡터의 다양한 요소는 세포로부터 cDNA 및 게놈 라이브러리를 스크리닝하거나 이를 포함하는 것으로 알려진 벡터로부터 폴리뉴클레오타이드를 유도하는 것과 같은 재조합 방법을 사용하여 얻을 수 있다.

[0346] 본 명세서에 제공된 벡터의 다양한 요소는 또한 공지된 서열에 기초하여 클로닝되기 보다는 합성에 의해서 생산될 수 있다. 완전한 서열은 표준 방법에 의해 제조된 중첩 올리고뉴클레오타이드로부터 조립되고, 완전한 서열로 조립될 수 있다. 예를 들어, 문헌[Edge, Nature (1981) 292:756; Nambair *et al.*, Science (1984) 223 : 1299; 및 Jay *et al.*, J. Biol. Chem. (1984) 259:631 1] 참조.

[0347] 따라서, 특정 뉴클레오타이드 서열은 목적하는 서열을 보유하는 벡터로부터 얻어지거나, 절절한 경우 부위 지정 돌연변이유발 및 중합효소 연쇄 반응(PCR) 기술과 같은 당업계에 공지된 다양한 올리고뉴클레오타이드 합성 기술을 사용하여 완전히 또는 부분적으로 합성될 수 있다. 목적하는 벡터 요소를 암호화하는 뉴클레오타이드 서열을 얻는 한 가지 방법은 종래의 자동화된 폴리뉴클레오타이드 합성기에서 생성된 중첩 합성 올리고뉴클레오타이드의 상보적 세트를 어닐링하고, 그 다음 적절한 DNA 리가제로 결찰시키고, PCR을 통해 결찰된 뉴클레오타이드 서열을 증폭시키는 것이다. 예를 들어, 문헌[Jayaraman *et al.*, Proc. Natl. Acad. Sci. USA (1991) 88:4084-4088] 참조. 또한, 올리고뉴클레오타이드-지정 합성(문헌[Jones *et al.*, Nature (1986) 54:75-82]), 기존 뉴클레오타이드 영역의 올리고뉴클레오타이드 지정 돌연변이유발(문헌[Riechmann *et al.*, Nature (1988) 332:323-327 및 Verhoeyen *et al.*, Science (1988) 239: 1534-1536]) 및 T4 DNA 중합효소를 사용한 갭핑된 올리고뉴클레오타이드의 효소적 충전(문헌[Queen *et al.*, Proc. Natl. Acad. Sci. USA (1989) 86: 10029-10033])이 사용될 수 있다.

[0348] 본 명세서에서 제공된 전구체 RNA는 벡터에 의해 암호화되는 전구체 RNA의 전사를 허용하는 조건 하에서 본 명세서에서 제공된 벡터를 인큐베이션시킴으로써 생성될 수 있다. 예를 들어, 일부 실시형태에서 전구체 RNA는 5' 듀플렉스 형성 영역 및/또는 발현 서열의 상류에 RNA 중합효소 프로모터를 포함하는 본 명세서에 제공된 벡터를 시험관내 전사를 허용하는 조건 하에서 상용성 RNA 중합효소와 인큐베이션시킴으로써 합성된다. 일부 실시형태에서, 벡터는 박테리오파지 RNA 중합효소에 의해 세포 내부에서 또는 숙주 RNA 중합효소 II에 의해 세포 핵에서 인큐베이션된다.

[0349] 특정 실시형태에서, 본 명세서에 제공된 벡터를 주형으로 사용하여 시험관내 전사를 수행함으로써 전구체 RNA를 생성시키는 방법이 본 명세서에 제공된다(예를 들어, 5' 듀플렉스 형성의 상류에 위치한 RNA 중합효소 프로모터와 함께 본 명세서에 제공된 벡터).

[0350] 특정 실시형태에서, 생성된 전구체 RNA는 RNA 원형화가 일어나는 온도(예를 들어, 20 ㎖ 및 60 ㎖)에서 마그네슘 이온 및 구아노신 뉴클레오타이드 또는 뉴클레오사이드의 존재 하에서 생성된 전구체 RNA를 인큐베이션시킴으로써 원형 RNA(예를 들어, 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드)를 생성시키는 데 사용될 수 있다.

[0351] 따라서, 특정 실시형태에서 원형 RNA를 제조하는 방법이 본 명세서에 제공된다. 특정 실시형태에서, 이 방법은 주형으로서 본 명세서에 제공된 벡터(예를 들어, 포스트 스플라이싱 3' 그룹 I 인트론 단편, 내부 리보솜 유입 부위(IRES), 발현 서열, 절단 부위를암호화하는 폴리뉴클레오타이드 서열, 제2 발현 서열 및 5' 그룹 I 인트론 단편)를 사용한 전사(예를 들어, 유출 전사)에 의해서 전구체 RNA를 합성하는 단계 및 생성된 전구체 RNA를 2가 양이온(예를 들어, 마그네슘 이온) 및 GTP의 존재 하에서 인큐베이션시켜, 이것을 원형화하여 원형 RNA를 형성하는 단계를 포함한다. 일부 실시형태에서, 본 명세서에 개시된 전구체 RNA는 마그네슘 이온 및 GTP의 부재 하에서 그리고/또는 마그네슘 이온 및 GTP와의 인큐베이션 단계 없이 원형화할 수 있다. 적어도 부분적으로 mRNA가 면역원성 5' 캡을 함유하기 때문에 원형 RNA가 상응하는 mRNA에 비해 면역원성을 감소시킨다는 것을 발견하였다. 특정 프로모터(예를 들어, T7 프로모터)로부터 DNA 벡터를 전사시켜 전구체 RNA를 생산하는 경우, 전구체 RNA의 5' 단부가 G인 것으로 이해된다. 낮은 수준의 오염물 선형 mRNA를 함유하는 원형 RNA 조성물의 면역원성을 감소시키기 위해서, 대부분의 전사체가 캡핑될 수 없는 5' GMP를 함유하도록 전사 동안 GTP에 비해 과량의

GMP가 제공될 수 있다. 따라서, 일부 실시형태에서, 전사는 과량의 GMP의 존재 하에서 수행된다. 일부 실시형태에서, 전사는 GMP 농도 대 GTP 농도의 비율이 약 3:1 내지 약 15:1, 예를 들어, 약 3:1 내지 약 10:1, 약 3:1 내지 약 5:1, 약 3:1, 약 4:1 또는 약 5:1의 범위 내에서 수행된다.

[0352] 일부 실시형태에서, 원형 RNA를 포함하는 조성물이 정제되었다. 원형 RNA는 칼럼 크로마토그래피, 겔 여과 크로마토그래피 및 크기 배제 크로마토그래피와 같이 당업계에 일반적으로 사용되는 공지된 방법에 의해 정제될 수 있다. 일부 실시형태에서, 정제는 다음 단계 중 하나 이상을 포함한다: 포스파타제 처리, HPLC 크기 배제 정제, 및 RNase R 소화. 일부 실시형태에서, 정제는 다음 단계를 순서대로 포함한다: RNase R 소화, 포스파타제 처리, 및 HPLC 크기 배제 정제. 일부 실시형태에서, 정제는 역상 HPLC를 포함한다. 일부 실시형태에서, 정제된 조성물은 비정제된 RNA보다 더 적은 이중 가닥 RNA, DNA 스플린트, 삼인산화된 RNA, 포스파타제 단백질, 단백질 리가제, 캡핑 효소 및/또는 니킹된 RNA를 함유한다. 일부 실시형태에서, 정제된 조성물은 비정제된 조성물보다 덜 면역원성이다. 일부 실시형태에서, 정제된 조성물에 노출된 면역 세포는 비정제된 조성물에 노출된 면역 세포보다 TNFα, RIG-I, IL-2, IL-6, IFNɣ 및/또는 1형 인터페론, 예를 들어, IFN-β1을 더 적게 생성한다.

[0353] **7. 나노입자**

[0354] 특정 양상에서, 본 명세서에 제공된 원형 RNA를 포함하는 약제학적 조성물이 본 명세서에 제공된다. 특정 실시형태에서, 이러한 약제학적 조성물은 전달을 용이하게 하기 위해 나노입자와 함께 제형화된다.

[0355] 특정 실시형태에서, 본 명세서에 제공된 원형 RNA는 전달 비히클, 예를 들어, 나노입자 또는 나노입자를 포함하는 조성물로 세포에 전달되고/되거나 표적화될 수 있다. 일부 실시형태에서, 원형 RNA는 또한 전달 비히클 또는 전달 비히클을 포함하는 조성물로 대상체에게 전달될 수 있다. 일부 실시형태에서, 전달 비히클은 나노입자이다. 일부 실시형태에서, 나노입자는 지질 나노입자, 고체 지질 나노입자, 중합체 코어-쉘 나노입자 또는 생분해성 나노입자이다. 일부 실시형태에서, 전달 비히클은 하나 이상의 양이온성 지질, 비-양이온성 지질, 이온화 가능한 지질, PEG-변형된 지질, 폴리글루탐산 중합체, 히알루론산 중합체, 폴리 β-아미노 에스터, 폴리 베타 아미노 펩타이드, 또는 양으로 하전된 펩타이드를 포함하거나 이들로 코팅된다.

[0356] 일 실시형태에서, 전달 비히클은 표적 세포로의 circRNA의 전달을 최적화하도록 선택 및/또는 제조될 수 있다. 예를 들어, 표적 세포가 간세포인 경우, 전달 비히클의 특성(예를 들어, 크기, 전하 및/또는 pH)은 이러한 전달 비히클을 표적 세포에 효과적으로 전달하고, 면역 제거를 감소시키고/시키거나 표적 세포 내의 체류를 촉진하도록 최적화될 수 있다.

[0357] 표적 세포로의 핵산의 전달을 용이하게 하기 위한 전달 비히클의 사용이 본 발명에 의해 고려된다. 리포솜(예를 들어, 리포솜 지질 나노입자)은 일반적으로 연구, 산업 및 의학의 다양한 응용 분야, 특히 생체내 진단 또는 치료 화합물의 전달 비히클로서의 사용에 유용하고(문헌[Lasic, Trends Biotechnol., 16: 307-321, 1998; Drummond *et al.*, Pharmacol. Rev., 51: 691-743, 1999]), 일반적으로 하나 이상의 이중층의 막에 의해서 외부 매질로부터 격리되는 내부 수성 공간을 갖는 미세한 소포를 특징으로 한다. 리포솜의 이중층 막은 일반적으로 공간적으로 분리된 친수성 및 소수성 도메인을 포함하는 합성 또는 자연 기원의 지질과 같은 양친매성 분자에 의해 형성된다(문헌[Lasic, Trends Biotechnol., 16: 307-321, 1998]). 리포솜의 이중층 막은 또한 양친매성 중합체 및 계면활성제(예를 들어, 폴리머로솜, 니오솜 등)에 의해 형성될 수 있다.

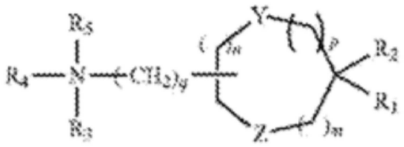
[0358] 본 발명의 맥락에서, 전달 비히클은 전형적으로 circRNA를 표적 세포로 수송하는 역할을 한다. 본 발명의 목적을 위해, 전달 비히클은 목적하는 핵산을 함유하거나 캡슐화하도록 준비된다. 목적하는 실체(예를 들어, 핵산)를 리포솜에 혼입하는 과정은 종종 로딩이라고 지칭된다(문헌[Lasic, *et al.*, FEBS Lett., 312: 255-258, 1992]). 리포솜-혼입된 핵산은 리포솜의 내부 공간, 리포솜의 이중층 막 내부에 완전히 또는 부분적으로 위치하거나 리포솜 막의 외부 표면과 회합될 수 있다. circRNA를 리포솜과 같은 전달 비히클에 혼입하는 목적은 종종 핵산을 분해하는 효소 또는 화학물질 및/또는 핵산의 빠른 배설을 유발하는 시스템 또는 수용체를 함유할 수 있는 환경으로부터 핵산을 보호하는 것이다. 따라서, 본 발명의 일 실시형태에서, 선택된 전달 비히클은 내부에 함유된 circRNA의 안정성을 향상시킬 수 있다. 리포솜은 캡슐화된 circRNA가 표적 세포에 도달하도록 할 수 있거나, 대안적으로 투여된 circRNA의 존재가 쓸모없거나 또는 바람직하지 않을 수 있는 다른 부위 또는 세포로의 이러한 circRNA의 전달을 제한할 수 있다. 또한, 예를 들어, 양이온성 리포솜과 같은 전달 비히클에 circRNA를 혼입하는 것은 또한 표적 세포 내로의 이러한 circRNA의 전달을 촉진한다. 일부 실시형태에서, 본 명세서에 개시된 전달 비히클은 예를 들어, 전달 비히클(예를 들어, 지질 나노입자) 내에 캡슐화된 내용물의 엔도솜 또는 리소좀 방출을 촉진하는 역할을 할 수 있다.

[0359] 이상적으로, 전달 비히클은 조성물이 높은 형질주입 효율 및 향상된 안정성을 나타내도록 하나 이상의 목적하는 circRNA를 캡슐화하도록 제조된다. 리포솜은 핵산을 표적 세포에 도입하는 것을 용이하게 할 수 있지만, 일부 예에서, 공중합체로서 다중양이온(예를 들어, 폴리 L-라이신 및 프로타민)을 첨가하면 시험관내 그리고 생체내 둘 다에서 다수의 세포주에서 여러 유형의 양이온성 리포솜의 형질주입 효율을 2 내지 28배만큼 현저하게 향상시킬 수 있다(문헌[N J. Caplen, *et al.*, Gene Ther. 1995; 2: 603; S. Li, *et al.*, Gene Ther. 1997; 4, 891]).

[0360] 특정 실시형태에서 본 명세서에는 (예를 들어, 이러한 표적 세포의 지질 막을 침투하거나 또는 이와 융합함으로써) 원형 RNA가 하나 이상의 표적 세포로 전달되고, 방출되는 것을 용이하게 하거나 향상시키기 위한 전달 비히클의 성분으로서 사용될 수 있는 이온화 가능한 지질이 개시된다. 특정 실시형태에서, 이온화 가능한 지질은 (예를 들어, 산화성, 환원성 또는 산성 조건 하에서) 예를 들어, 친수성 작용성 헤드기를 화합물의 친유성 작용성 테일기로부터 해리시킴으로써, 하나 이상의 표적 세포의 지질 이중층에서 상 전이를 용이하게 하는 하나 이상의 절단 가능한 작용기(예를 들어, 다이설파이드)를 포함한다.

[0361] 일부 실시형태에서, 이온화 가능한 지질은 국제 특허 출원 제PCT/US2018/058555호에 기재된 바와 같은 지질이다.

[0362] 실시형태 중 일부에서, 양이온성 지질은 하기 화학식을 갖는다:

[0363] 

[0364] 식 중,

[0365] R_1 및 R_2는 동일하거나 상이하고, 독립적으로 선택적으로 치환된 C_{10}-C_{24} 알킬, 선택적으로 치환된 C_{10}-C_{24} 알켄일, 선택적으로 치환된 C_{10}-C_{24} 알킨일 또는 선택적으로 치환된 C_{10}-C_{24} 아실이고;

[0366] R_3 및 R_4는 동일하거나 상이하고, 독립적으로 선택적으로 치환된 C_1-C_6 알킬, 선택적으로 치환된 C_2-C_6 알켄일 또는 선택적으로 치환된 C_2-C_6 알킨일이거나 R_3과 R_4는 합쳐져서 4 내지 6개 탄소 원자 및 질소 및 산소로부터 선택된 1 또는 2개의 헤테로원자의 선택적으로 치환된 헤테로사이클릭 고리를 형성할 수 있고;

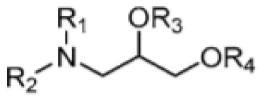
[0367] R_5는 존재하지 않거나 존재하고, 존재하는 경우 수소 또는 C_1-C_6 알킬이고; m, n 및 p는 동일하거나 상이하고, 독립적으로 0 또는 1이되, 단 m, n 및 p는 동시에 0은 아니고; q는 0, 1, 2, 3 또는 4이고;

[0368] Y 및 Z는 동일하거나 상이하고, 독립적으로 O, S 또는 NH이다.

[0369] 일 실시형태에서, R_1 및 R_2는 각각 리놀레일이고, 아미노 지질은 다이리놀레일 아미노 지질이다.

[0370] 일 실시형태에서, 아미노 지질은 다이리놀레일 아미노 지질이다.

[0371] 다양한 다른 실시형태에서, 양이온성 지질은 하기 구조 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체를 갖는다:

[0372] 

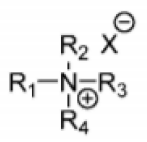
[0373] 식 중,

[0374] R_1 및 R_2는 각각 독립적으로 H 및 C_1-C_3 알킬로 이루어진 군으로부터 선택되고;

[0375] R_3 및 R_4는 각각 독립적으로 약 10 내지 약 20개 탄소 원자를 갖는 알킬기이되, R_3 및 R_4 중 적어도 하나는 적어도 2개의 불포화 부위를 포함한다.

[0376] 일부 실시형태에서, R_3 및 R_4는 각각 독립적으로 도데카다이엔일, 테트라데카다이엔일, 헥사데카다이엔일, 리놀레일 및 이코사다이엔일로부터 선택된다. 실시형태에서, R_3 및 R_4는 둘 다 리놀레일이다. 일부 실시형태에서, R_3 및/또는 R_4는 적어도 3개의 불포화 부위를 포함할 수 있다(예를 들어, R_3 및/또는 R_4는 예를 들어, 도데카트라이엔일, 테트라데카트라이엔일, 헥사데카트라이엔일, 리놀렌일 및 이코사트라이엔일일 수 있다).

[0377] 일부 실시형태에서, 양이온성 지질은 하기 구조, 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체를 갖는다:

[0378] 

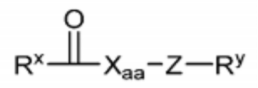
[0379] 식 중,

[0380] R_1 및 R_2는 각각 독립적으로 H 및 C_1-C_3 알킬로부터 선택되고;

[0381] R_3 및 R_4는 각각 독립적으로 약 10 내지 약 20개 탄소 원자를 갖는 알킬기이되, R_3 및 R_4 중 적어도 하나는 적어도 2개의 불포화 부위를 포함한다.

[0382] 일 실시형태에서, R_3 및 R_4는 동일하고, 예를 들어, 일부 실시형태에서 R_3 및 R_4 둘 다는 리놀레일(C_{18}-알킬)이다. 또 다른 실시형태에서, R_3 및 R_4는 상이하고, 예를 들어, 일부 실시형태에서, R_3은 테트라데카트라이엔일(C_{14}-알킬)이고, R_4는 리놀레일(C_{18}-알킬)이다. 바람직한 실시형태에서, 본 발명의 양이온성 지질(들)은 비대칭이고, 즉, R_3 및 R_4는 동일하다. 또 다른 바람직한 실시형태에서, R_3 및 R_4 둘 다는 적어도 2개의 불포화 부위를 포함한다. 일부 실시형태에서, R_3 및 R_4는 각각 독립적으로 도데카다이엔일, 테트라데카다이엔일, 헥사데카다이엔일, 리놀레일 및 이코사다이엔일로부터 선택된다. 실시형태에서, R_3 및 R_4는 둘 다 리놀레일이다. 일부 실시형태에서, R_3 및/또는 R_4는 적어도 3개의 불포화 부위를 포함하고, 각각 독립적으로 도데카트라이엔일, 테트라데카트라이엔일, 헥사데카트라이엔일, 리놀렌일 및 이코사트라이엔일로부터 선택된다.

[0383] 다양한 실시형태에서, 양이온성 지질은 하기 화학식, 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체을 갖는다:

[0384] 

[0385] 식 중,

[0386] X_{aa}는 화학식 $-NR^N-CR^1R^2-C(C=O)-$를 갖는 D- 또는 L-아미노산 잔기, 또는 펩타이드 또는 화학식 $-\{NR^N-CR^1R^2-C(C=O)\}_n-$를 갖는 아미노산 잔기의 펩타이드이되, n은 2 내지 20의 정수이고;

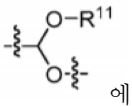
[0387] R^1은 독립적으로, 각각의 경우에 대해서, 비-수소 또는 아미노산의 치환된 또는 비치환된 측쇄이고;

[0388] R^2 및 R^N은 독립적으로, 각각의 경우에 대해서, 수소, 1 내지 20개의 탄소 원자를 갖는 탄소, 산소, 질소, 황 및 수소 원자로 이루어진 유기기 또는 이들의 조합물, $C_{(1-5)}$알킬, 사이클로알킬, 사이클로알킬알킬, $C_{(1-5)}$알켄일, $C_{(1-5)}$알킨일, $C_{(1-5)}$알칸오일, $C_{(1-5)}$알칸오일옥시, $C_{(1-5)}$알콕시, $C_{(1-5)}$알콕시- $C_{(1-5)}$알킬, $C_{(1-5)}$알콕시- $C_{(1-5)}$알콕시, $C_{(1-5)}$알킬-아미노- $C_{(1-5)}$알킬-, $C_{(1-5)}$다이알킬-아미노- $C_{(1-5)}$알킬-, 나이트로-$C_{(1-5)}$알킬, 사이아노-$C_{(1-5)}$알킬, 아릴-$C_{(1-5)}$알킬, 4-바이페닐-$C_{(1-5)}$알킬, 카복실 또는 하이드록실이고;

[0389] Z는 -NH-, -O-, -S-, $-CH_2S-$, $-CH_2S(O)-$, 또는 수소, 탄소, 산소, 질소 및 황 원자로부터 선택된 1 내지 40개의 원자로 이루어진 유기 링커이고(바람직하게는 Z는 -NH- 또는 -O-임);

[0390] R^x 및 R^y는, 독립적으로, (i) 지질(이것은 자연 발생이거나 합성일 수 있음), 예를 들어, 인지질, 당지질, 트라이아실글리세롤, 글리세로인지질, 스핑고지질, 세라마이드, 스핑고미엘린, 세레브로사이드 또는 강글리오사이드로부터 유래된 친유성 테일(여기서 테일은 선택적으로 스테로이드를 포함함); (ii) 수소, 하이드록실, 아미노 및 유기 보호기로부터 선택된 아미노산 말단기; 또는 (iii) 치환된 또는 비치환된 $C_{(3-22)}$알킬, $C_{(6-12)}$사이클로알킬, $C_{(6-12)}$사이클로알킬- $C_{(3-22)}$알킬, $C_{(3-22)}$알켄일, $C_{(3-22)}$알킨일, $C_{(3-22)}$알콕시 또는 $C_{(6-12)}$-알콕시 $C_{(3-22)}$알킬이다.

[0391] 일부 실시형태에서, R^x 및 R^y 중 하나는 상기에 정의된 바와 같은 친유성 테일이고, 나머지는 아미노산 말단기이다. 일부 실시형태에서, R^x 및 R^y 둘 다는 친유성 테일이다.

[0392] 일부 실시형태에서, R^x 및 R^y 중 적어도 하나는 하나 이상의 생분해성 기(예를 들어, -OC(O)-, -C(O)O-, -SC(O)-, -C(O)S-, -OC(S)-, -C(S)O-, -S-S-, -C(O)(NR^5)-, -N(R^5)C(O)-, -C(S)(NR^5)-, -N(R^5)C(O)-, -N(R^5)C(O)N(R^5)-, -OC(O)O-, -OSi(R^5)$_2$O-, -C(O)(CR^3R^4)C(O)O-, -OC(O)(CR^3R^4)C(O)- 또는  에 의해서 개재된다.

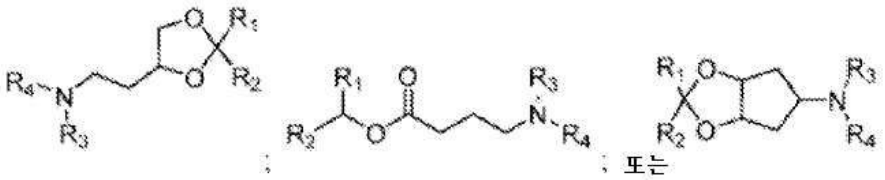
[0393] 일부 실시형태에서, R^{11}은 C_2-C_8알킬 또는 알켄일이다.

[0394] 일부 실시형태에서, R^5의 각각의 경우는, 독립적으로, H 또는 알킬이다.

[0395] 일부 실시형태에서, R^3 및 R^4의 각각의 경우는, 독립적으로 H, 할로겐, OH, 알킬, 알콕시, $-NH_2$, 알킬아미노 또는 다이알킬아미노이거나; R_3과 R_4는이들이 직접 부착되는 탄소 원자와 함께, 사이클로알킬기를 형성한다. 일부 특정 실시형태에서, R^3 및 R^4의 각각의 경우는, 독립적으로 H 또는 C_1-C_4알킬이다.

[0396] 일부 실시형태에서, R^x 및 R^y는 각각 독립적으로, 하나 이상의 탄소-탄소 이중 결합을 갖는다.

[0397] 일부 실시형태에서, 양이온성 지질은 하기 중 하나, 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체이다:



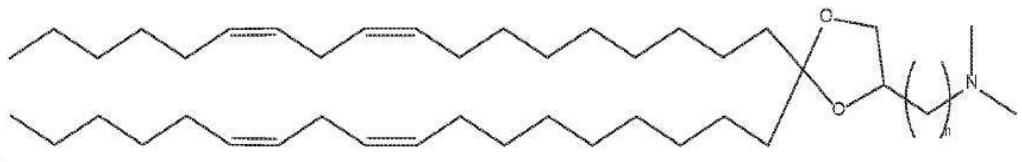
[0398]

[0399] 식 중,

[0400] R_1 및 R_2는 각각 독립적으로 알킬, 알켄일 또는 알킨일이고, 이들 각각은 선택적으로 치환될 수 있고;

[0401] R_3 및 R_4는 각각 독립적으로 C_1-C_6 알킬이거나, R_3과 R_4는 함께 합쳐져서 선택적으로 치환된 헤테로사이클릭 고리를 형성한다.

[0402] 대표적인 유용한 다이리놀레일 아미노 지질은 하기 화학식을 갖는다:

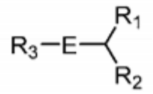


[0403]

[0404] 식 중, n은 0, 1, 2, 3 또는 4이다.

[0405] 일 실시형태에서, 양이온성 지질은 DLin-K-DMA이다. 일 실시형태에서, 양이온성 지질은 DLin-KC2-DMA(상기 DLin-K-DMA, 식 중, n은 2임)이다.

[0406] 일 실시형태에서, 양이온성 지질은 하기 구조, 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체를 갖는다:

[0407] 

[0408] 식 중,

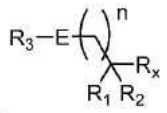
[0409] R_1 및 R_2는 각각 독립적으로 각각의 경우에 대해서 선택적으로 치환된 C_{10}-C_{30} 알킬, 선택적으로 치환된 C_{10}-C_{30} 알켄일, 선택적으로 치환된 C_{10}-C_{30} 알킨일 또는 선택적으로 치환된 C_{10}-C_{30} 아실이고;

[0410] R_3은 H, 선택적으로 치환된 C_2-C_{10} 알킬, 선택적으로 치환된 C_2-C_{10} 알켄일, 선택적으로 치환된 C_2-C_{10} 알킬릴, 알킬헤테로사이클, 알킬포스페이트, 알킬포스포로티오에이트, 알킬포스포로다이티오에이트, 알킬포스포네이트, 알킬아민, 하이드록시알킬, ω-아미노알킬, ω-(치환된)아미노알킬, ω-포스포알킬, ω-티오포스포알킬, 선택적으로 치환된 폴리에틸렌 글리콜(PEG, mw 100 내지 40K), 선택적으로 치환된 mPEG(mw 120 내지 40K), 헤테로아릴 또는 헤테로사이클 또는 링커 리간드이고, 예를 들어, 일부 실시형태에서, R_3은 $(CH_3)_2N(CH_2)_n$-이되, n은 1, 2, 3 또는 4이고;

[0411] E는 O, S, N(Q), C(O), OC(O), C(O)O, N(Q)C(O), C(O)N(Q), (Q)N(CO)O, O(CO)N(Q), S(O), $NS(O)_2N(Q)$, $S(O)_2$, $N(Q)S(O)_2$, SS, O=N, 아릴, 헤테로아릴, 사이클릭 또는 헤테로사이클, 예를 들어, -C(O)O이되, -는 R_3에 대한 연결 지점이고;

[0412] Q는 H, 알킬, ω-아미노알킬, ω-(치환된)아미노알킬, ω-포스포알킬 또는 ω-티오포스포알킬이다.

[0413] 구체적인 일 실시형태에서, 실시형태 1, 2, 3, 4 또는 5의 양이온성 지질은 하기 구조, 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체를 갖는다:

[0414] 

[0415] 식 중,

[0416] E는 O, S, N(Q), C(O), N(Q)C(O), C(O)N(Q), (Q)N(CO)O, O(CO)N(Q), S(O), $NS(O)_2N(Q)$, $S(O)_2$, $N(Q)_S(O)_2$, SS, O=N, 아릴, 헤테로아릴, 사이클릭 또는 헤테로사이클이고;

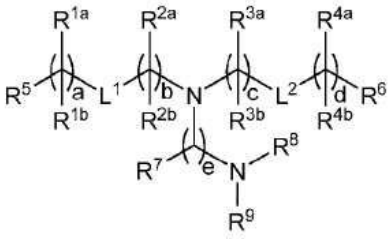
[0417] Q는 H, 알킬, ω-아미노알킬, ω-(치환된)아미노알킬, ω-포스포알킬 또는 ω-티오포스포알킬이고;

[0418] R_1 및 R_2 및 R_x는 각각 독립적으로 각각의 경우에 대해서 H, 선택적으로 치환된 C_1-C_{10} 알킬, 선택적으로 치환된 C_{10}-C_{30} 알킬, 선택적으로 치환된 C_{10}-C_{30} 알켄일, 선택적으로 치환된 C_{10}-C_{30}알킨일, 선택적으로 치환된 C_{10}-C_{30} 아실 또는 링커-리간드이되, 단, R_1, R_2 및 R_x 중 하나는 H가 아니고;

[0419] R_3은 H, 선택적으로 치환된 C_1-C_{10} 알킬, 선택적으로 치환된 C_2-C_{10} 알켄일, 선택적으로 치환된 C_2-C_{10} 알킨일, 알킬헤테로사이클, 알킬포스페이트, 알킬포스포로티오에이트, 알킬포스포로다이티오에이트, 알킬아민, 하이드록시알킬, ω-아미노알킬, ω-(치환된)아미노알킬, ω-포스포알킬, ω- 티오포스포알킬, 선택적으로 치환된 폴리에틸렌 글리콜(PEG, mw 100 내지 40K), 선택적으로 치환된 mPEG(mw 120 내지 40K), 헤테로아릴, 또는 헤테로사이클 또는 링커-리간드이고;

[0420] n은 0, 1, 2 또는 3이고이다.

[0421] 일 실시형태에서, 실시형태 1, 2, 3, 4 또는 5의 양이온성 지질은 하기 화학식 I의 구조 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체를 갖는다:



I

[0422]

[0423] 식 중,

[0424] L^1 또는 L^2 중 하나는 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_X-$, -S-S-, -C(=O)S-, SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 또는 $-NR^aC(=O)O-$이고, L^1 또는 L^2 중 나머지는 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_X-$, -S-S-, -C(=O)S-, SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 또는 $-NR^aC(=O)O-$ 또는 직접 결합이고;

[0425] R^a는 H 또는 C_1-C_{12} 알킬이고;

[0426] R^{1a} 및 R^{1b}는 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{1a}는 H 또는 C_1-C_{12} 알킬이고, R^{1b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{1b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0427] R^{2a} 및 R^{2b}는 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{2a}는 H 또는 C_1-C_{12} 알킬이고, R^{2b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{2b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0428] R^{3a} 및 R^{3b}는 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{3a}는 H 또는 C_1-C_{12} 알킬이고, R^{3b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{3b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0429] R^{4a} 및 R^{4b}는 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{4a}는 H 또는 C_1-C_{12} 알킬이고, R^{4b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{4b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0430] R^5 및 R^6은 각각 독립적으로 메틸 또는 사이클로알킬이고;

[0431] R^7은, 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0432] R^8 및 R^9는 각각 독립적으로 비치환된 C_1-C_{12} 알킬이거나; R^8과 R^9는, 이들이 부착되는 질소 원자와 함께 1개의 질소 원자를 포함하는 5, 6 또는 7-원 헤테로사이클릭 고리를 형성하고;

[0433] a 및 d는 각각 독립적으로 0 내지 24의 정수이고;

[0434] b 및 c는 각각 독립적으로 1 내지 24의 정수이고;

[0435] e는 1 또는 2이고;

[0436] x는 0, 1 또는 2이다.

[0437] 화학식 I의 일부 실시형태에서, L^1 및 L^2는 독립적으로- O(C=O)- 또는 -(C=O)O-이다.

[0438] 화학식 I의 특정 실시형태에서, R^{1a}, R^{2a}, R^{3a} 또는 R^{4a} 중 적어도 하나는 C_1-C_{12} 알킬이거나, L^1 또는 L^2 중 적어도 하나는 -O(C=O)- 또는 -(C=O)O-이다. 다른 실시형태에서, a가 6인 경우 R^{1a} 및 R^{1b}는 아이소프로필이 아니거나 a가 8인 경우 n-부틸이 아니다.

[0439] 화학식 I의 더 추가의 실시형태에서, R^{1a}, R^{2a}, R^{3a} 또는 R^{4a} 중 적어도 하나는 C_1-C_{12} 알킬이거나, L^1 또는 L^2 중 적어도 하나는 -O(C=O)- 또는 -(C=O)O-이고;

[0440] a가 6인 경우 R^{1a} 및 R^{1b}는 아이소프로필이 아니거나 a가 8인 경우 n-부틸이 아니다.

[0441] 화학식 I의 다른 실시형태에서, R^8 및 R^9는 각각 독립적으로 비치환된 C_1-C_{12} 알킬이거나; R^8과 R^9는, 이들이 부착되는 질소 원자와 함께 1개의 질소 원자를 포함하는 5, 6 또는 7-원 헤테로사이클릭 고리를 형성하고;

[0442] 화학식 I의 특정 실시형태에서, L^1 또는 L^2 중 임의의 하나는 -O(C=O)- 또는 탄소-탄소 이중 결합일 수 있다. L^1 및 L^2는 각각 -O(C=O)-일 수 있거나 각각 탄소-탄소 이중 결합일 수 있다.

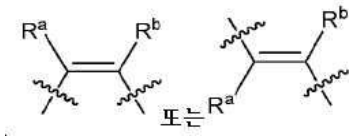
[0443] 화학식 I의 일부 실시형태에서, L^1 또는 L^2 중 하나는 -O(C=O)-이다. 다른 실시형태에서, L^1 및 L^2 둘 다는 -O(C=O)-이다.

[0444] 화학식 I의 일부 실시형태에서, L^1 또는 L^2 중 하나는 -(C=O)O-이다. 다른 실시형태에서, L^1 및 L^2 둘 다는 -(C=O)O-이다.

[0445] 화학식 I의 일부 다른 실시형태에서, L^1 또는 L^2 중 하나는 탄소-탄소 이중 결합이다. 다른 실시형태에서, L^1 및 L^2 둘 다는 탄소-탄소 이중 결합이다.

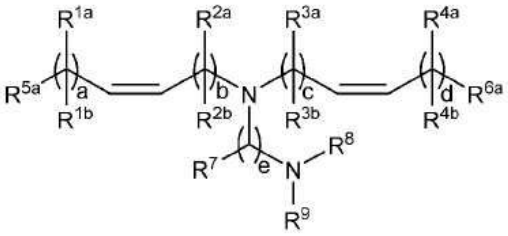
[0446] 화학식 I의 추가의 다른 실시형태에서, L^1 또는 L^2 중 하나는 -O(C=O)-이고, L^1 또는 L^2 중 나머지는 -(C=O)O-이다. 추가의 실시형태에서, L^1 또는 L^2 중 하나는 -O(C=O)-이고, L^1 또는 L^2 중 나머지는 탄소-탄소 이중 결합이다. 더 추가의 실시형태에서, L^1 또는 L^2 중 하나는 -(C=O)O-이고, L^1 또는 L^2 중 나머지는 탄소-탄소 이중 결합이다.

[0447] 본 명세서 전체에서 사용되는 바와 같은 "탄소-탄소" 이중 결합은 하기 구조 중 하나를 지칭한다는 것이 이해된다:

[0448] 

[0449] 식 중, R^a 및 R^b는 각각의 경우에, 독립적으로 H 또는 치환체이다. 예를 들어, 일부 실시형태에서 R^a 및 R^b는, 각각의 경우에, 독립적으로 H, C_1-C_{12} 알킬 또는 사이클로알킬, 예를 들어, H 또는 C_1-C_{12} 알킬이다.

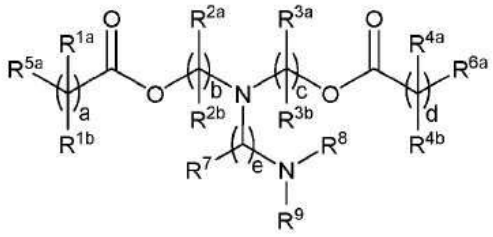
[0450] 다른 실시형태에서, 화학식 I의 지질 화합물은 하기 화학식 (Ia)를 갖는다:



(Ia)

[0451]

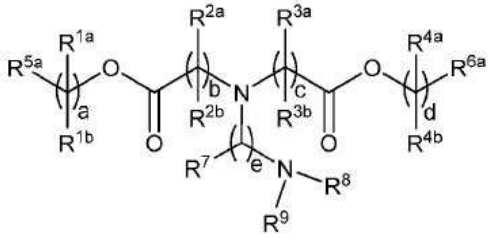
[0452] 다른 실시형태에서, 화학식 I의 지질 화합물은 하기 화학식 (Ib)를 갖는다:



(Ib)

[0453]

[0454] 추가의 다른 실시형태에서, 화학식 I의 지질 화합물은 하기 화학식 (Ic)를 갖는다:



(Ic)

[0455]

[0456] 화학식 I의 지질 화합물의 특정 실시형태에서, a, b, c 및 d는 각각 독립적으로 2 내지 12의 정수 또는 4 내지 12의 정수이다. 다른 실시형태에서, a, b, c 및 d는 각각 독립적으로 8 내지 12 또는 5 내지 9의 정수이다. 일부 특정 실시형태에서, a는 0이다. 일부 실시형태에서, a는 1이다. 다른 실시형태에서, a는 2이다. 추가의 실시형태에서, a는 3이다. 추가의 다른 실시형태에서, a는 4이다. 일부 실시형태에서, a는 5이다. 다른 실시형태에서, a는 6이다. 추가의 실시형태에서, a는 7이다. 추가의 다른 실시형태에서, a는 8이다. 일부 실시형태에서, a는 9이다. 다른 실시형태에서, a는 10이다. 추가의 실시형태에서, a는 11이다. 추가의 다른 실시형태에서, a는 12이다. 일부 실시형태에서, a는 13이다. 다른 실시형태에서, a는 14이다. 추가의 실시형태에서, a는 15이다. 추가의 다른 실시형태에서, a는 16이다.

[0457] 화학식 I의 일부 다른 실시형태에서, b는 1이다. 다른 실시형태에서, b는 2이다. 추가의 실시형태에서, b는 3이다. 추가의 다른 실시형태에서, b는 4이다. 일부 실시형태에서, b는 5이다. 다른 실시형태에서, b는 6이다. 추가의 실시형태에서, b는 7이다. 추가의 다른 실시형태에서, b는 8이다.

[0458] 일부 실시형태에서, b는 9이다. 다른 실시형태에서, b는 10이다. 추가의 실시형태에서, b는 11이다. 추가의 다른 실시형태에서, b는 12이다. 일부 실시형태에서, b는 13이다. 다른 실시형태에서, b는 14이다. 추가의 실시형태에서, b는 15이다. 추가의 다른 실시형태에서, b는 16이다.

[0459] 화학식 I의 일부 추가의 실시형태에서, c는 1이다. 다른 실시형태에서, c는 2이다. 추가의 실시형태에서, c는 3이다. 추가의 다른 실시형태에서, c는 4이다. 일부 실시형태에서, c는 5이다. 다른 실시형태에서, c는 6이다. 추가의 실시형태에서, c는 7이다. 추가의 다른 실시형태에서, c는 8이다. 일부 실시형태에서, c는 9이다. 다른

실시형태에서, c는 10이다. 추가의 실시형태에서, c는 11이다. 추가의 다른 실시형태에서, c는 12이다. 일부 실시형태에서, c는 13이다. 다른 실시형태에서, c는 14이다. 추가의 실시형태에서, c는 15이다. 추가의 다른 실시형태에서, c는 16이다.

[0460] 화학식 I의 일부 특정 다른 실시형태에서, d는 0이다. 일부 실시형태에서, d는 1이다. 다른 실시형태에서, d는 2이다. 추가의 실시형태에서, d는 3이다. 추가의 다른 실시형태에서, d는 4이다. 일부 실시형태에서, d는 5이다. 다른 실시형태에서, d는 6이다. 추가의 실시형태에서, d는 7이다. 추가의 다른 실시형태에서, d는 8이다. 일부 실시형태에서, d는 9이다. 다른 실시형태에서, d는 10이다. 추가의 실시형태에서, d는 11이다. 추가의 다른 실시형태에서, d는 12이다. 일부 실시형태에서, d는 13이다. 다른 실시형태에서, d는 14이다. 추가의 실시형태에서, d는 15이다. 추가의 다른 실시형태에서, d는 16이다.

[0461] 화학식 I의 일부 다른 다양한 실시형태에서, a 및 d는 동일하다. 일부 다른 실시형태에서, b 및 c는 동일하다. 일두 다른 구체적인 실시형태에서, a 및 d는 동일하고, b 및 c는 동일하다.

[0462] 화학식 I에서 a와 b의 합 및 c와 d의 합은 목적하는 특성을 갖는 화학식 I의 지질을 얻기 위해서 달라질 수 있는 인자이다. 일 실시형태에서, a 및 b는 이들의 합이 14 내지 24의 범위의 정수이도록 선택된다. 다른 실시형태에서, c 및 d는 이들의 합이 14 내지 24의 범위의 정수이도록 선택된다. 추가 실시형태에서, a와 b의 합 및 c와 d의 합은 동일하다. 예를 들어, 일부 실시형태에서 a와 b의 합 및 c와 d의 합 둘 다는 14 내지 24의 범위일 수 있는 동일한 정수이다. 더 추가의 실시형태에서, a, b, c 및 d는 a와 b의 합 및 c와 d의 합이 12 이상이도록 선택된다.

[0463] 화학식 I의 일부 실시형태에서, e는 1이다. 다른 실시형태에서, e는 2이다.

[0464] 화학식 I의 R^{1a}, R^{2a}, R^{3a} 및 R^{4a}에서 치환체는 특별히 제한되지 않는다. 특정 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a}는 각각의 경우에 H이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_{12} 알킬이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_8 알킬이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_6 알킬이다. 상기 실시형태 중 일부에서, C_1-C_8 알킬은 메틸, 에틸, n-프로필, 아이소-프로필, n-부틸, 아이소-부틸, tert-부틸, n-헥실 또는 n-옥틸이다.

[0465] 화학식 I의 특정 실시형태에서, R^{1a}, R^{1b}, R^{4a} 및 R^{4b}는 각각의 경우에 C_1-C_{12} 알킬이다.

[0466] 화학식 I의 추가 실시형태에서, R^{1b}, R^{2b}, R^{3b} 및 R^{4b} 중 적어도 하나는 H이거나 R^{1b}, R^{2b}, R^{3b} 및 R^{4b}는 각각의 경우에 H이다.

[0467] 화학식 I의 특정 실시형태에서, R^{1b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{1b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다. 상기 중 다른 실시형태에서, R^{4b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{4b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0468] 화학식 I의 R^5 및 R^6에서 치환체는 상기 실시형태에서 특별히 제한되지 않는다. 특정 실시형태에서 R^5 또는 R^6 중 하나 또는 둘 다는 메틸이다. 특정 실시형태에서 R^5 또는 R^6 중 하나 또는 둘 다는 사이클로알킬 예를 들어, 사이클로헥실이다. 이들 실시형태에서 사이클로알킬은 치환될 수 있거나 치환되지 않을 수 있다. 특정 실시형태에서 사이클로알킬은 C_1-C_{12} 알킬, 예를 들어, tert-부틸로 치환된다.

[0469] R^7에서 치환체는 화학식 I의 상기 실시형태에서 특별히 제한되지 않는다. 특정 실시형태에서 적어도 하나의 R^7은 H이다. 일부 다른 실시형태에서, R^7은 각각의 경우에 H이다. 특정 실시형태에서 R^7은 C_1-C_{12} 알킬이다.

[0470] 화학식 I의 상기 실시형태 중 특정 다른 것에서, R^8 또는 R^9 중 하나는 메틸이다. 다른 실시형태에서, R^8 및 R^9 둘 다는 메틸이다.

[0471] 화학식 I의 일부 상이한 실시형태에서, R^8 및 R^9는 이들이 부착되는 질소 원자와 함께 5, 6 또는 7-원 헤테로사이클릭 고리를 형성한다. 상기 중 일부 실시형태에서, R^8 및 R^9는 이들이 부착되는 질소 원자와 함께 5-원 헤테

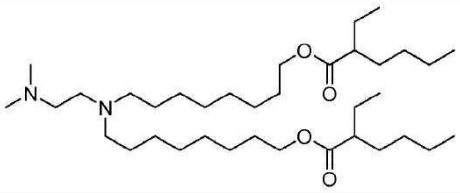
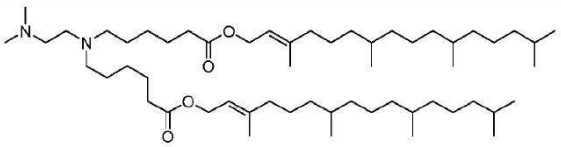
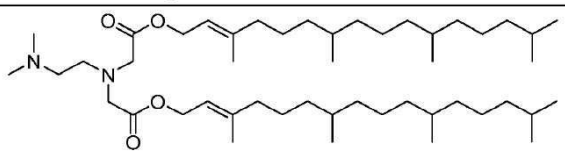
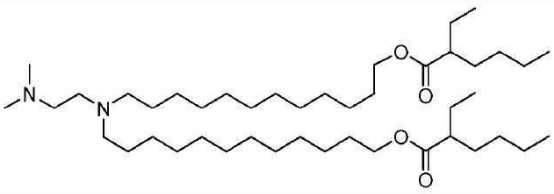
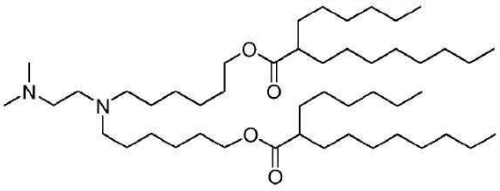
로사이클릭 고리, 예를 들어, 피롤리딘일 고리를 형성한다.

[0472] 실시형태 3의 일부 실시형태에서, 제1 스페이서 및 제2 스페이서양이온성 지질은 하기 화학식 I의 지질로부터 각각 독립적으로 선택된다.

[0473] 다양한 상이한 실시형태에서, 화학식 I의 지질은 하기 표 1에 제시된 구조 중 하나를 갖는다

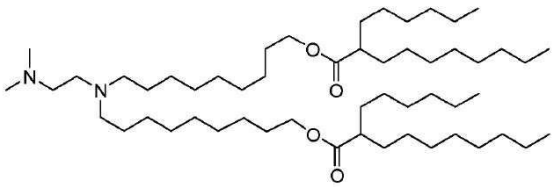
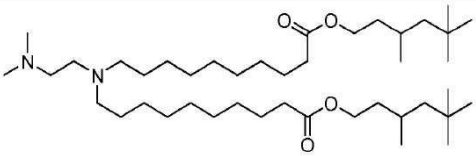
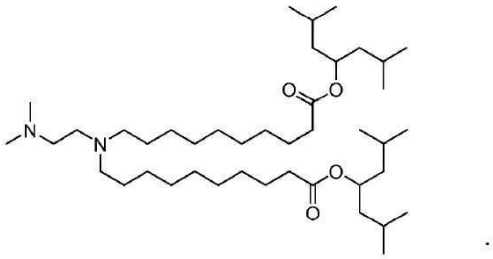
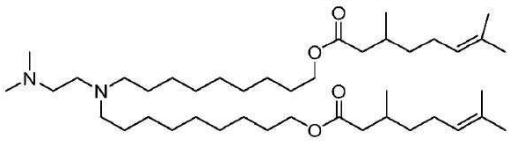
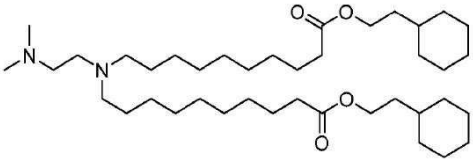
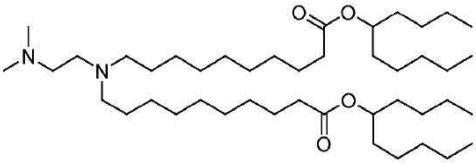
표 1

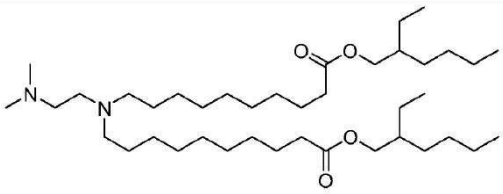
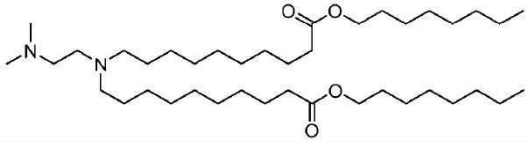
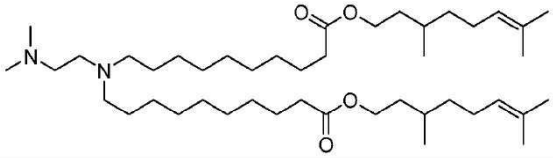
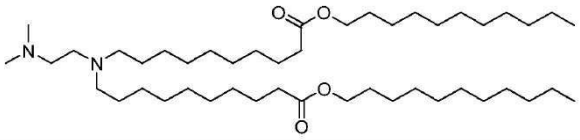
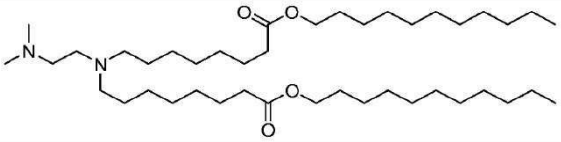
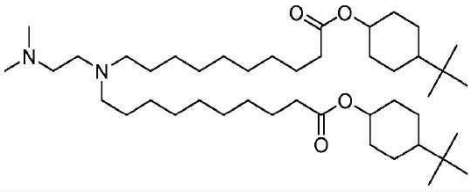
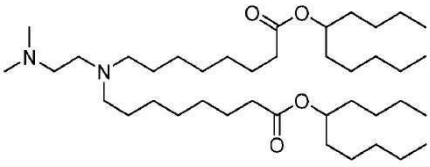
화학식 I의 대표적인 지질

번호	구조	pKa
I-1		-
I-2		5.64
I-3		7.15
I-4		6.43
I-5		6.28

[0474]

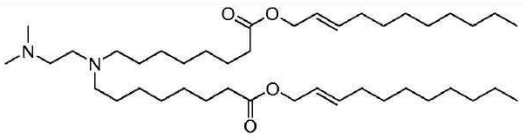
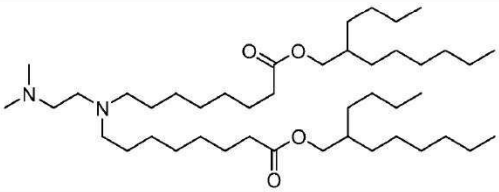
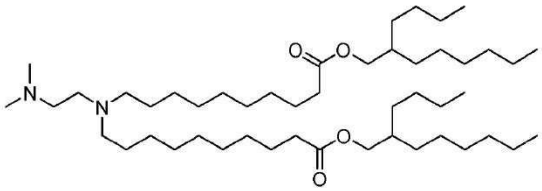
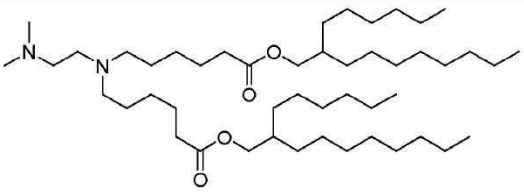
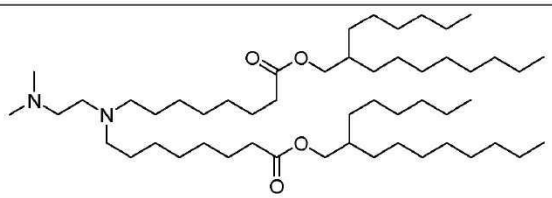
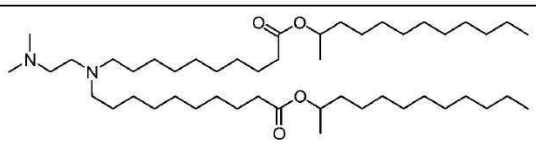
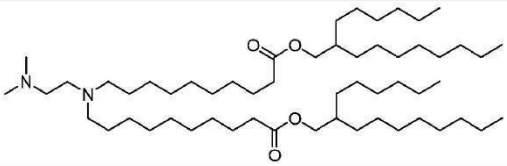
[0475]

번호	구조	pKa
I-6		6.12
I-7		-
I-8		-
I-9		-
I-10		-
I-11		6.36

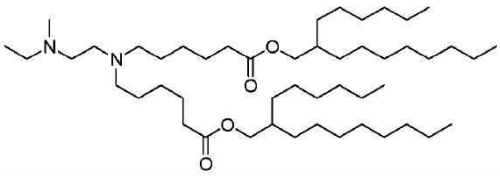
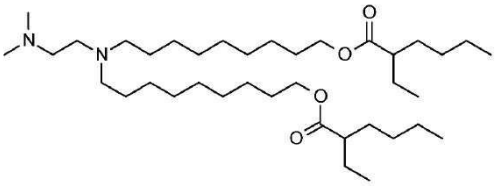
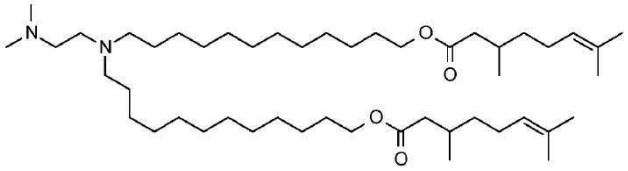
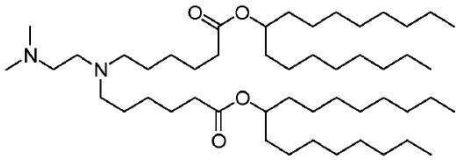
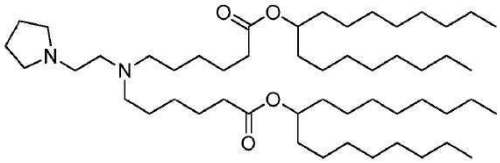
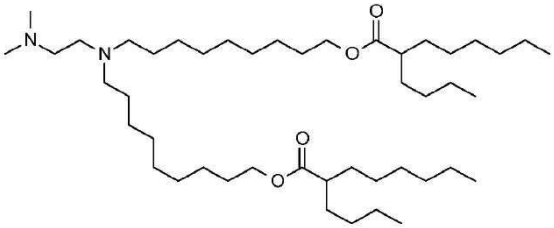
번호	구조	pKa
I-12		-
I-13		6.51
I-14		-
I-15		6.30
I-16		6.63
I-17		-
I-18		-

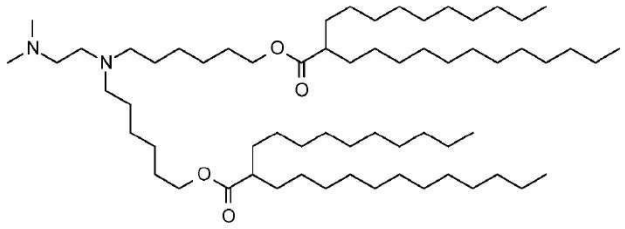
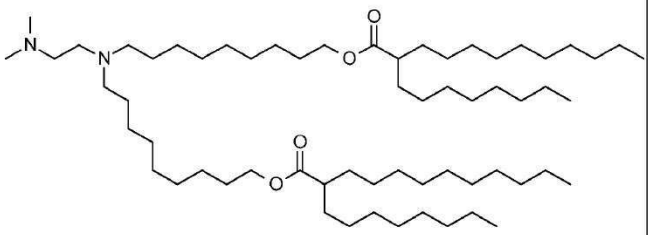
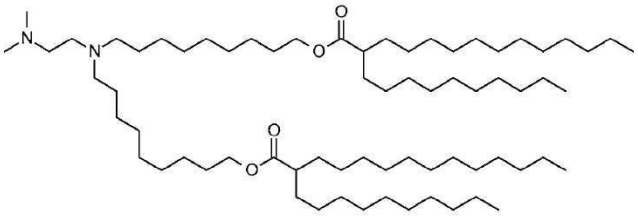
[0476]

[0477]

번호	구조	pKa
I-19		6.72
I-20		6.44
I-21		6.28
I-22		6.53
I-23		6.24
I-24		6.28
I-25		6.20

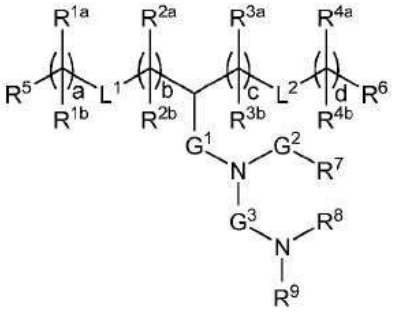
[0478]

번호	구조	pKa
I-33		6.27
I-34		-
I-35		6.21
I-36		-
I-37		-
I-38		6.24

번호	구조	pKa
I-39		5.82
I-40		6.38
I-41		5.91

[0479]

[0480] 일부 실시형태에서, 실시형태 1, 2, 3, 4 또는 5의 양이온성 지질은 하기 화학식 II의 구조, 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체를 갖는다:



II

[0481]

[0482] 식 중,

[0483] L^1 또는 L^2 중 하나는 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_X-$, -S-S-, -C(=O)S-, SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 또는 $-NR^aC(=O)O-$이고, L^1 또는 L^2 중 나머지는 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_X-$, -S-S-, -C(=O)S-, SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 또는 $-NR^aC(=O)O-$ 또는 직접 결합이고;

[0484] G^1은 C_1-C_2 알킬렌, -(C=O)-, -O(C=O)-, -SC(=O)-, $-NR^aC(=O)-$ 또는 직접 결합이고;

[0485] G^2는 -C(=O)-, -(C=O)O-, -C(=O)S-, $-C(=O)NR^a-$ 또는 직접 결합이고;

[0486] G^3은 C_1-C_6 알킬렌이고;

[0487] R^a는 H 또는 C_1-C_{12} 알킬이고;

[0488] R^{1a} 및 R^{1b}는 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{1a}는 H 또는 C_1-C_{12} 알킬이고, R^{1b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{1b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0489] R^{2a} 및 R^{2b}는 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{2a}는 H 또는 C_1-C_{12} 알킬이고, R^{2b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{2b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0490] R^{3a} 및 R^{3b}는 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{3a}는 H 또는 C_1-C_{12} 알킬이고, R^{3b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{3b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0491] R^{4a} 및 R^{4b}는 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{4a}는 H 또는 C_1-C_{12} 알킬이고, R^{4b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{4b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0492] R^5 및 R^6은 각각 독립적으로 H 또는 메틸이고;

[0493] R^7은 C_4-C_{20} 알킬이고;

[0494] R^8 및 R^9는 각각 독립적으로 C_1-C_{12} 알킬이거나; R^8과 R^9는, 이들이 부착되는 질소 원자와 함께 5, 6 또는 7-원 헤테로사이클릭 고리를 형성하고;

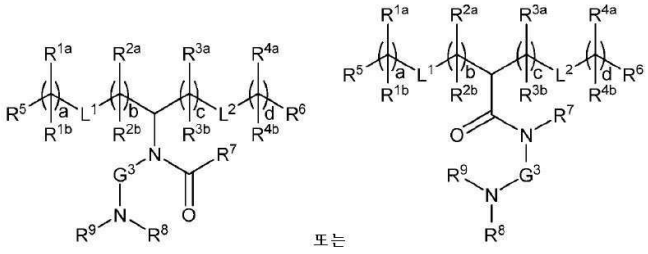
[0495] a, b, c 및 d는 각각 독립적으로 1 내지 24의 정수이고;

[0496] x는 0, 1 또는 2이다.

[0497] 화학식 (II)의 일부 실시형태에서, L^1 및 L^2는 각각 독립적으로 -O(C=O)-, -(C=O)O- 또는 직접 결합이다. 다른 실시형태에서, G^1 및 G^2는 각각 독립적으로 -(C=O)- 또는 직접 결합이다. 일부 상이한 실시형태에서, L^1 및 L^2는 각각 독립적으로 -O(C=O)-, -(C=O)O- 또는 직접 결합이고; G^1 및 G^2는 각각 독립적으로 -(C=O)- 또는 직접 결합이다.

[0498] 화학식 (II)의 일부 상이한 실시형태에서, L^1 및 L^2는 각각 독립적으로 -C(=O)-, -O-, $-S(O)_x-$, -S-S-, -C(=O)S-, -SC(=O)-, $-NR^a-$, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $-NR^aC(=O)NR^a$, $-OC(=O)NR^a-$, $-NR^aC(=O)O-$, $-NR^aS(O)_xNR^a-$, $-NR^aS(O)_x-$ 또는 $-S(O)_xNR^a-$이다.

[0499] 화학식 (II)의 상기 실시형태 중 다른 것에서, 지질 화합물은 하기 화학식 (IIA) 또는 (IIB) 중 하나를 갖는다:



(IIA) (IIB)

[0500]

[0501] 화학식 (II)의 일부 실시형태에서, 지질 화합물은 화학식 (IIA)를 갖는다. 다른 실시형태에서, 지질 화합물은 화학식 (IIB)를 갖는다.

[0502] 화학식 (II)의 상기 실시형태 중 임의의 것에서, L^1 또는 L^2 중 하나는 -O(C=O)-이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 -O(C=O)-이다.

[0503] 화학식 (II)의 일부 상이한 실시형태에서, L^1 또는 L^2 중 하나는 -(C=O)O-이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 -(C=O)O-이다.

[0504] 화학식 (II)의 상이한 실시형태에서, L^1 또는 L^2 중 하나는 직접 결합이다. 본 명세서에서 사용되는 바와 같이, "직접 결합"은 기(예를 들어, L^1 또는 L^2)가 존재하지 않는다는 것을 의미한다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 직접 결합이다.

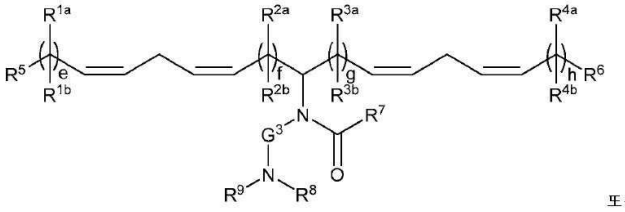
[0505] 화학식 (II)의 다른 상이한 실시형태에서, R^{1a} 및 R^{1b} 중 적어도 하나의 경우에 대해서, R^{1a}는 H 또는 C_1-C_{12} 알킬이고, R^{1b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{1b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0506] 화학식 (II)의 추가의 다른 상이한 실시형태에서, R^{4a} 및 R^{4b} 중 적어도 하나의 경우에 대해서, R^{4a}는 H 또는 C_1-C_{12} 알킬이고, R^{4b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{4b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0507] 화학식 (II)의 추가의 실시형태에서, R^{2a} 및 R^{2b} 중 적어도 하나의 경우에 대해서, R^{2a}는 H 또는 C_1-C_{12} 알킬이고, R^{2b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{2b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

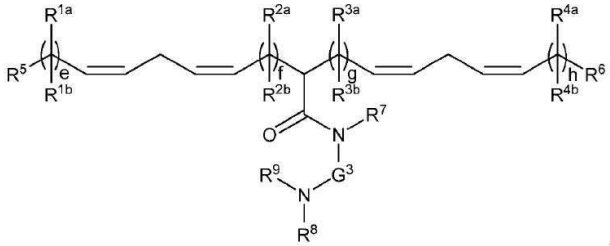
[0508] 화학식 (II)의 다른 상이한 실시형태에서, R^{3a} 및 R^{3b} 중 적어도 하나의 경우에 대해서, R^{3a}는 H 또는 C_1-C_{12} 알킬이고, R^{3b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{3b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0509] 화학식 (II)의 다양한 다른 실시형태에서, 지질 화합물은 하기 화학식 (IIC) 또는 (IID) 중 하나를 갖는다:



또는

(IIC)



,

(IID)

[0510]

[0511] 식 중, e, f, g 및 h는 각각 독립적으로 1 내지 12의 정수이다.

[0512] 화학식 (II)의 일부 실시형태에서, 지질 화합물은 화학식 (IIC)를 갖는다. 다른 실시형태에서, 지질 화합물은 화학식 (IID)를 갖는다.

[0513] 화학식 (IIC) 또는 (IID)의 다양한 실시형태에서, e, f, g 및 h는 각각 독립적으로 4 내지 10의 정수이다.

[0514] 화학식 (II)의 특정 실시형태에서, a, b, c 및 d는 각각 독립적으로 2 내지 12의 정수 또는 4 내지 12의 정수이다. 다른 실시형태에서, a, b, c 및 d는 각각 독립적으로 8 내지 12 또는 5 내지 9의 정수이다. 일부 특정 실시형태에서, a는 0이다. 일부 실시형태에서, a는 1이다. 다른 실시형태에서, a는 2이다. 추가의 실시형태에서, a는 3이다. 추가의 다른 실시형태에서, a는 4이다. 일부 실시형태에서, a는 5이다. 다른 실시형태에서, a는 6이다. 추가의 실시형태에서, a는 7이다. 추가의 다른 실시형태에서, a는 8이다. 일부 실시형태에서, a는 9이다. 다른 실시형태에서, a는 10이다. 추가의 실시형태에서, a는 11이다. 추가의 다른 실시형태에서, a는 12이다. 일부 실시형태에서, a는 13이다.

[0515] 다른 실시형태에서, a는 14이다. 추가의 실시형태에서, a는 15이다. 추가의 다른 실시형태에서, a는 16이다.

[0516] 화학식 (II)의 일부 실시형태에서, b는 1이다. 다른 실시형태에서, b는 2이다. 추가의 실시형태에서, b는 3이다. 추가의 다른 실시형태에서, b는 4이다. 일부 실시형태에서, b는 5이다. 다른 실시형태에서, b는 6이다. 추가의 실시형태에서, b는 7이다. 추가의 다른 실시형태에서, b는 8이다.

[0517] 일부 실시형태에서, b는 9이다. 다른 실시형태에서, b는 10이다. 추가의 실시형태에서, b는 11이다. 추가의 다른 실시형태에서, b는 12이다. 일부 실시형태에서, b는 13이다. 다른 실시형태에서, b는 14이다. 추가의 실시형태에서, b는 15이다. 추가의 다른 실시형태에서, b는 16이다.

[0518] 화학식 (II)의 일부 실시형태에서, c는 1이다. 다른 실시형태에서, c는 2이다. 추가의 실시형태에서, c는 3이다. 추가의 다른 실시형태에서, c는 4이다. 일부 실시형태에서, c는 5이다. 다른 실시형태에서, c는 6이다. 추가의 실시형태에서, c는 7이다. 추가의 다른 실시형태에서, c는 8이다. 일부 실시형태에서, c는 9이다. 다른 실시형태에서, c는 10이다. 추가의 실시형태에서, c는 11이다. 추가의 다른 실시형태에서, c는 12이다. 일부 실시형태에서, c는 13이다. 다른 실시형태에서, c는 14이다. 추가의 실시형태에서, c는 15이다. 추가의 다른 실시형태에서, c는 16이다.

[0519] 화학식 (II)의 일부 특정 실시형태에서, d는 0이다. 일부 실시형태에서, d는 1이다. 다른 실시형태에서, d는 2이다. 추가의 실시형태에서, d는 3이다. 추가의 다른 실시형태에서, d는 4이다. 일부 실시형태에서, d는 5이다. 다른 실시형태에서, d는 6이다. 추가의 실시형태에서, d는 7이다. 추가의 다른 실시형태에서, d는 8이다. 일부 실시형태에서, d는 9이다. 다른 실시형태에서, d는 10이다. 추가의 실시형태에서, d는 11이다. 추가의 다른 실시형태에서, d는 12이다.

[0520] 일부 실시형태에서, d는 13이다. 다른 실시형태에서, d는 14이다. 추가의 실시형태에서, d는 15이다. 추가의 다른 실시형태에서, d는 16이다.

[0521] 화학식 (II)의 일부 실시형태에서, e는 1이다. 다른 실시형태에서, e는 2이다. 추가의 실시형태에서, e는 3이다. 추가의 다른 실시형태에서, e는 4이다. 일부 실시형태에서, e는 5이다. 다른 실시형태에서, e는 6이다. 추가의 실시형태에서, e는 7이다. 추가의 다른 실시형태에서, e는 8이다. 일부 실시형태에서, e는 9이다. 다른 실시형태에서, e는 10이다. 추가의 실시형태에서, e는 11이다. 추가의 다른 실시형태에서, e는 12이다.

[0522] 화학식 (II)의 일부 실시형태에서, f는 1이다. 다른 실시형태에서, f는 2이다. 추가의 실시형태에서, f는 3이다. 추가의 다른 실시형태에서, f는 4이다. 일부 실시형태에서, f는 5이다. 다른 실시형태에서, f는 6이다. 추가의 실시형태에서, f는 7이다. 추가의 다른 실시형태에서, f는 8이다. 일부 실시형태에서, f는 9이다. 다른 실시형태에서, f는 10이다. 추가의 실시형태에서, f는 11이다. 추가의 다른 실시형태에서, f는 12이다.

[0523] 화학식 (II)의 일부 실시형태에서, g는 1이다. 다른 실시형태에서, g는 2이다. 추가의 실시형태에서, g는 3이다. 추가의 다른 실시형태에서, g는 4이다. 일부 실시형태에서, g는 5이다. 다른 실시형태에서, g는 6이다. 추가의 실시형태에서, g는 7이다. 추가의 다른 실시형태에서, g는 8이다. 일부 실시형태에서, g는 9이다. 다른 실시형태에서, g는 10이다. 추가의 실시형태에서, g는 11이다. 추가의 다른 실시형태에서, g는 12이다.

[0524] 화학식 (II)의 일부 실시형태에서, h는 1이다. 다른 실시형태에서, e는 2이다. 추가의 실시형태에서, h는 3이다. 추가의 다른 실시형태에서, h는 4이다. 일부 실시형태에서, e는 5이다. 다른 실시형태에서, h는 6이다. 추가의 실시형태에서, h는 7이다. 추가의 다른 실시형태에서, h는 8이다. 일부 실시형태에서, h는 9이다. 다른

실시형태에서, h는 10이다. 추가의 실시형태에서, h는 11이다. 추가의 다른 실시형태에서, h는 12이다.

[0525] 화학식 (II)의 일부 다른 다양한 실시형태에서, a 및 d는 동일하다. 일부 다른 실시형태에서, b 및 c는 동일하다. 일두 다른 구체적인 실시형태에서, a 및 d는 동일하고, b 및 c는 동일하다.

[0526] 화학식 (II)에서 a와 b의 합 및 c와 d의 합은 목적하는 특성을 갖는 지질을 얻기 위해서 달라질 수 있는 인자이다. 일 실시형태에서, a 및 b는 이들의 합이 14 내지 24의 범위의 정수이도록 선택된다. 다른 실시형태에서, c 및 d는 이들의 합이 14 내지 24의 범위의 정수이도록 선택된다. 추가 실시형태에서, a와 b의 합 및 c와 d의 합은 동일하다. 예를 들어, 일부 실시형태에서 a와 b의 합 및 c와 d의 합 둘 다는 14 내지 24의 범위일 수 있는 동일한 정수이다. 더 추가의 실시형태에서, a, b, c 및 d는 a와 b의 합 및 c와 d의 합이 12 이상이도록 선택된다.

[0527] 화학식 (II)의 R^{1a}, R^{2a}, R^{3a} 및 R^{4a}에서 치환체는 특별히 제한되지 않는다. 일부 실시형태에서, R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 H이다. 특정 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a}는 각각의 경우에 H이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_{12} 알킬이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_8 알킬이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_6 알킬이다. 상기 실시형태 중 일부에서, C_1-C_8 알킬은 메틸, 에틸, n-프로필, 아이소-프로필, n-부틸, 아이소-부틸, tert-부틸, n-헥실 또는 n-옥틸이다.

[0528] 화학식 (II)의 특정 실시형태에서, R^{1a}, R^{1b}, R^{4a} 및 R^{4b}는 각각의 경우에 C_1-C_{12} 알킬이다.

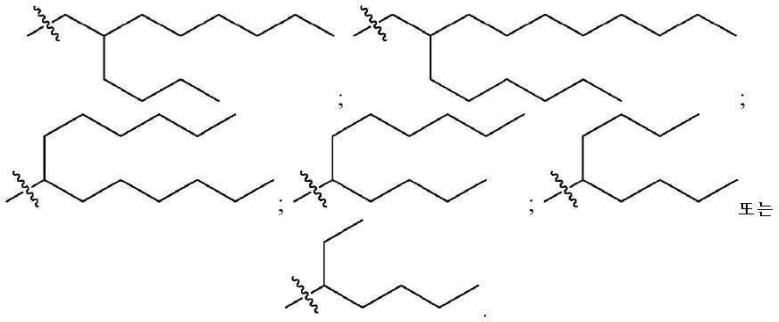
[0529] 화학식 (II)의 추가 실시형태에서, R^{1b}, R^{2b}, R^{3b} 및 R^{4b} 중 적어도 하나는 H이거나 R^{1b}, R^{2b}, R^{3b} 및 R^{4b}는 각각의 경우에 H이다.

[0530] 화학식 (II)의 특정 실시형태에서, R^{1b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{1b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다. 상기 중 다른 실시형태에서, R^{4b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{4b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0531] 화학식 (II)의 R^5 및 R^6에서 치환체는 상기 실시형태에서 특별히 제한되지 않는다. 특정 실시형태에서 R^5 또는 R^6 중 하나는 메틸이다. 다른 실시형태에서 R^5 또는 R^6 각각은 메틸이다.

[0532] 화학식 (II)의 R^7에서 치환체는 상기 실시형태에서 특별히 제한되지 않는다. 특정 실시형태에서 R^7은 C_6-C_{16} 알킬이다. 일부 다른 실시형태에서, R^7은 C_6-C_9 알킬이다. 이들 실시형태 중 일부에서, R^7은 $-(C=O)OR^b$, $-O(C=O)R^b$, $-C(=O)R^b$, $-OR^b$, $-S(O)_xR^b$, $-S-SR^b$, $-C(=O)SR^b$, $-SC(=O)R^b$, $-NR^aR^b$, $-NR^aC(=O)R^b$, $-C(=O)NR^aR^b$, $-NR^aC(=O)NR^aR^b$, $-OC(=O)NR^aR^b$, $-NR^aC(=O)OR^b$, $-NR^aS(O)_xNR^aR^b$, $-NR^aS(O)_xR^b$ 또는 $-S(O)_xNR^aR^b$로 치환되고, R^a는 H 또는 C_1-C_{12} 알킬이고; R^b는 C_1-C_{15} 알킬이고; x는 0, 1 또는 2이다. 예를 들어, 일부 실시형태에서 R^7은 $-(C=O)OR^b$ 또는 $-O(C=O)R^b$로 치환된다.

[0533] 화학식 (II)의 상기 실시형태 중 일부에서, R^b는 분지형 C_1-C_{16} 알킬이다. 예를 들어, 일부 실시형태에서 R^b는 하기 구조 중 하나를 갖는다:



[0534]

[0535] 화학식 (II)의 상기 실시형태 중 특정 다른 것에서, R^8 또는 R^9 중 하나는 메틸이다. 다른 실시형태에서, R^8 및 R^9 둘 다는 메틸이다.

[0536] 화학식 (II)의 일부 상이한 실시형태에서, R^8 및 R^9는 이들이 부착되는 질소 원자와 함께 5, 6 또는 7-원 헤테로사이클릭 고리를 형성한다.

[0537] 상기 중 일부 실시형태에서, R^8 및 R^9는 이들이 부착되는 질소 원자와 함께 5-원 헤테로사이클릭 고리, 예를 들어, 피롤리딘일 고리를 형성한다. 상기 중 일부 상이한 실시형태에서, R^8 및 R^9는 이들이 부착되는 질소 원자와 함께 6-원 헤테로사이클릭 고리, 예를 들어, 피페라진일 고리를 형성한다.

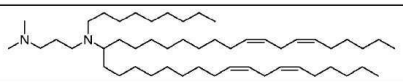
[0538] 실시형태 3의 특정 실시형태에서, 제1 스페이서 및 제2 스페이서양이온성 지질은 하기 화학식 II의 지질로부터 각각 독립적으로 선택된다.

[0539] 화학식 (II)의 상기 지질 중 추가의 다른 실시형태에서, G^3은 C_2-C_4 알킬렌, 예를 들어, C_3 알킬렌이다. 다양한 상이한 실시형태에서, 지질 화합물은 하기 표 2에 제시된 구조 중 하나를 갖는다.

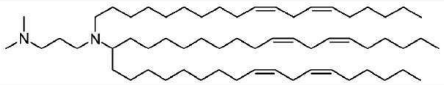
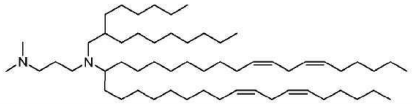
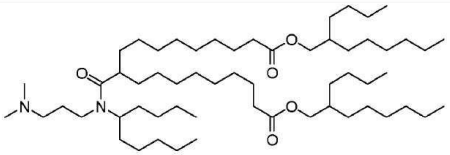
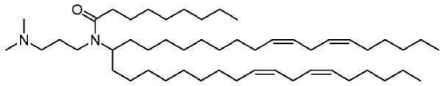
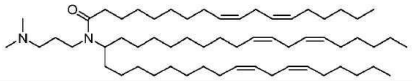
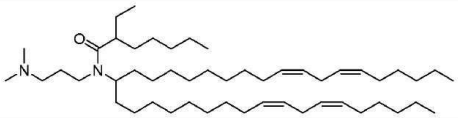
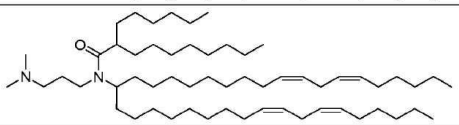
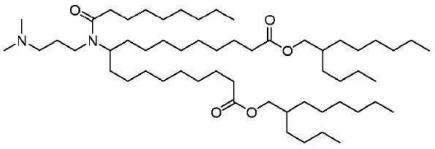
표 2

화학식 (II)의 대표적인 지질

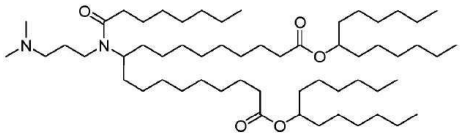
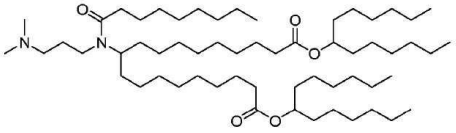
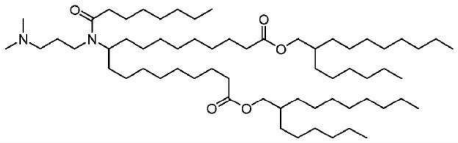
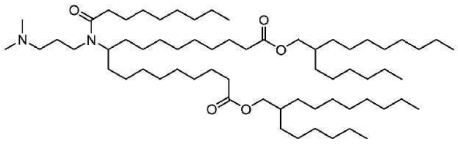
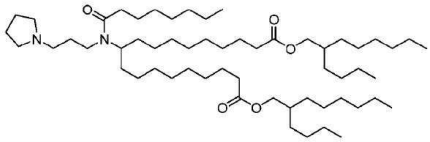
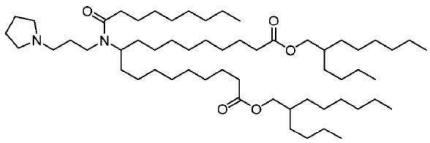
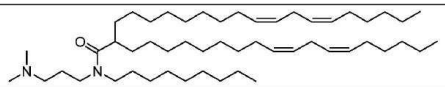
[0540]

번호	구조	pKa
II-1		5.64

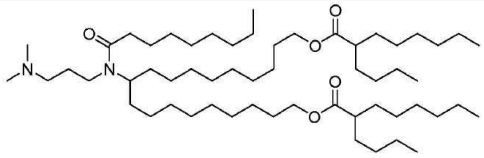
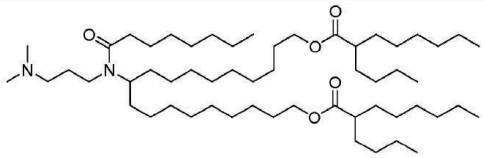
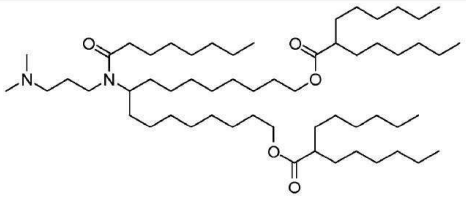
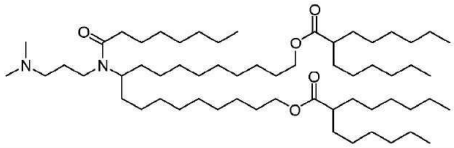
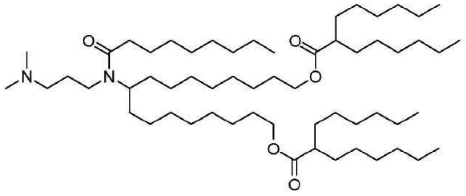
[0541]

번호	구조	pKa
II-2		-
II-3		-
II-4		-
II-5		6.27
II-6		6.14
II-7		5.93
II-8		5.35
II-9		6.27

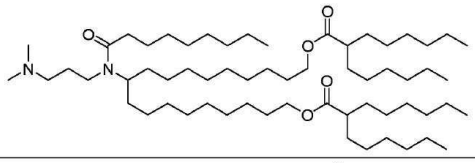
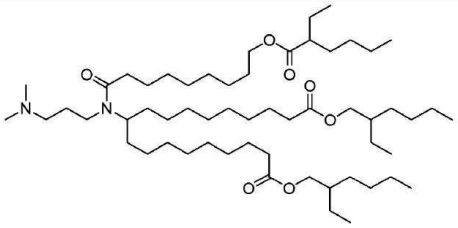
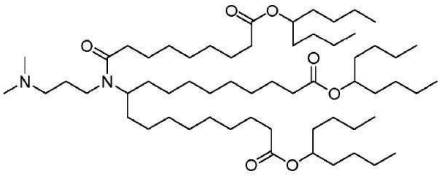
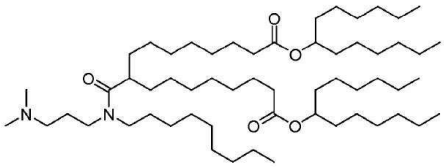
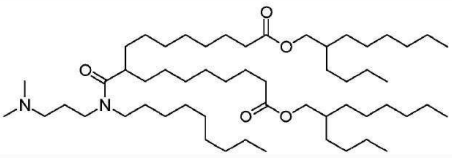
[0542]

번호	구조	pKa
II-10		6.16
II-11		6.13
II-12		6.21
II-13		6.22
II-14		6.33
II-15		6.32
II-16		6.37

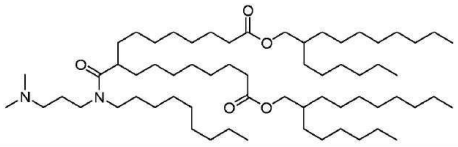
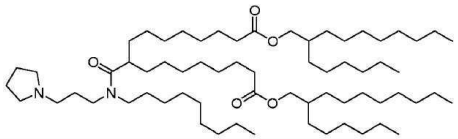
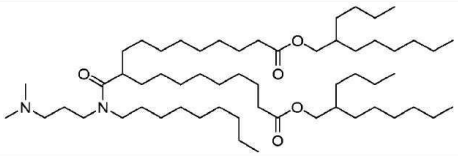
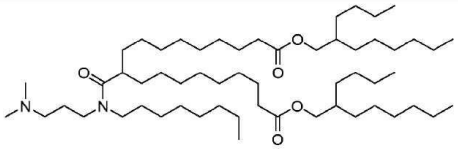
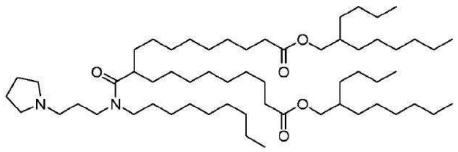
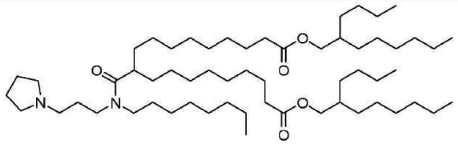
[0543]

번호	구조	pKa
II-17		6.27
II-18		-
II-19		-
II-20		-
II-21		-

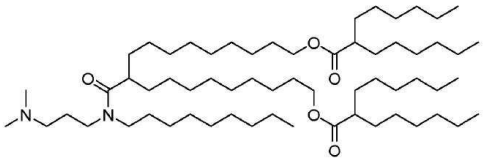
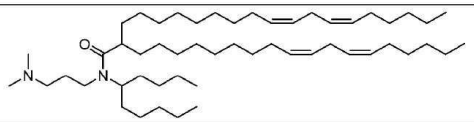
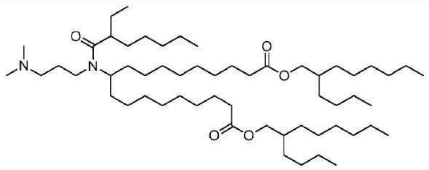
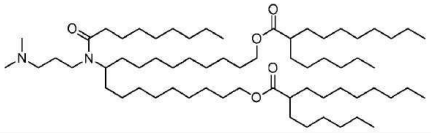
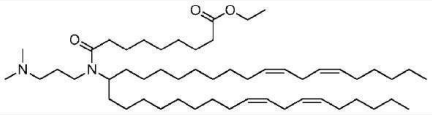
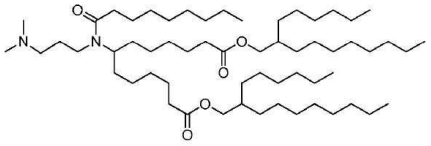
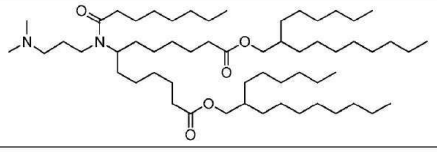
[0544]

번호	구조	pKa
II-22		-
II-23		-
II-24		6.14
II-25		-
II-26		-

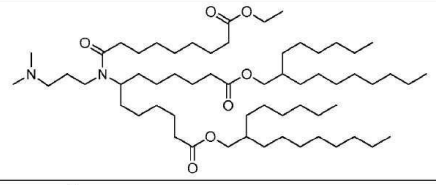
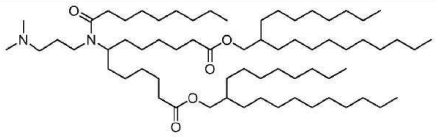
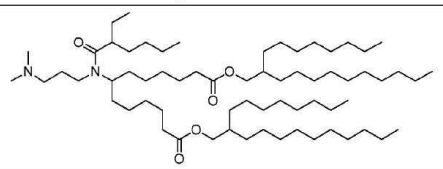
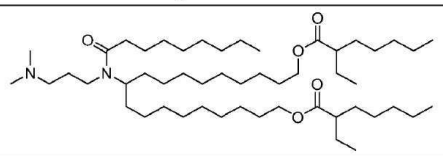
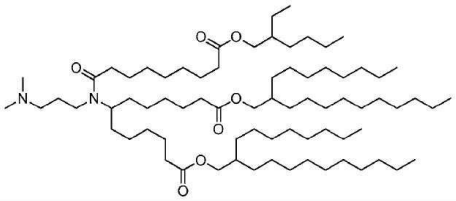
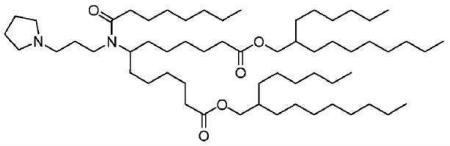
[0545]

번호	구조	pKa
II-27		-
II-28		-
II-29		-
II-30		-
II-31		-
II-32		-

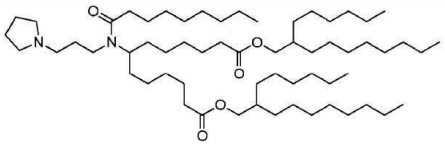
[0546]

번호	구조	pKa
II-33		-
II-34		-
II-35		5.97
II-36		6.13
II-37		5.61
II-38		6.45
II-39		6.45

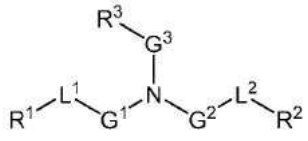
[0547]

번호	구조	pKa
II-40		6.57
II-41		-
II-42		-
II-43		-
II-44		-
II-45		-

[0548]

번호	구조	pKa
II-46		-

[0549] 일부 다른 실시형태에서, 실시형태 1, 2, 3, 4 또는 5의 양이온성 지질은 하기 화학식 III의 구조를 갖는다:

[0550] 

III

[0551] 또는 이의 약제학적으로 허용 가능한 염, 전구약물 또는 입체이성질체, 식 중, L^1 또는 L^2 중 하나는 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_X-$, -S-S-, -C(=O)S-, SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 또는 $-NR^aC(=O)O-$이고, L^1 또는 L^2 중 나머지는 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_X-$, -S-S-, -C(=O)S-, SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 또는 $-NR^aC(=O)O-$ 또는 직접 결합이고;

[0552] G^1 및 G^2는 각각 독립적으로 비치환된 C_1-C_{12} 알킬렌 또는 C_1-C_{12} 알켄일렌이고;

[0553] G^3은 C_1-C_{24} 알킬렌, C_1-C_{24} 알켄일렌, C_3-C_8 사이클로알킬렌, C_3-C_8 사이클로알켄일렌이고;

[0554] R^a는 H 또는 C_1-C_{12} 알킬이고;

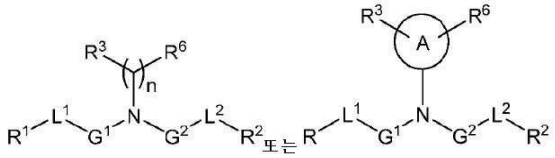
[0555] R^1 및 R^2는 각각 독립적으로 C_6-C_{24} 알킬 또는 C_6-C_{24} 알켄일이고;

[0556] R^3은 H, OR^5, CN, $-C(=O)OR^4$, $-OC(=O)R^4$ 또는 $-NR^5C(=O)R^4$이고;

[0557] R^4는 C_1-C_{12} 알킬이고;

[0558] R^5는 H 또는 C_1-C_6 알킬이고;

[0559] x는 0, 1 또는 2이다.화학식 (III)의 상기 실시형태 중 일부에서, 지질은 하기 화학식 (IIIA) 또는 (IIIB) 중 하나를 갖는다:



(IIIA) (IIIB)

[0560]

[0561] 식 중,

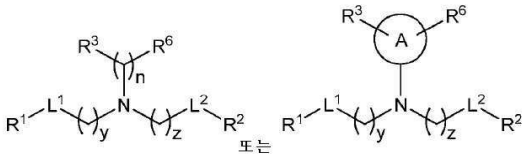
[0562] A는 3 내지 8-원 사이클로알킬 또는 사이클로알킬렌 고리이고;

[0563] R^6은, 각각의 경우에, 독립적으로 H, OH 또는 C_1-C_{24} 알킬이고;

[0564] n은 1 내지 15 범위의 정수이다.

[0565] 화학식 (III)의 상기 실시형태 중 일부에서, 지질은 화학식 (IIIA)를 갖고, 다른 실시형태에서, 지질은 화학식 (IIIB)를 갖는다.

[0566] 화학식 (III)의 다른 실시형태에서, 지질은 하기 화학식 (IIIC) 또는 (IIID) 중 하나를 갖는다:



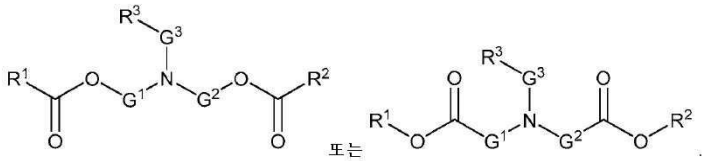
(IIIC) (IIID)

[0567]

[0568] 식 중, y 및 z는 각각 독립적으로 1 내지 12 범위의 정수이다.

[0569] 화학식 (III)의 상기 실시형태 중 임의의 것에서, L^1 또는 L^2 중 하나는 -O(C=O)-이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 -O(C=O)-이다. 상기 중 임의의 것의 일부 상이한 실시형태에서, L^1 및 L^2는 각각 독립적으로 -(C=O)O- 또는 -O(C=O)-이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 -(C=O)O-이다.

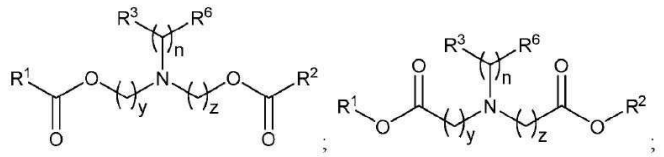
[0570] 화학 (III)의 일부 다른 실시형태에서, 지질은 하기 화학식 (IIIE) 또는 (IIIF) 중 하나를 갖는다:



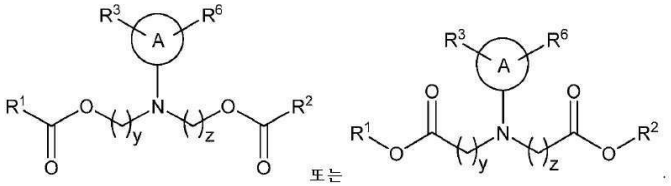
(IIIE) (IIIF)

[0571]

[0572] 화학식 (III)의 상기 실시형태 중 일부에서, 지질은 하기 화학식 (IIIG), (IIIH), (IIII) 또는 (IIIJ) 중 하나를 갖는다:



(IIIG) (IIIH)



(III) (IIIJ)

[0573]

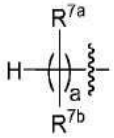
[0574] 화학식 (III)의 상기 실시형태 중 일부에서, n은 2 내지 12, 예를 들어, 2 내지 8 또는 2 내지 4 범위의 정수이다. 예를 들어, 일부 실시형태에서, n은 3, 4, 5 또는 6이다. 일부 실시형태에서, n은 3이다. 일부 실시형태에서, n은 4이다. 일부 실시형태에서, n은 5이다. 일부 실시형태에서, n은 6이다.

[0575] 화학식 (III)의 상기 실시형태 중 일부 다른 것에서, y 및 z는 각각 독립적으로 2 내지 10 범위의 정수이다. 예를 들어, 일부 실시형태에서, y 및 z는 각각 독립적으로 4 내지 9 또는 4 내지 6 범위의 정수이다.

[0576] 화학식 (III)의 상기 실시형태 중 일부에서, R^6은 H이다. 상기 실시형태 중 다른 것에서, R^6은 C_1-C_{24} 알킬이다. 다른 실시형태에서, R^6은 OH이다.

[0577] 화학식 (III)의 일부 실시형태에서, G^3은 비치환된다. 다른 실시형태에서, G3은 치환된다. 다양한 상이한 실시형태에서, G^3은 선형 C_1-C_{24} 알킬렌 또는 선형 C_1-C_{24} 알켄일렌이다.

[0578] 화학식 (III)의 일부 다른 상기 실시형태에서, R^1 또는 R^2, 또는 둘 다는 C_6-C_{24} 알켄일이다. 예를 들어, 일부 실시형태에서, R^1 및 R^2는 각각 독립적으로 하기 구조를 갖는다:



[0579]

[0580] 식 중,

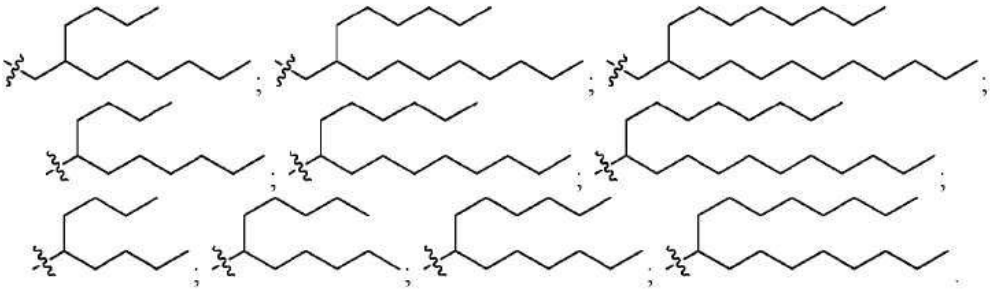
[0581] R^{7a} 및 R^{7b}는 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0582] a는 2 내지 12이고,

[0583] R^{7a}, $R7^{b}$ 및 a는 R^{1} 및 R^{2}가 각각 독립적으로 6 내지 20개의 탄소 원자를 포함하도록 각각 선택된다. 예를 들어, 일부 실시형태에서 a는 5 내지 9 또는 8 내지 12 범위의 정수이다.

[0584] 화학식 (III)의 상기 실시형태 중 일부에서, R^{7a}의 적어도 하나의 경우는 H이다. 예를 들어, 일부 실시형태에서, R^{7a}는 각각의 경우에 H이다. 상기 중 다른 상이한 실시형태에서, R^{7b} 중 적어도 하나의 경우는 C_1-C_8 알킬이다. 예를 들어, 일부 실시형태에서, C_1-C_8 알킬은 메틸, 에틸, n-프로필, 아이소프로필, n-부틸, 아이소-부틸, tert-부틸, n-헥실 또는 n-옥틸이다.

[0585] 화학식 (III)의 상이한 실시형태에서, R^{1} 또는 R^{2}, 또는 이들 둘 다는 하기 구조 중 하나를 갖는다:

[0586] 

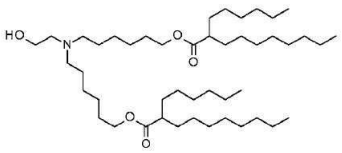
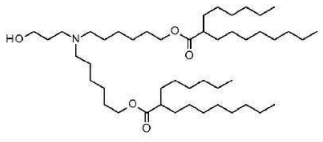
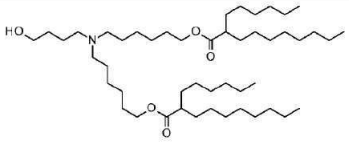
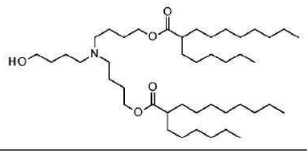
[0587] 화학식 (III)의 상기 실시형태 중 일부에서, R^{3}은 OH, CN, -C(=O)OR^{4}, -OC(=O)R^{4} 또는 -NHC(=O)R^{4}이다. 일부 실시형태에서, R^{4}는 메틸 또는 에틸이다.

[0588] 실시형태 3의 일부 구체적인 실시형태에서, 제1 스페이서 및 제2 스페이서양이온성 지질은 각각 독립적으로 화학식 III의 지질로부터 선택된다.

[0589] 다양한 상이한 실시형태에서, 화학식 (III)의 개시된 실시형태 중 임의의 하나의 양이온성 지질(예를 들어, 양이온성 지질, 제1 양이온성 지질, 제2 양이온성 지질)은 하기 표 3에 제시된 구조 중 하나를 갖는다.

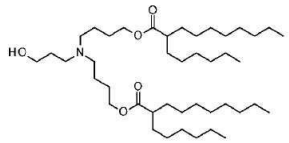
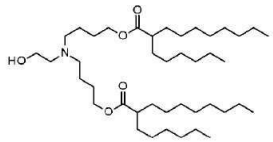
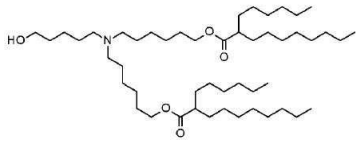
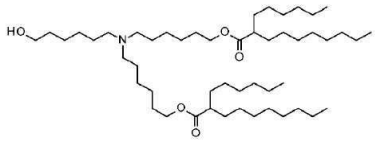
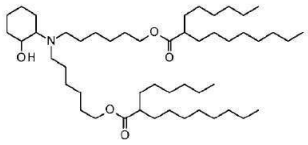
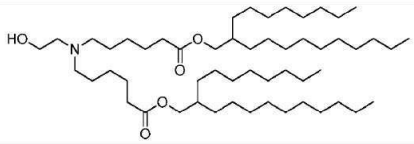
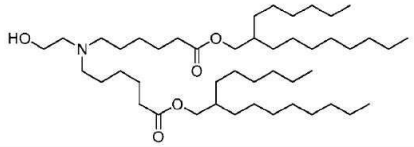
표 3

화학식 (III)의 대표적인 화합물

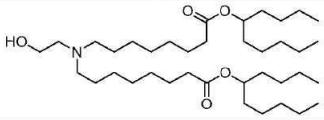
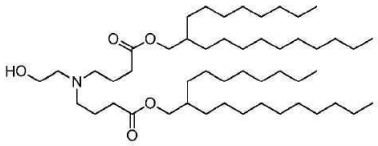
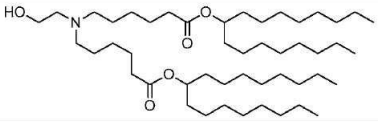
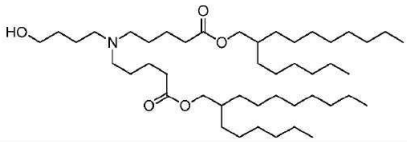
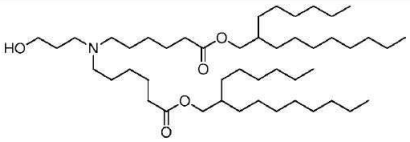
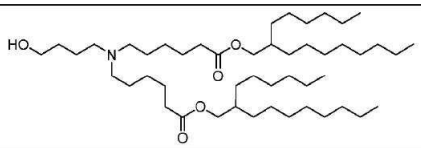
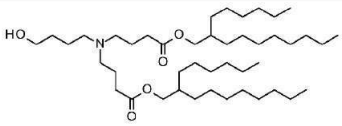
번호	구조	pKa
III-1		5.89
III-2		6.05
III-3		6.09
III-4		5.60

[0590]

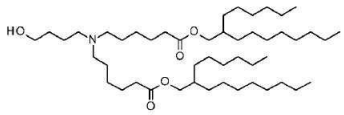
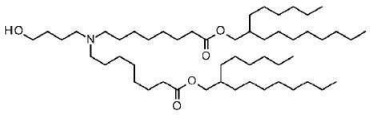
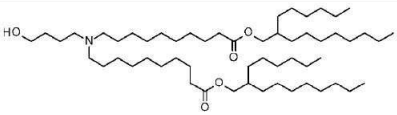
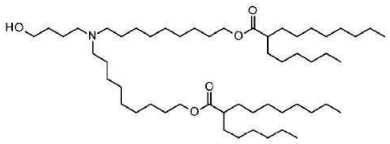
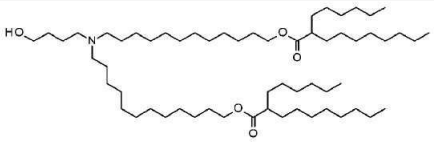
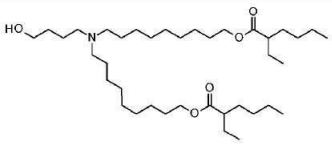
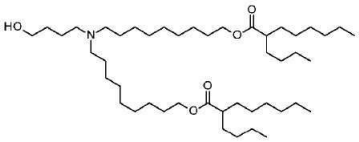
[0591]

번호	구조	pKa
III-5		5.59
III-6		5.42
III-7		6.11
III-8		5.84
III-9		-
III-10		-
III-11		-

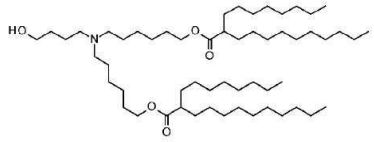
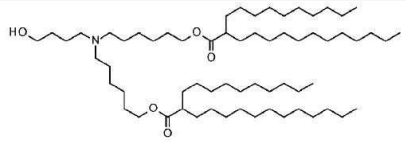
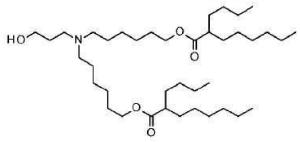
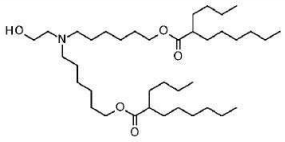
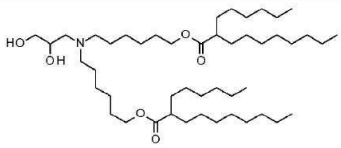
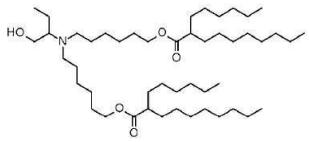
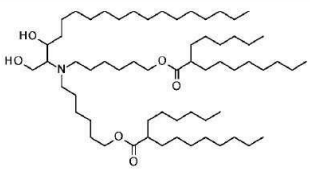
[0592]

번호	구조	pKa
III-12		-
III-13		-
III-14		-
III-15		6.14
III-16		6.31
III-17		6.28
III-18		-

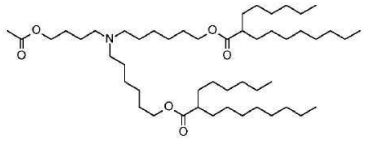
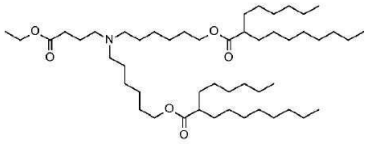
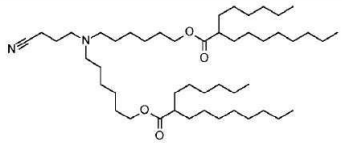
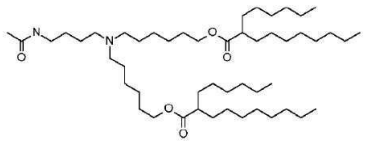
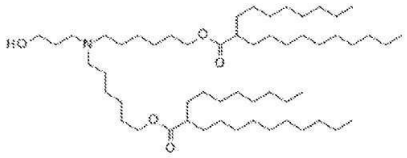
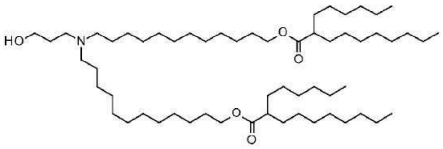
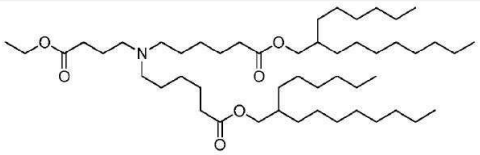
[0593]

번호	구조	pKa
III-19		-
III-20		6.36
III-21		-
III-22		6.10
III-23		5.98
III-24		-
III-25		6.22

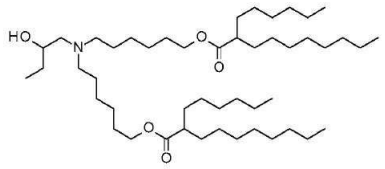
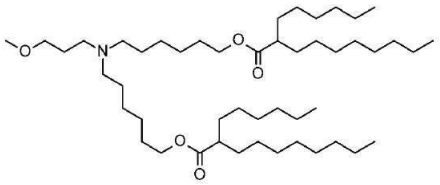
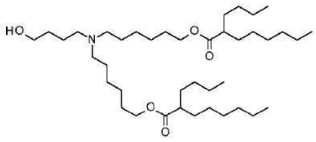
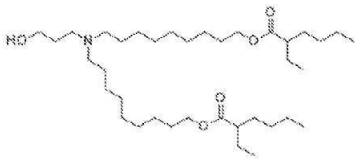
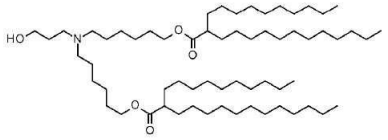
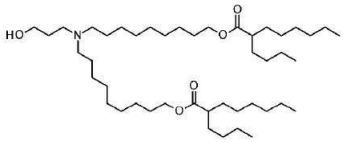
[0594]

번호	구조	pKa
III-26		5.84
III-27		5.77
III-28		-
III-29		-
III-30		6.09
III-31		-
III-32		-

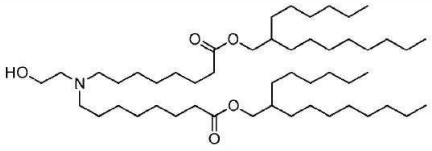
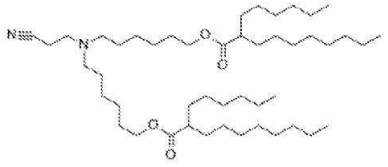
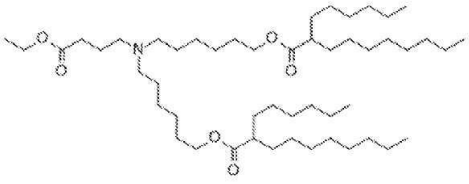
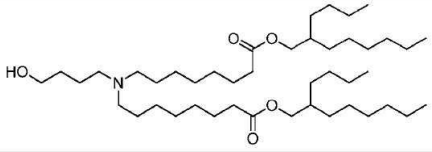
[0595]

번호	구조	pKa
III-33		-
III-34		-
III-35		-
III-36		-
III-37		-
III-38		-
III-39		-

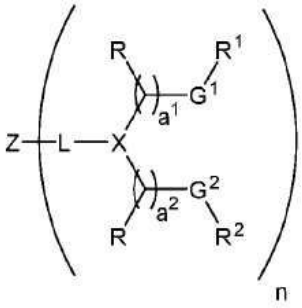
[0596]

번호	구조	pKa
III-40		-
III-41		-
III-42		-
III-43		-
III-44		-
III-45		-

[0597]

번호	구조	pKa
III-46		-
III-47		-
III-48		-
III-49		-

[0598] 일 실시형태에서, 실시형태 1, 2, 3, 4 또는 5 중 어느 하나의 양이온성 지질은 하기 화학식 (IV)의 구조, 또는 이의 약제학적으로 허용 가능한 염, 전구약물 또는 입체이성질체를 갖는다:



(IV)

[0599]

[0600] 식 중, G^1 또는 G^2 중 하나는 각각의 경우에, -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_y-$, -S-S-, -C(=O)S-, SC(=O)-, $-N(R^a)C(=O)-$, $-C(=O)N(R^a)-$, $-N(R^a)C(=O)N(R^a)-$, $-OC(=O)N(R^a)-$ 또는 $-N(R^a)C(=O)O-$이고, G^1 또는 G^2 중 나머지는, 각각의 경우에, -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_y-$, -S-S-, -C(=O)S-, -SC(=O)-, $-N(R^a)C(=O)-$, $-C(=O)N(R^a)-$, $-N(R^a)C(=O)N(R^a)-$, $-OC(=O)N(R^a)-$ 또는 $-N(R^a)C(=O)O-$ 또는 직접 결합이고;

[0601] L은, 각각의 경우에, ~O(C=O)-이되, ~는 X에 대한 공유 결합을 나타내고;

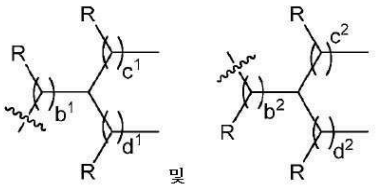
[0602] X는 CR^a이고;

[0603] Z는 n이 1인 경우에는 알킬, 사이클로알킬 또는 적어도 하나의 극성 작용기를 포함하는 1가 모이어티이거나; Z는 n이 1보다 큰 경우에는 알킬렌, 사이클로알킬렌 또는 적어도 하나의 극성 작용기를 포함하는 다가 모이어티이고;

[0604] R^a는, 각각의 경우에, 독립적으로 H, C_1-C_{12} 알킬, C_1-C_{12} 하이드록실알킬, C_1-C_{12} 아미노알킬, C_1-C_{12} 알킬아미닐알킬, C_1-C_{12} 알콕시알킬, C_1-C_{12} 알콕시카보닐, C_1-C_{12} 알킬카보닐옥시, C_1-C_{12} 알킬카보닐옥시알킬 또는 C_1-C_{12} 알킬카보닐이고;

[0605] R은 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R은 이것이 결합되는 탄소 원자와 함께 인접한 R 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0606] R^1 및 R^2는, 각각의 경우에, 각각 하기 구조를 갖는다:



R^1 및 R^2

[0607]

[0608] a^1 및 a^2는, 각각의 경우에, 독립적으로 3 내지 12의 정수이고; b^1 및 b^2는, 각각의 경우에, 독립적으로 0 또는 1이고;

[0609] c^1 및 c^2는, 각각의 경우에, 독립적으로 5 내지 10이고; d^1 및 d^2는, 각각의 경우에, 독립적으로 5 내지 10의 정수이고;

[0610] y는, 각각의 경우에, 독립적으로 0 내지 2의 정수이고;

[0611] n은 1 내지 6의 정수이되, 각각의 알킬, 알킬렌, 하이드록실알킬, 아미노알킬, 알킬아미닐알킬, 알콕시알킬, 알

콕시카보닐, 알킬카보닐옥시, 알킬카보닐옥시알킬 및 알킬카보닐은 하나 이상의 치환체로 선택적으로 치환된다.

[0612] 화학식 (IV)의 일부 실시형태에서, G^1 및 G^2는 각각 독립적으로 -O(C=O)- 또는 -(C=O)O-이다.

[0613] 화학식 (IV)의 다른 실시형태에서, X는 CH이다.

[0614] 화학식 (IV)의 상이한 실시형태에서, $a^1 + b^1 + c^1$의 합 또는 $a^2 + b^2 + c^2$의 합은 12 내지 26의 정수이다.

[0615] 화학식 (IV)의 추가의 다른 실시형태, a^1 및 a^2는 독립적으로 3 내지 10의 정수이다. 예를 들어, 일부 실시형태에서 a^1 및 a^2는 독립적으로 4 내지 9의 정수이다.

[0616] 화학식 (IV)의 다양한 실시형태에서, b^1 및 b^2는 0이다. 상이한 실시형태에서, b^1 및 b^2는 1이다.

[0617] 화학식 (IV)의 추가의 실시형태에서, c^1, c^2, d^1 및 d^2는 독립적으로 6 내지 8의 정수이다.

[0618] 화학식 (IV)의 다른 실시형태에서, c^1 및 c^2는, 각각의 경우에, 독립적으로 6 내지 10이고; d^1 및 d^2는, 각각의 경우에, 독립적으로 6 내지 10의 정수이다.

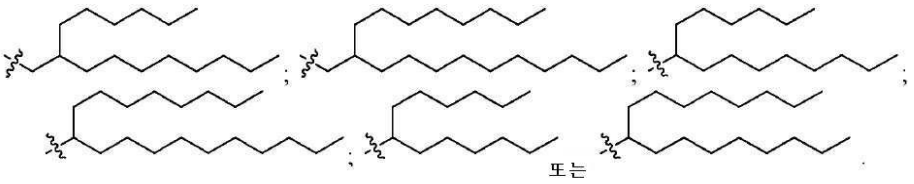
[0619] 화학식 (IV)의 다른 실시형태에서, c^1 및 c^2는, 각각의 경우에, 독립적으로 5 내지 9이고; d^1 및 d^2는, 각각의 경우에, 독립적으로 5 내지 9의 정수이다.

[0620] 화학식 (IV)의 추가의 실시형태에서, n이 1인 경우 Z는 알킬, 사이클로알킬 또는 적어도 하나의 극성 작용기를 포함하는 1가 모이어티이다. 다른 실시형태에서, Z는 알킬이다.

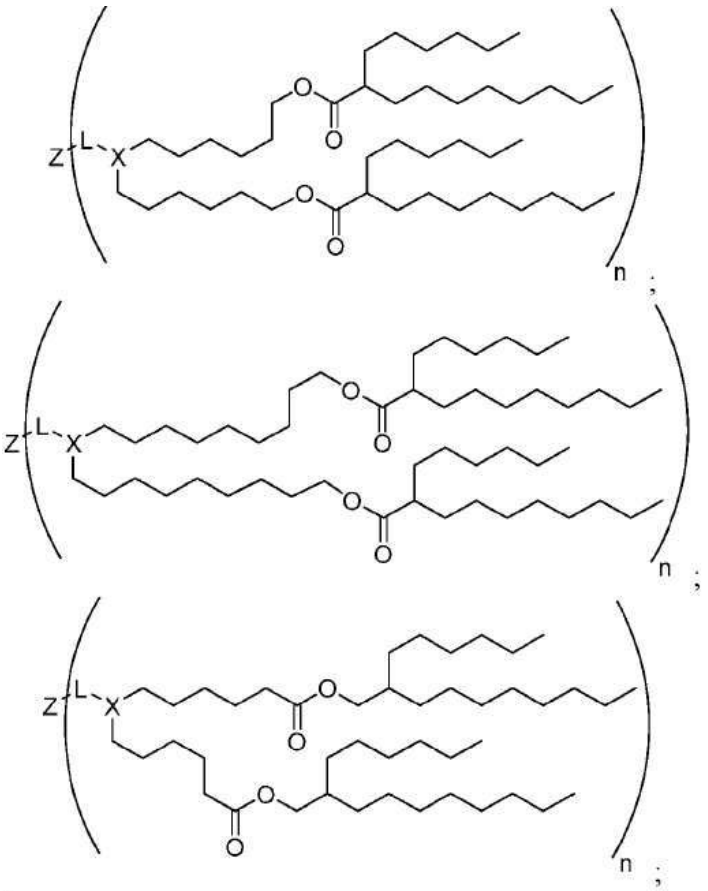
[0621] 상기 화학식 (IV)의 다양한 실시형태에서, R은 각각의 경우에, 독립적으로 (a) H 또는 메틸이거나, (b) R은 이것이 결합되는 탄소 원자와 함께 인접한 R 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0622] 특정 실시형태에서, 각각의 R은 H이다. 다른 실시형태에서, 적어도 하나의 R은 이것이 결합되는 탄소 원자와 함께 인접한 R 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

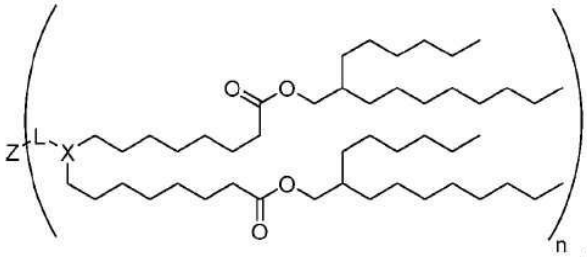
[0623] 화학식 (IV)의 화합물의 다른 실시형태에서, R^1 및 R^2는 독립적으로 하기 구조 중 하나를 갖는다:

[0624] 

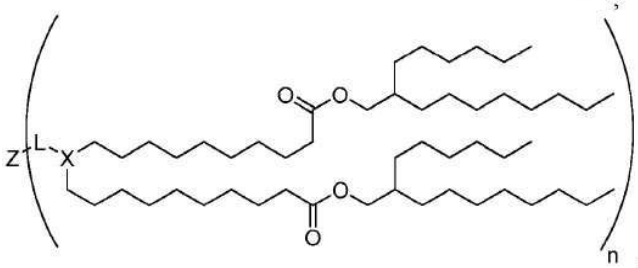
[0625] 화학식 (IV)의 특정 실시형태에서, 화합물은 하기 구조 중 하나를 갖는다:



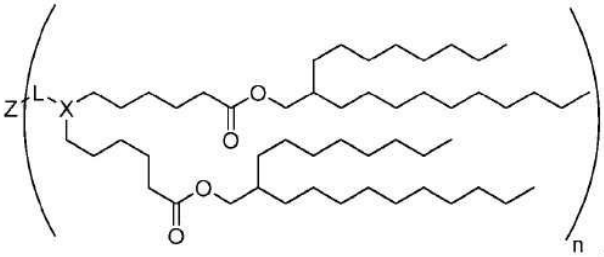
[0626]



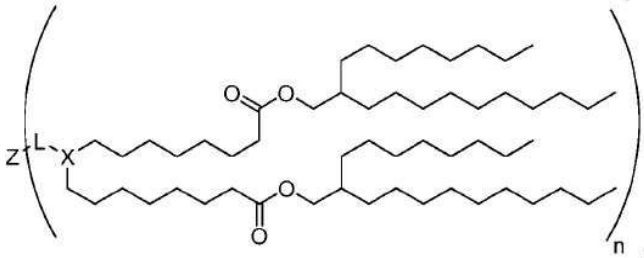
;



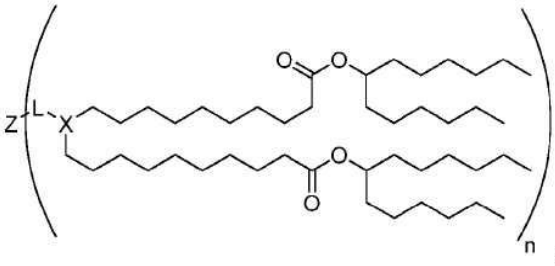
;



;

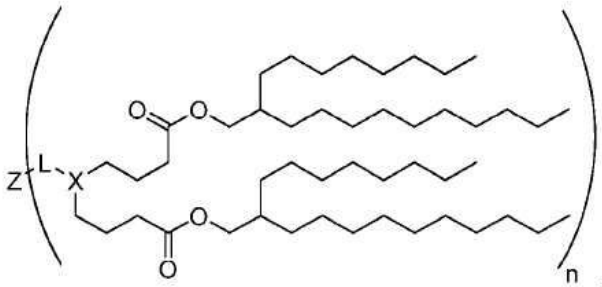


;

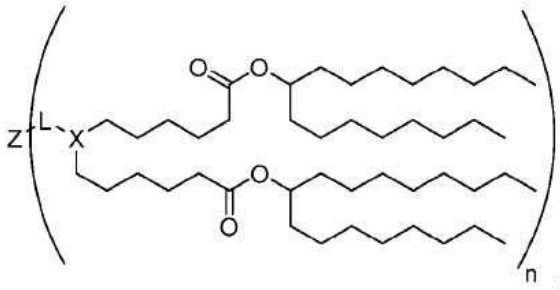


;

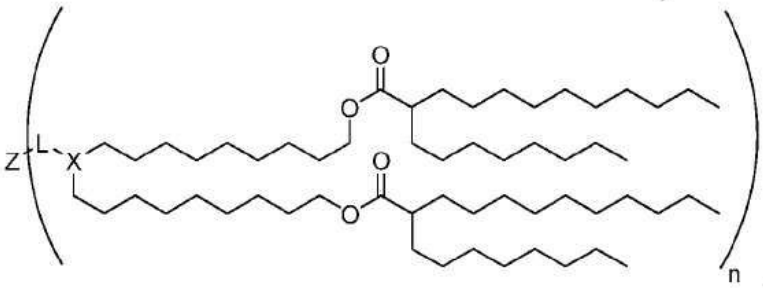
[0627]



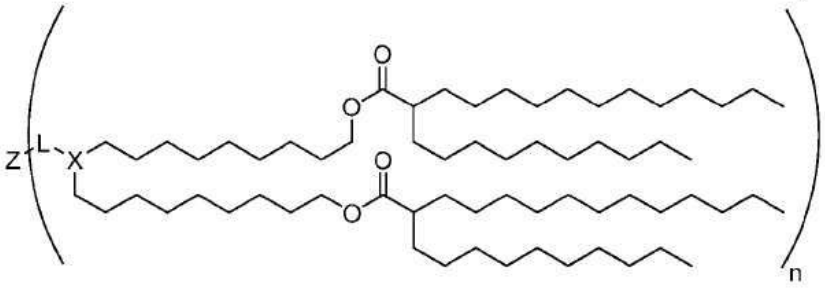
;



;

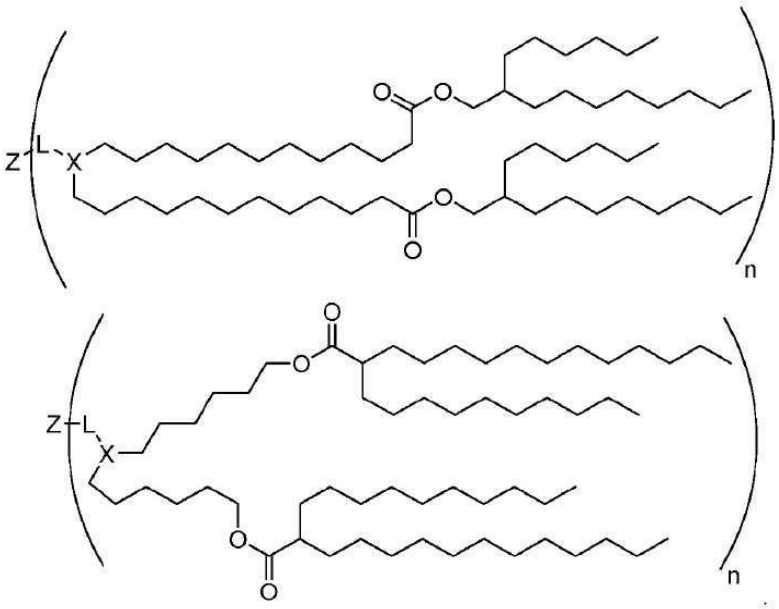


;



;

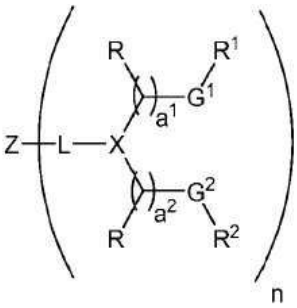
[0628]



또는

[0629]

[0630] 추가의 상이한 실시형태에서 실시형태 1, 2, 3, 4 또는 5의 양이온성 지질은 하기 화학식 (V)의 구조, 또는 이의 약제학적으로 허용 가능한 염, 전구약물 또는 입체이성질체를 갖는다:



(V)

[0631]

[0632] 식 중,

[0633] G^1 또는 G^2 중 하나는, 각각의 경우에, -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_y-$, -S-S-, -C(=O)S-, SC(=O)-, $-N(R^a)C(=O)-$, $-C(=O)N(R^a)-$, $-N(R^a)C(=O)N(R^a)-$, $-OC(=O)N(R^a)-$ 또는 $-N(R^a)C(=O)O-$이고, G^1 또는 G^2 중 나머지는, 각각의 경우에, -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, -S(O)y-, -S-S-, -C(=O)S-, -SC(=O)-, $-N(R^a)C(=O)-$, $-C(=O)N(R^a)-$, $-N(R^a)C(=O)N(R^a)-$, $-OC(=O)N(R^a)-$ 또는 $-N(R^a)C(=O)O-$ 또는 직접 결합이고;

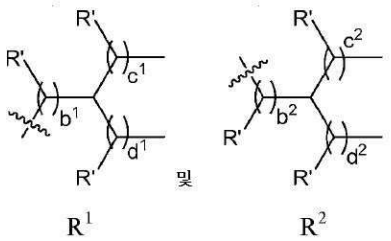
[0634] L은, 각각의 경우에, ~O(C=O)-이되, ~는 X에 대한 공유 결합을 나타내고;

[0635] X는 CR^a이고; Z는 n이 1인 경우에는 알킬, 사이클로알킬 또는 적어도 하나의 극성 작용기를 포함하는 1가 모이어티이거나; Z는 n이 1보다 큰 경우에는 알킬렌, 사이클로알킬렌 또는 적어도 하나의 극성 작용기를 포함하는 다가 모이어티이고;

[0636] R^a는, 각각의 경우에, 독립적으로 H, C_1-C_{12} 알킬, C_1-C_{12} 하이드록실알킬, C_1-C_{12} 아미노알킬, C_1-C_{12} 알킬아미닐알킬, C_1-C_{12} 알콕시알킬, C_1-C_{12} 알콕시카보닐, C_1-C_{12} 알킬카보닐옥시, C_1-C_{12} 알킬카보닐옥시알킬 또는 C_1-C_{12} 알킬카보닐이고;

[0637] R은 각각의 경우에, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R은 이것이 결합되는 탄소 원자와 함께 인접한 R 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0638] R^1 및 R^2는, 각각의 경우에, 각각 하기 구조를 갖는다:



R^1 R^2

[0639]

[0640] R'는, 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0641] a^1 및 a^2는, 각각의 경우에, 독립적으로 3 내지 12의 정수이고; b^1 및 b^2는, 각각의 경우에, 독립적으로 0 또는 1이고;

[0642] c^1 및 c^2는, 각각의 경우에, 독립적으로 2 내지 12이고; d^1 및 d^2는, 각각의 경우에, 독립적으로 2 내지 12의 정수이고;

[0643] y는, 각각의 경우에, 독립적으로 0 내지 2의 정수이고;

[0644] n은 1 내지 6의 정수이고,

[0645] a^1, a^2, c^1, c^2, d^1 및 d^2는 $a^1+c^1+d^1$의 합이 18 내지 30의 정수이고, $a^2+c^2+d^2$의 합이 18 내지 30의 정수이도록 선택되고, 각각의 알킬, 알킬렌, 하이드록실알킬, 아미노알킬, 알킬아미닐알킬, 알콕시알킬, 알콕시카보닐, 알킬카보닐옥시, 알킬카보닐옥시알킬 및 알킬카보닐은 하나 이상의 치환체로 선택적으로 치환된다.

[0646] 화학식 (V)의 특정 실시형태에서, G^1 및 G^2는 각각 독립적으로 -O(C=O)- 또는 -(C=O)O-이다.

[0647] 화학식 (V)의 다른 실시형태에서, X는 CH이다.

[0648] 화학식 (V)의 일부 실시형태에서, $a^1+c^1+d^1$의 합은 20 내지 30이고, $a^2+c^2+d^2$의 합은 18 내지 30이다. 다른 실시형태에서, $a^1+c^1+d^1$의 합은 20 내지 30의 정수이고, $a^2+c^2+d^2$의 합은 20 내지 30의 정수이다. 화학식 (V)의 추가의 실시형태에서, $a^1+b^1+c^1$의 합 또는 $a^2+b^2+c^2$의 합은 12 내지 26의 정수이다. 다른 실시형태에서, a^1, a^2, c^1, c^2, d^1 및 d^2는 $a^1+c^1+d^1$의 합이 18 내지 28의 정수이고, $a^2+c^2+d^2$의 합이 18 내지 28의 정수이도록 선택된다.

[0649] 화학식 (V)의 추가의 다른 실시형태에서, a^1 및 a^2는 독립적으로 3 내지 10의 정수, 예를 들어, 4 내지 9의 정수이다.

[0650] 화학식 (V)의 추가의 다른 실시형태에서, b^1 및 b^2는 0이다. 상이한 실시형태에서 b^1 및 b^2는 1이다.

[0651] 화학식 (V)의 특정 다른 실시형태에서, c^1, c^2, d^1 및 d^2는 독립적으로 6 내지 8의 정수이다.

[0652] 화학식 (V)의 상이한 다른 실시형태, Z는 n이 1인 경우에는 알킬 또는 적어도 하나의 극성 작용기를 포함하는 1가 모이어티이거나; Z는 n이 1보다 큰 경우에는 알킬렌 또는 적어도 하나의 극성 작용기를 포함하는 다가 모이어티이다.

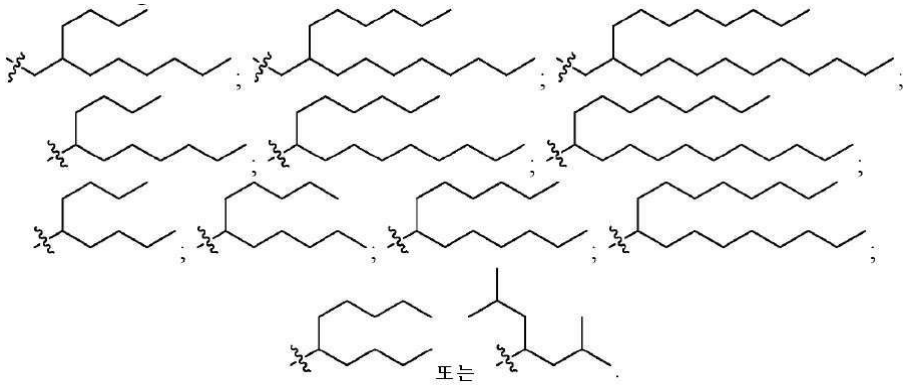
[0653] 화학식 (V)의 추가의 실시형태에서, n이 1인 경우 Z는 알킬, 사이클로알킬 또는 적어도 하나의 극성 작용기를 포함하는 1가 모이어티이다. 다른 실시형태에서, Z는 알킬이다.

[0654] 화학식 (V)의 다른 상이한 실시형태에서, R은 각각의 경우에, 독립적으로 (a) H 또는 메틸이거나, (b) R은 이것이 결합되는 탄소 원자와 함께 인접한 R 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다. 예를 들어, 일부 실시형태에서, 각각의 R은 H이다. 다른 실시형태에서, 적어도 하나의 R은 이것이 결합되는 탄소 원자와 함께 인접한 R 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0655] 추가의 실시형태에서, 각각의 R'는 H이다.

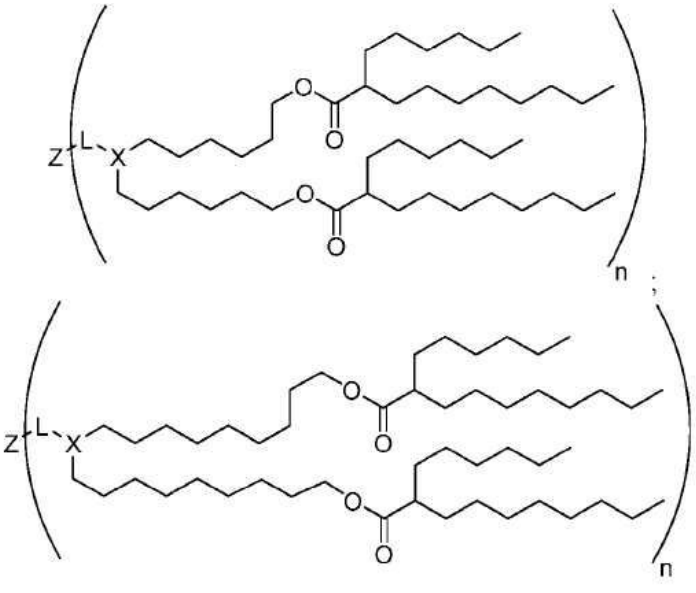
[0656] 화학식 (V)의 일부 실시형태에서, $a^1+c^1+d^1$의 합은 20 내지 25의 정수이고, $a^2+c^2+d^2$의 합은 20 내지 25의 정수이다.

[0657] 화학식 (V)의 다른 실시형태에서, R^1 및 R^2는 독립적으로 하기 구조 중 하나를 갖는다:

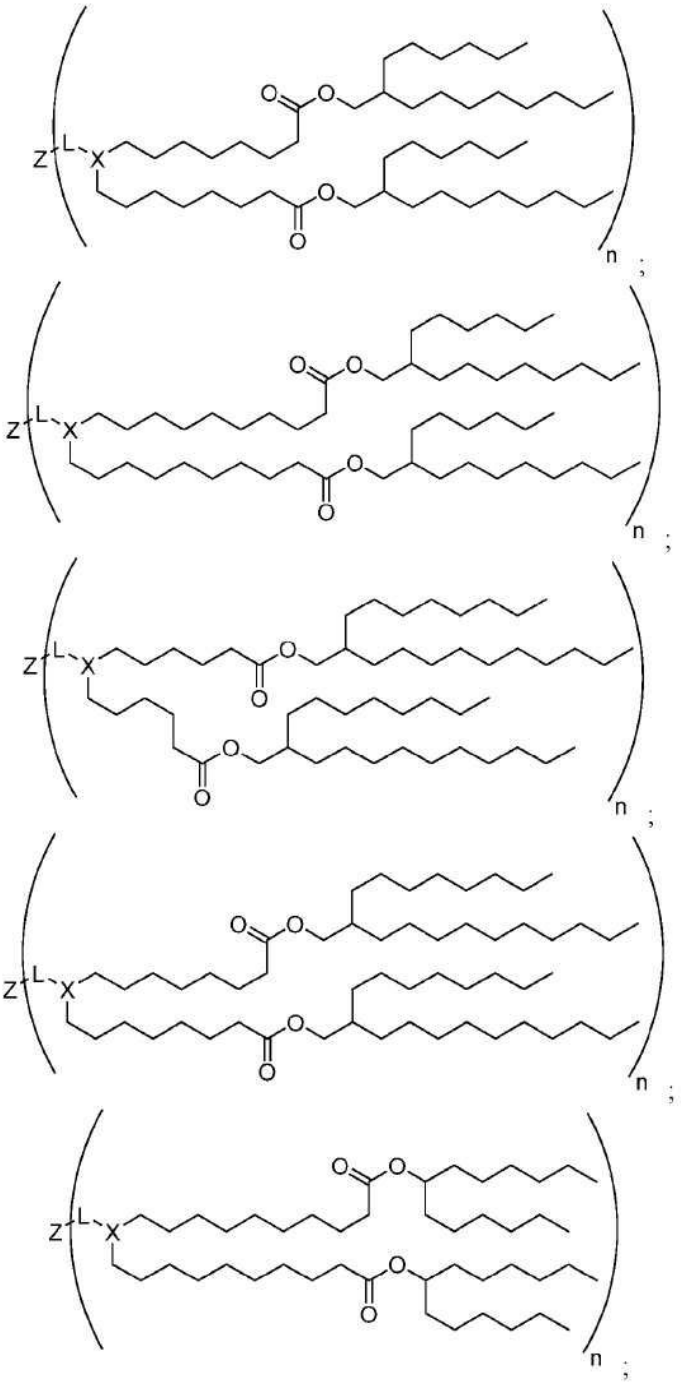


[0658] 또는

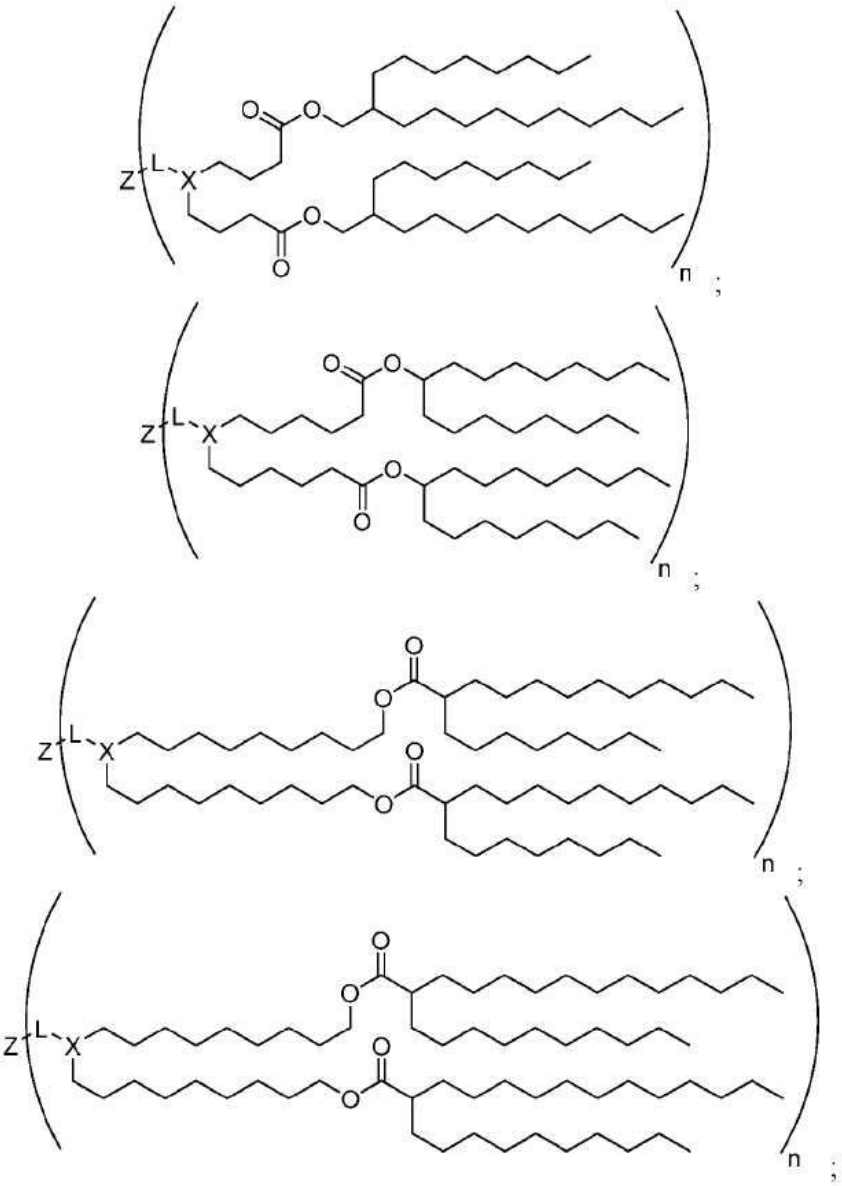
[0659] 화학식 (V)의 추가의 실시형태에서, 화합물은 하기 구조 중 하나를 갖는다:



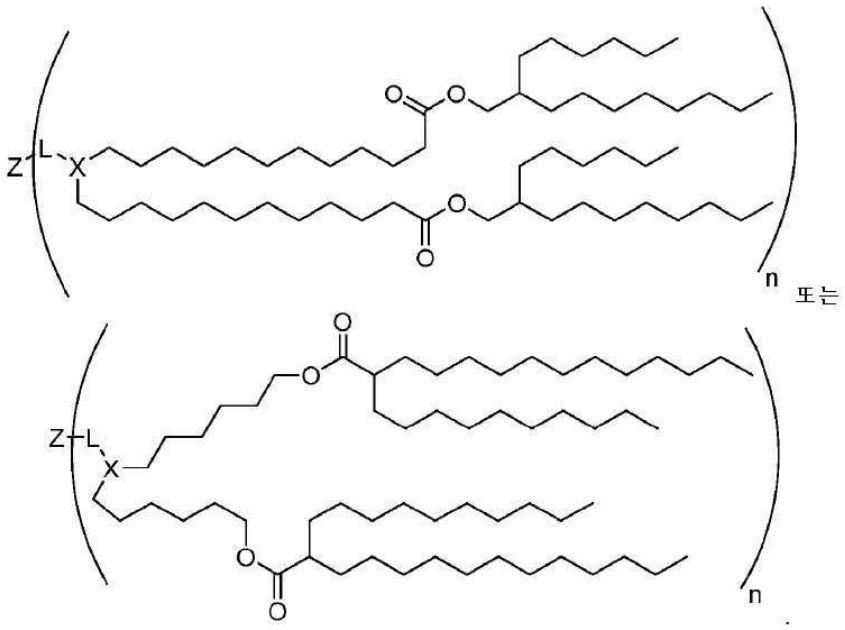
[0660]



[0661]



[0662]



[0663]

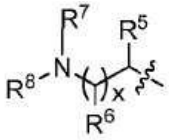
[0664] 화학식 (IV) 또는 (V)의 상기 실시형태 중 임의의 것에서, n은 1이다. 화학식 (IV) 또는 (V)의 상기 실시형태 중 나머지 것에서, n은 1 초과이다.

[0665] 화학식 (IV) 또는 (V)의 상기 실시형태 중 임의의 것에서, Z는 적어도 하나의 극성 작용기를 포함하는 1가 또는 다가 모이어티이다. 일부 실시형태에서, Z는 적어도 하나의 극성 작용기를 포함하는 1가 모이어티이다. 다른 실시형태에서, Z는 적어도 하나의 극성 작용기를 포함하는 다가 모이어티이다.

[0666] 화학식 (IV) 또는 (V)의 상기 실시형태 중 임의의 것에서, 극성 작용기는 하이드록실, 알콕시, 에스터, 사이아노, 아마이드, 아미노, 알킬아미닐, 헤테로사이클릴 또는 헤테로아릴 작용기이다.

[0667] 화학식 (IV) 또는 (V)의 상기 실시형태 중 임의의 것에서, Z는 하이드록실, 하이드록실알킬, 알콕시알킬, 아미노, 아미노알킬, 알킬아미닐, 알킬아미닐알킬, 헤테로사이클릴 또는 헤테로사이클릴알킬이다.

[0668] 화학식 (IV) 또는 (V)의 일부 다른 실시형태에서, Z는 하기 구조를 갖는다:



[0669]

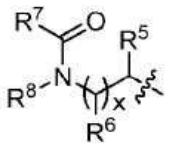
[0670] 식 중,

[0671] R^5 및 R^6은 독립적으로 H 또는 C_1-C_6 알킬이고;

[0672] R^7 및 R^8은 독립적으로 H 또는 C_1-C_6 알킬이거나 또는 R^7과 R^8은, 이들이 부착되는 질소 원자와 함께 합쳐져서 3 내지 7원 헤테로사이클릭 고리를 형성하고;

[0673] x는 0 내지 6의 정수이다.

[0674] 화학식 (IV) 또는 (V)의 추가의 상이한 실시형태에서, Z는 하기 구조를 갖는다:



[0675]

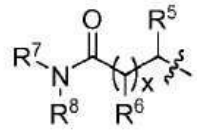
[0676] 식 중,

[0677] R^5 및 R^6은 독립적으로 H 또는 C_1-C_6 알킬이고;

[0678] R^7 및 R^8은 독립적으로 H 또는 C_1-C_6 알킬이거나 또는 R^7과 R^8은, 이들이 부착되는 질소 원자와 함께 합쳐져서 3 내지 7원 헤테로사이클릭 고리를 형성하고;

[0679] x는 0 내지 6의 정수이다.

[0680] 화학식 (IV) 또는 (V)의 추가의 상이한 실시형태에서, Z는 하기 구조를 갖는다:

[0681] 

[0682] 식 중,

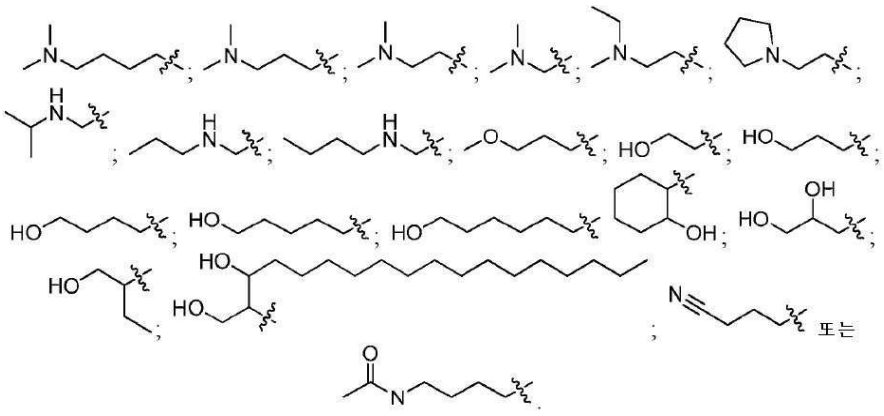
[0683] R^5 및 R^6은 독립적으로 H 또는 C_1-C_6 알킬이고;

[0684] R^7 및 R^8은 독립적으로 H 또는 C_1-C_6 알킬이거나 또는 R^7과 R^8은, 이들이 부착되는 질소 원자와 함께 합쳐져서 3 내지 7원 헤테로사이클릭 고리를 형성하고;

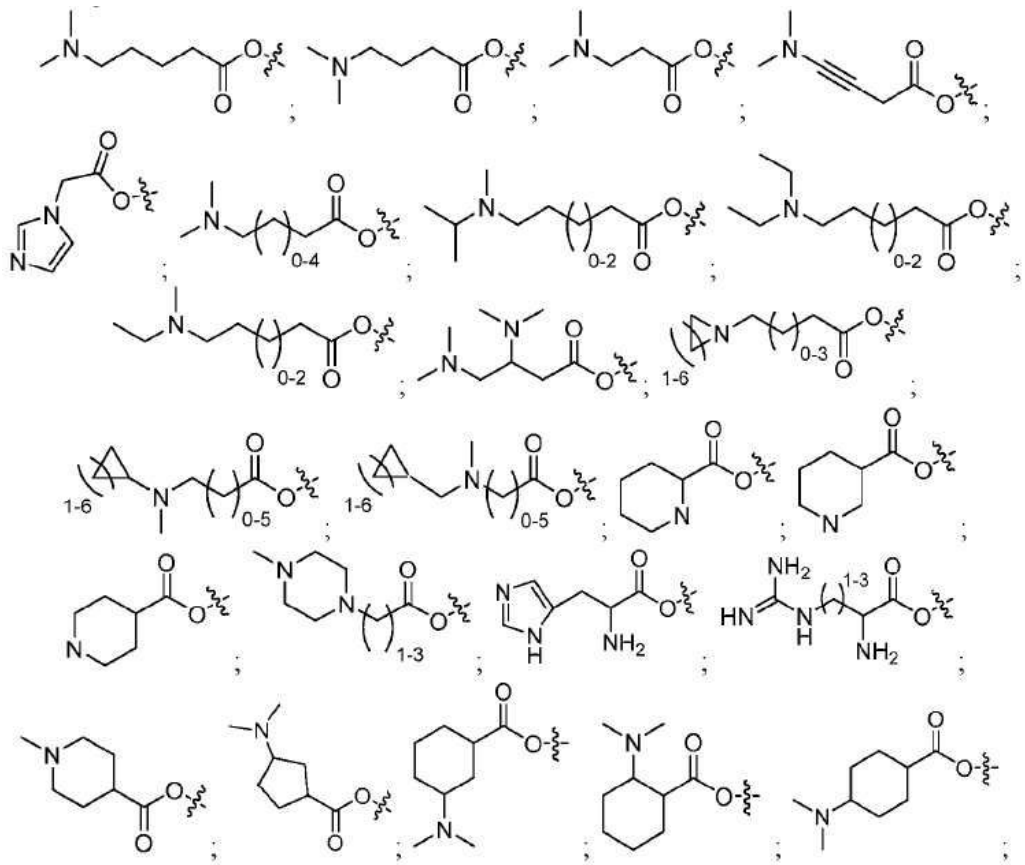
[0685] x는 0 내지 6의 정수이다.

[0686] 화학식 (IV) 또는 (V)의 일부 다른 실시형태에서, Z는 하이드록실알킬, 사이아노알킬 또는 하나 이상의 에스터 또는 아마이드기로 치환된 알킬이다.

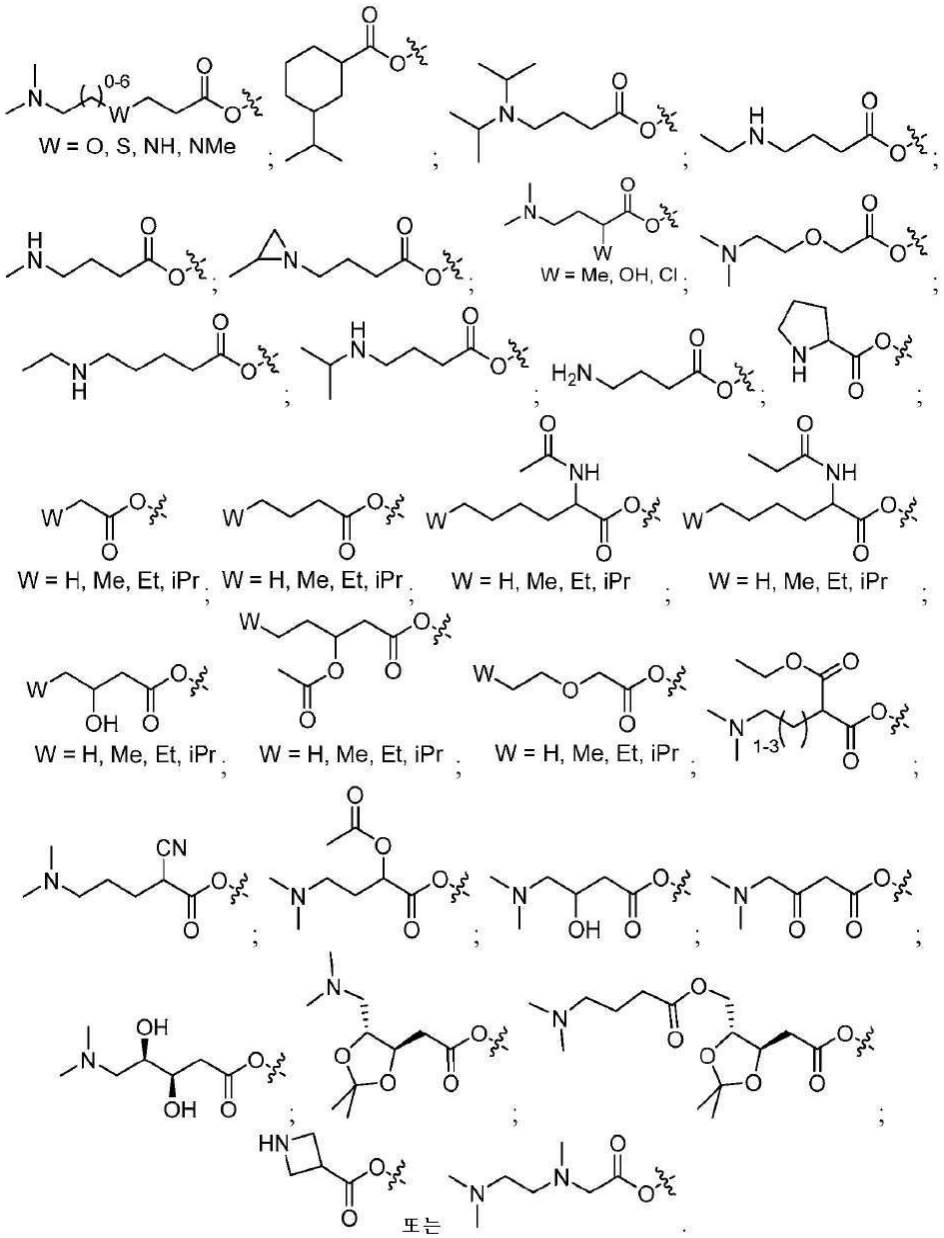
[0687] 예를 들어, 화학식 (IV) 또는 (V)의 상기 실시형태 중 임의의 것에서, Z는 하기 구조 중 하나를 갖는다:

[0688] 

[0689] 화학식 (IV) 또는 (V)의 다른 실시형태에서, Z-L은 하기 구조 중 하나를 갖는다:

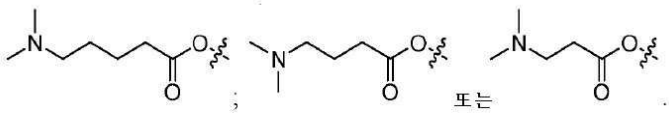


[0690]



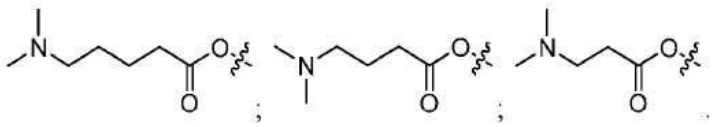
[0691]

[0692] 다른 실시형태에서, Z-L은 하기 구조 중 하나를 갖는다:



[0693]

[0694] 추가의 다른 실시형태에서, X는 CH이고, Z-L은 하기 구조 중 하나를 갖는다:

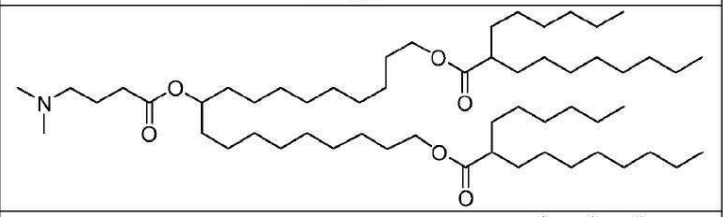
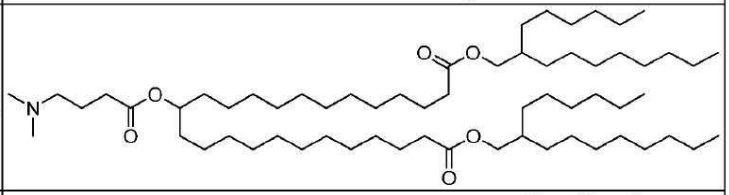
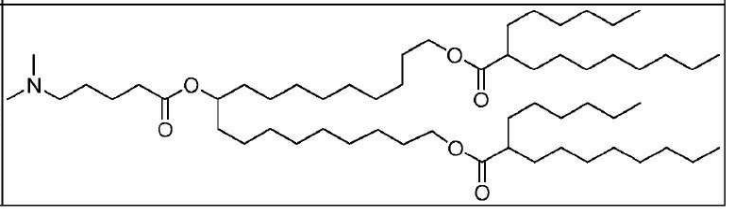


[0695]

[0696] 다양한 상이한 실시형태에서, 임의의 하나의 실시형태 1, 2, 3, 4 또는 5의 양이온성 지질은 하기 표 4에 제시된 구조 중 하나를 갖는다.

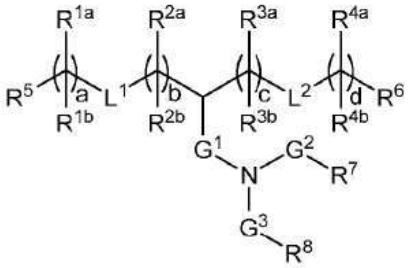
표 4

화학식 (IV) 또는 (V)의 대표적인 화합물

번호	구조
IV-1	
IV-2	
IV-3	

[0697]

[0698] 일 실시형태에서, 양이온성 지질은 하기 구조 (VI)을 갖는 화합물, 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체이다:



(VI)

[0699]

[0700] 식 중,

[0701] L^1 및 L^2는 각각 독립적으로 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_X-$, -S-S-, -C(=O)S-, -SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $-NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$, $-NR^aC(=O)O-$ 또는 직접 결합이고;

[0702] G^1은 C_1-C_2 알킬렌, -(C=O)-, -O(C=O)-, -SC(=O)-, $-NR^aC(=O)-$ 또는 직접 결합이고;

[0703] G^2는 -C(=O)-, -(C=O)O-, -C(=O)S-, $-C(=O)NR^a-$ 또는 직접 결합이고;

[0704] G^3은 C_1-C_6 알킬렌이고;

[0705] R^a는 H 또는 C_1-C_{12} 알킬이고;

[0706] R^{1a} 및 R^{1b}는 각각의 경우에서, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{1a}는 H 또는 C_1-C_{12} 알킬이고, R^{1b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{1b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결

합을 형성하고;

[0707] R^{2a} 및 R^{2b}는 각각의 경우에서, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나, (b) R^{2a}는 H 또는 C_1-C_{12} 알킬이고, R^{2b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{2b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0708] R^{3a} 및 R^{3b}는 각각의 경우에서, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나; (b) R^{3a}는 H 또는 C_1-C_{12} 알킬이고, R^{3b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{3b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0709] R^{4a} 및 R^{4b}는 각각의 경우에서, 독립적으로 (a) H 또는 C_1-C_{12} 알킬이거나; (b) R^{4a}는 H 또는 C_1-C_{12} 알킬이고, R^{4b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{4b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성하고;

[0710] R^5 및 R^6은 각각 독립적으로 H 또는 메틸이고;

[0711] R^7은 H 또는 C_1-C_{20} 알킬이고;

[0712] R^8은 OH, $-N(R^9)(C=O)R^{10}$, $-(C=O)NR^9R^{10}$, $-NR^9R^{10}$, $-(C=O)OR^{11}$ 또는 $-O(C=O)R^{11}$이되, 단 R^8이 $-NR^9R^{10}$인 경우 G^3은 C_4-C_6 알킬렌이고;

[0713] R^9 및 R^{10}은 각각 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0714] R^{11}은 아르알킬이고;

[0715] a, b, c 및 d는 각각 독립적으로 1 내지 24의 정수이고;

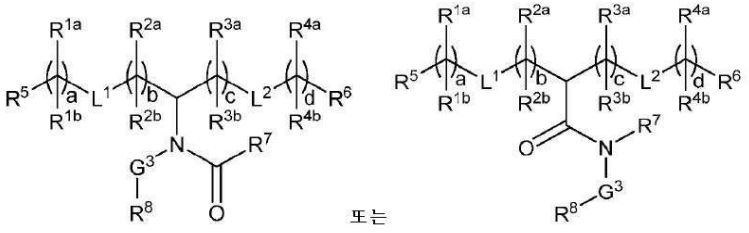
[0716] x는 0, 1 또는 2이고,

[0717] 각각의 알킬, 알킬렌 및 아르알킬은 선택적으로 치환된다.

[0718] 구조 (VI)의 일부 실시형태에서, L^1 및 L^2는 각각 독립적으로 -O(C=O)-, -(C=O)O- 또는 직접 결합이다. 다른 실시형태에서, G^1 및 G^2는 각각 독립적으로 -(C=O)- 또는 직접 결합이다. 일부 상이한 실시형태에서, L^1 및 L^2는 각각 독립적으로 -O(C=O)-, -(C=O)O- 또는 직접 결합이고; G^1 및 G^2는 각각 독립적으로 -(C=O)- 또는 직접 결합이다.

[0719] 구조식 (VI)의 일부 상이한 실시형태에서, L^1 및 L^2는 각각 독립적으로 -C(=O)-, -O-, $-S(O)_x$-, -S-S-, -C(=O)S-, -SC(=O)-, $-NR^a$-, $-NR^aC(=O)$-, $-C(=O)NR^a$-, $-NR^aC(=O)NR^a$, $-OC(=O)NR^a$-, $-NR^aC(=O)O$-, $-NR^aS(O)_xNR^a$-, $-NR^aS(O)_x$- 또는 $-S(O)_xNR^a$-이다.

[0720] 화학식 (VI)의 상기 실시형태 중 다른 것에서, 화합물은 하기 구조식 (VIA) 또는 (VIB) 중 하나를 갖는다:



(VIA) (VIB)

[0721]

[0722] 일부 실시형태에서, 화합물은 구조식 (VIA)를 갖는다. 다른 실시형태에서, 화합물은 구조식 (VIB)를 갖는다.

[0723] 구조식 (VI)의 상기 실시형태 중 임의의 것에서, L^1 또는 L^2 중 하나는 -O(C=O)-이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 -O(C=O)-이다.

[0724] 상기 중 임의의 것의 일부 상이한 실시형태에서, L^1 또는 L^2 중 하나는 -(C=O)O-이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 -(C=O)O-이다.

[0725] 구조식 (VI)의 상이한 실시형태에서, L^1 또는 L^2 중 하나는 직접 결합이다. 본 명세서에서 사용되는 바와 같이, "직접 결합"은 기(예를 들어, L^1 또는 L^2)가 존재하지 않는다는 것을 의미한다 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 직접 결합이다.

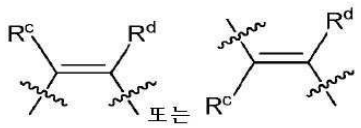
[0726] 상기의 다른 상이한 실시형태에서, R^{1a} 및 R^{1b} 중 적어도 하나의 경우에 대해서, R^{1a}는 H 또는 C_1-C_{12} 알킬이고, R^{1b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{1b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0727] 구조식 (VI)의 추가의 다른 상이한 실시형태에서, R^{4a} 및 R^{4b} 중 적어도 하나의 경우에 대해서, R^{4a}는 H 또는 C_1-C_{12} 알킬이고, R^{4b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{4b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0728] 구조식 (VI)의 추가의 실시형태에서, R^{2a} 및 R^{2b} 중 적어도 하나의 경우에 대해서, R^{2a}는 H 또는 C_1-C_{12} 알킬이고, R^{2b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{2b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

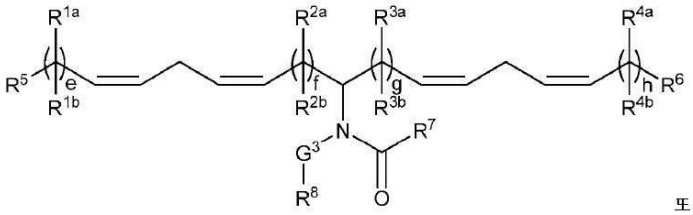
[0729] 상기 중 임의의 것의 다른 상이한 실시형태에서, R^{3a} 및 R^{3b} 중 적어도 하나의 경우에 대해서, R^{3a}는 H 또는 C_1-C_{12} 알킬이고, R^{3b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{3b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0730] "탄소-탄소" 이중 결합은 하기 구조 중 하나를 지칭한다는 것이 이해된다:

[0731] 

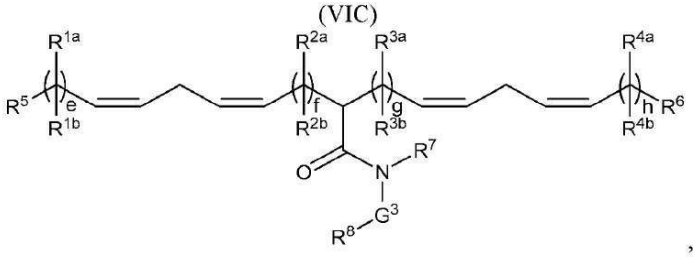
[0732] 식 중, R^c 및 R^d는 각각의 경우에, 독립적으로 H 또는 치환체이다. 예를 들어, 일부 실시형태에서 R^c 및 R^d는, 각각의 경우에, 독립적으로 H, C_1-C_{12} 알킬 또는 사이클로알킬, 예를 들어, H 또는 C_1-C_{12} 알킬이다.

[0733] 다양한 다른 실시형태에서, 화합물은 하기 구조식 (VIC) 또는 (VID) 중 하나를 갖는다:



또는

(VIC)



,

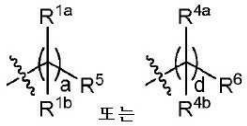
(VID)

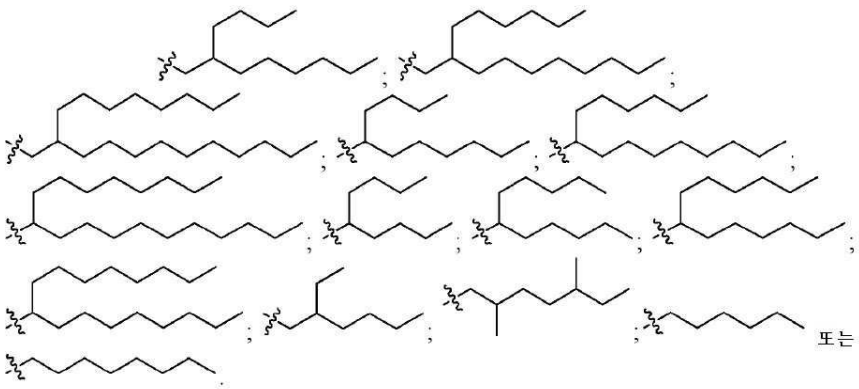
[0734]

[0735] 식 중, e, f, g 및 h는 각각 독립적으로 1 내지 12의 정수이다.

[0736] 일부 실시형태에서, 화합물은 구조식 (VIC)를 갖는다. 다른 실시형태에서, 화합물은 구조식 (VID)를 갖는다.

[0737] 구조식 (VIC) 또는 (VID)의 화합물의 다양한 실시형태에서, e, f, g 및 h는 각각 독립적으로 4 내지 10의 정수이다.

[0738] 다른 실시형태에서,  또는 이들 둘 다는 독립적으로 하기 구조 중 하나를 갖는다:



또는

[0739]

[0740] 상기 중 특정 실시형태에서, a, b, c 및 d는 각각 독립적으로 2 내지 12의 정수 또는 4 내지 12의 정수이다. 다른 실시형태에서, a, b, c 및 d는 각각 독립적으로 8 내지 12 또는 5 내지 9의 정수이다. 일부 특정 실시형태에서, a는 0이다. 일부 실시형태에서, a는 1이다. 다른 실시형태에서, a는 2이다. 추가의 실시형태에서, a는 3이다. 추가의 다른 실시형태에서, a는 4이다. 일부 실시형태에서, a는 5이다. 다른 실시형태에서, a는 6이다. 추가의 실시형태에서, a는 7이다. 추가의 다른 실시형태에서, a는 8이다. 일부 실시형태에서, a는 9이다. 다른 실시형태에서, a는 10이다. 추가의 실시형태에서, a는 11이다. 추가의 다른 실시형태에서, a는 12이다. 일부 실시형태에서, a는 13이다. 다른 실시형태에서, a는 14이다. 추가의 실시형태에서, a는 15이다. 추가의 다른 실시형태에서, a는 16이다.

[0741] 구조식 (VI)의 일부 실시형태에서, b는 1이다. 다른 실시형태에서, b는 2이다. 추가의 실시형태에서, b는 3이다. 추가의 다른 실시형태에서, b는 4이다. 일부 실시형태에서, b는 5이다. 다른 실시형태에서, b는 6이다. 추가의 실시형태에서, b는 7이다. 추가의 다른 실시형태에서, b는 8이다. 일부 실시형태에서, b는 9이다. 다른 실시형태에서, b는 10이다. 추가의 실시형태에서, b는 11이다. 추가의 다른 실시형태에서, b는 12이다. 일부 실

시형태에서, b는 13이다. 다른 실시형태에서, b는 14이다. 추가의 실시형태에서, b는 15이다. 추가의 다른 실시형태에서, b는 16이다.

[0742] 구조식 (VI)의 일부 실시형태에서, c는 1이다. 다른 실시형태에서, c는 2이다. 추가의 실시형태에서, c는 3이다. 추가의 다른 실시형태에서, c는 4이다. 일부 실시형태에서, c는 5이다. 다른 실시형태에서, c는 6이다. 추가의 실시형태에서, c는 7이다. 추가의 다른 실시형태에서, c는 8이다. 일부 실시형태에서, c는 9이다. 다른 실시형태에서, c는 10이다. 추가의 실시형태에서, c는 11이다. 추가의 다른 실시형태에서, c는 12이다. 일부 실시형태에서, c는 13이다. 다른 실시형태에서, c는 14이다. 추가의 실시형태에서, c는 15이다. 추가의 다른 실시형태에서, c는 16이다.

[0743] 구조식 (VI)의 일부 특정 실시형태에서, d는 0이다. 일부 실시형태에서, d는 1이다. 다른 실시형태에서, d는 2이다. 추가의 실시형태에서, d는 3이다. 추가의 다른 실시형태에서, d는 4이다. 일부 실시형태에서, d는 5이다. 다른 실시형태에서, d는 6이다. 추가의 실시형태에서, d는 7이다. 추가의 다른 실시형태에서, d는 8이다. 일부 실시형태에서, d는 9이다. 다른 실시형태에서, d는 10이다. 추가의 실시형태에서, d는 11이다. 추가의 다른 실시형태에서, d는 12이다.

[0744] 일부 실시형태에서, d는 13이다. 다른 실시형태에서, d는 14이다. 추가의 실시형태에서, d는 15이다. 추가의 다른 실시형태에서, d는 16이다.

[0745] 구조식 (VI)의 일부 실시형태에서, e는 1이다. 다른 실시형태에서, e는 2이다. 추가의 실시형태에서, e는 3이다. 추가의 다른 실시형태에서, e는 4이다. 일부 실시형태에서, e는 5이다. 다른 실시형태에서, e는 6이다. 추가의 실시형태에서, e는 7이다. 추가의 다른 실시형태에서, e는 8이다. 일부 실시형태에서, e는 9이다. 다른 실시형태에서, e는 10이다. 추가의 실시형태에서, e는 11이다. 추가의 다른 실시형태에서, e는 12이다.

[0746] 구조식 (VI)의 일부 실시형태에서, f는 1이다. 다른 실시형태에서, f는 2이다. 추가의 실시형태에서, f는 3이다. 추가의 다른 실시형태에서, f는 4이다. 일부 실시형태에서, f는 5이다. 다른 실시형태에서, f는 6이다. 추가의 실시형태에서, f는 7이다. 추가의 다른 실시형태에서, f는 8이다. 일부 실시형태에서, f는 9이다. 다른 실시형태에서, f는 10이다. 추가의 실시형태에서, f는 11이다. 추가의 다른 실시형태에서, f는 12이다.

[0747] 구조식 (VI)의 일부 실시형태에서, g는 1이다. 다른 실시형태에서, g는 2이다. 추가의 실시형태에서, g는 3이다. 추가의 다른 실시형태에서, g는 4이다. 일부 실시형태에서, g는 5이다. 다른 실시형태에서, g는 6이다. 추가의 실시형태에서, g는 7이다. 추가의 다른 실시형태에서, g는 8이다. 일부 실시형태에서, g는 9이다. 다른 실시형태에서, g는 10이다. 추가의 실시형태에서, g는 11이다. 추가의 다른 실시형태에서, g는 12이다.

[0748] 구조식 (VI)의 일부 실시형태에서, h는 1이다. 다른 실시형태에서, e는 2이다. 추가의 실시형태에서, h는 3이다. 추가의 다른 실시형태에서, h는 4이다. 일부 실시형태에서, e는 5이다. 다른 실시형태에서, h는 6이다. 추가의 실시형태에서, h는 7이다. 추가의 다른 실시형태에서, h는 8이다. 일부 실시형태에서, h는 9이다. 다른 실시형태에서, h는 10이다. 추가의 실시형태에서, h는 11이다. 추가의 다른 실시형태에서, h는 12이다.

[0749] 구조식 (VI)의 일부 다른 다양한 실시형태에서, a 및 d는 동일하다. 일부 다른 실시형태에서, b 및 c는 동일하다. 일두 다른 구체적인 실시형태에서, a 및 d는 동일하고, b 및 c는 동일하다.

[0750] a와 b의 합 및 c와 d의 합은 목적하는 특성을 갖는 지질을 얻기 위해서 달라질 수 있는 인자이다. 일 실시형태에서, a 및 b는 이들의 합이 14 내지 24의 범위의 정수이도록 선택된다. 다른 실시형태에서, c 및 d는 이들의 합이 14 내지 24의 범위의 정수이도록 선택된다. 추가 실시형태에서, a와 b의 합 및 c와 d의 합은 동일하다. 예를 들어, 일부 실시형태에서 a와 b의 합 및 c와 d의 합 둘 다는 14 내지 24의 범위일 수 있는 동일한 정수이다.

[0751] 더 추가의 실시형태에서, a, b, c 및 d는 a와 b의 합 및 c와 d의 합이 12 이상이도록 선택된다.

[0752] R^{1a}, R^{2a}, R^{3a} 및 R^{4a}에서 치환체는 특별히 제한되지 않는다. 일부 실시형태에서, R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 H이다. 특정 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a}는 각각의 경우에 H이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_{12} 알킬이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_8 알킬이다. 특정 다른 실시형태에서 R^{1a}, R^{2a}, R^{3a} 및 R^{4a} 중 적어도 하나는 C_1-C_6 알킬이다. 상기 실시형태 중 일부에서, C_1-C_8 알킬은 메틸, 에틸, n-프로필, 아이소-프로필, n-부틸, 아이소-부틸, tert-부틸, n-헥실 또는 n-옥틸이다.

[0753] 상기 중 특정 실시형태에서, R^{1a}, R^{1b}, R^{4a} 및 R^{4b}는 각각의 경우에 C_1-C_{12} 알킬이다.

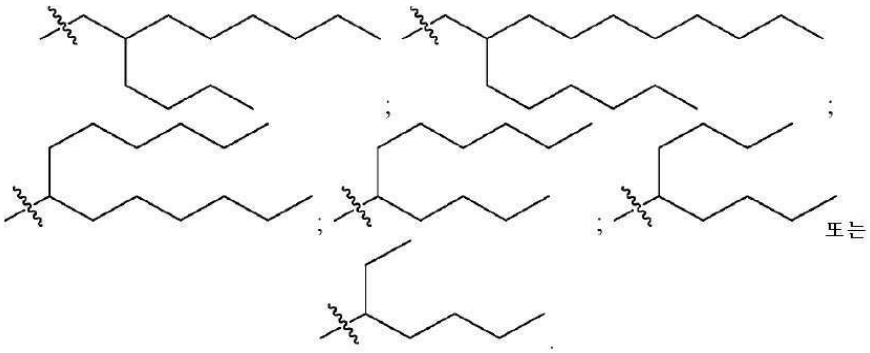
[0754] 상기 중 추가 실시형태에서, R^{1b}, R^{2b}, R^{3b} 및 R^{4b} 중 적어도 하나는 H이거나 R^{1b}, R^{2b}, R^{3b} 및 R^{4b}는 각각의 경우에 H이다.

[0755] 상기 중 특정 실시형태에서, R^{1b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{1b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다. 상기 중 다른 실시형태에서, R^{4b}는 이것이 결합되는 탄소 원자와 함께 인접한 R^{4b} 및 이것이 결합되는 탄소와 함께 취해져서 탄소-탄소 이중 결합을 형성한다.

[0756] R^5 및 R^6에서 치환체는 상기 실시형태에서 특별히 제한되지 않는다. 특정 실시형태에서 R^5 또는 R^6 중 하나는 메틸이다. 다른 실시형태에서 R^5 또는 R^6 각각은 메틸이다.

[0757] R^7에서 치환체는 상기 실시형태에서 특별히 제한되지 않는다. 특정 실시형태에서 R^7은 C_6-C_{16} 알킬이다. 일부 다른 실시형태에서, R^7은 C_6-C_9 알킬이다. 이들 실시형태 중 일부에서, R^7은 -(C=O)OR^b, -O(C=O)R^b, -C(=O)R^b, -OR^b, -S(O)$_x$$R^b$, -S-$SR^b$, -C(=O)$SR^b$, -SC(=O)$R^b$, -$NR^aR^b$, -$NR^aC$(=O)$R^b$, -C(=O)$NR^aR^b$, -$NR^aC$(=O)$NR^aR^b$, -OC(=O)$NR^aR^b$, -$NR^aC$(=O)$OR^b$, -$NR^aS(O)_x$$NR^aR^b$, -$NR^aS(O)_x$$R^b$ 또는 -S(O)$_x$$NR^aR^b$로 치환되고, R^a는 H 또는 C_1-C_{12} 알킬이고; R^b는 C_1-C_{15} 알킬이고; x는 0, 1 또는 2이다. 예를 들어, 일부 실시형태에서 R^7은 -(C=O)OR^b 또는 -O(C=O)R^b로 치환된다.

[0758] 구조식 (VI)의 상기 실시형태 중 다양한 것에서, R^b는 분지형 C_3-C_{15} 알킬이다. 예를 들어, 일부 실시형태에서 R^b는 하기 구조 중 하나를 갖는다:

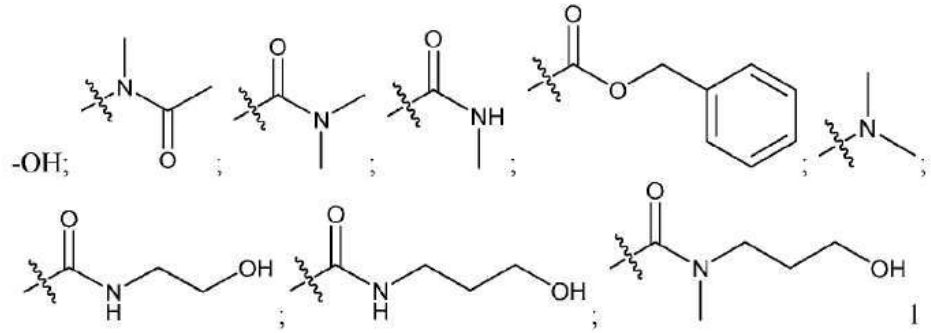
[0759] 

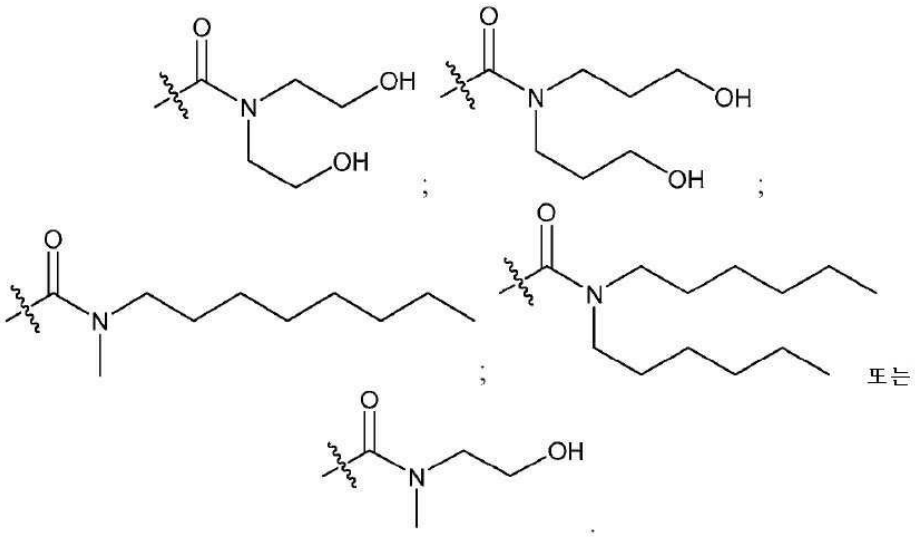
[0760] 특정 실시형태에서, R^8은 OH이다.

[0761] 구조식 (VI)의 다른 실시형태에서, R^8은 -N(R^9)(C=O)R^{10}이다. 일부 다른 실시형태에서, R^8은 -(C=O)NR^9R^{10}이다. 더 추가의 실시형태에서, R^8은 -NR^9R^{10}이다. 상기 실시형태 중 일부에서, R^9 및 R^{10}은 각각 독립적으로 H 또는 C_1-C_8 알킬, 예를 들어, H 또는 C_1-C_3 알킬이다. 이들 실시형태 중 보다 구체적인 것에서, C_1-C_8 알킬 또는 C_1-C_3 알킬은 비치환되거나 하이드록실로 치환된다. 이들 실시형태 중 다른 것에서, R^9 및 R^{10}은 각각 메틸이다.

[0762] 구조식 (VI)의 더 추가의 실시형태에서, R^8은 -(C=O)OR^{11}이다. 이들 실시형태 중 일부에서, R^{11}은 벤질이다.

[0763] 구조식 (VI)의 더 추가의 구체적인 실시형태에서, R^8은 하기 구조를 갖는다:

[0764] 

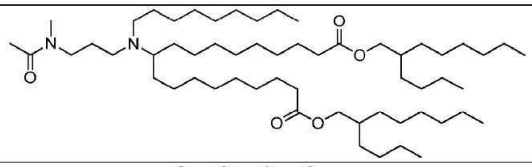
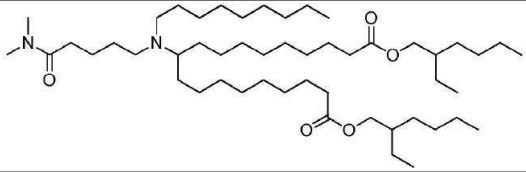
[0765] 

[0766] 상기 화합물의 추가의 다른 실시형태, G^3은 C_2-C_5 알킬렌, 예를 들어, C_2-C_4 알킬렌, C_3 알킬렌 또는 C_4 알킬렌이다. 이들 실시형태 중 일부에서, R^8은 OH이다. 다른 실시형태에서, G^2는 존재하지 않고, R^7은 C_1-C_2 알킬렌, 예컨대, 메틸이다.

[0767] 다양한 상이한 실시형태에서, 화합물은 하기 표 5에 제시된 구조 중 하나를 갖는다.

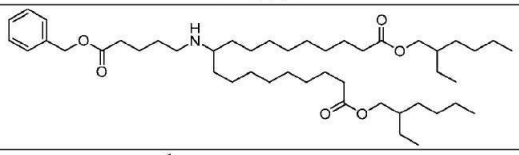
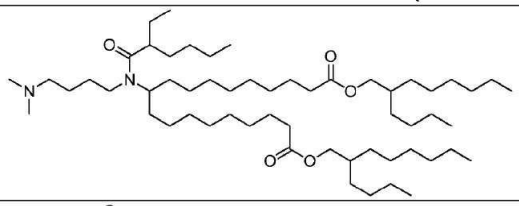
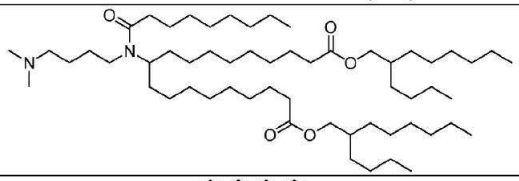
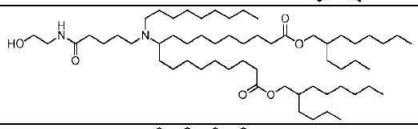
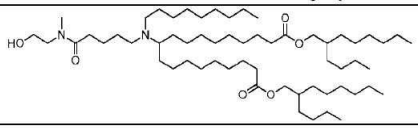
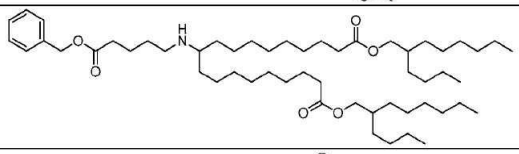
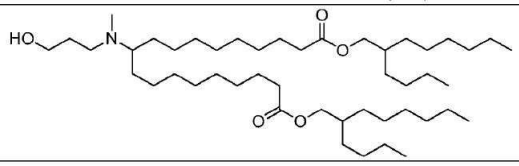
표 5

구조식 (VI)의 대표적인 양이온성 지질

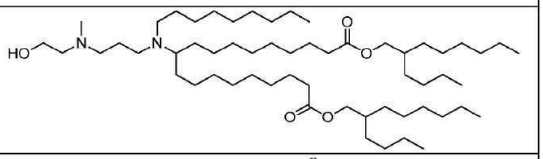
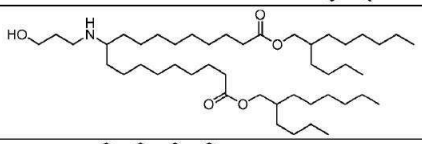
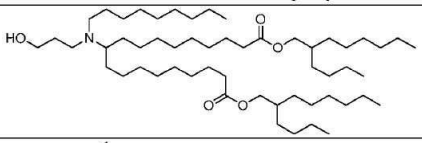
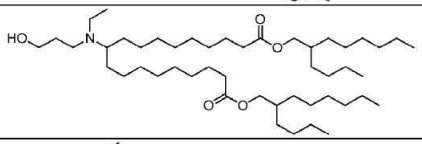
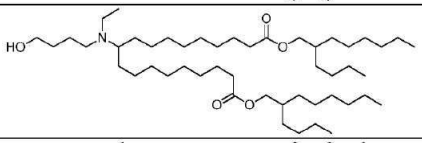
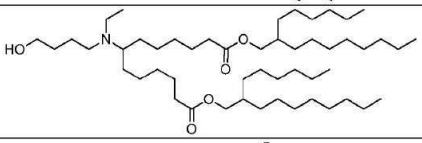
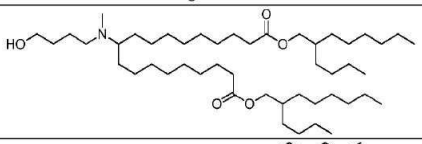
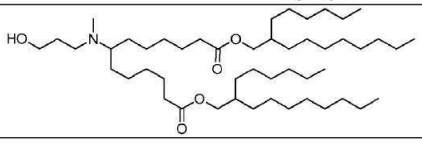
번호	구조
VI-1	
VI-2	

[0768]

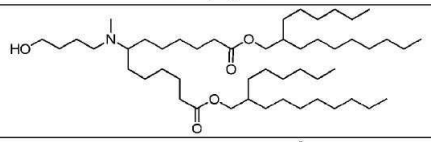
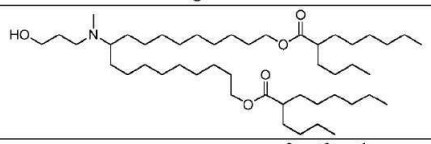
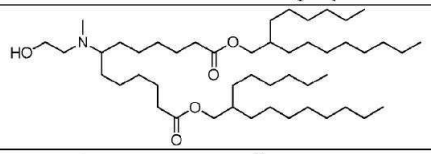
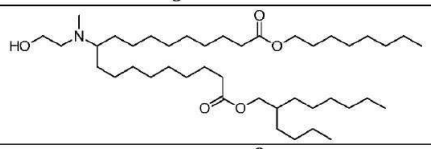
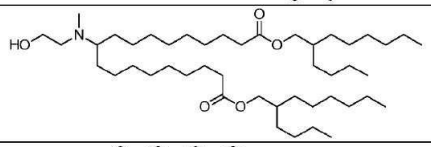
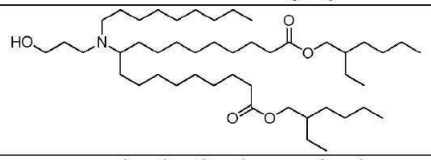
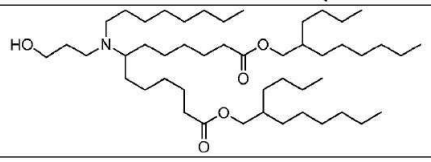
[0769]

번호	구조
VI-3	
VI-4	
VI-5	
VI-6	
VI-7	
VI-8	
VI-9	

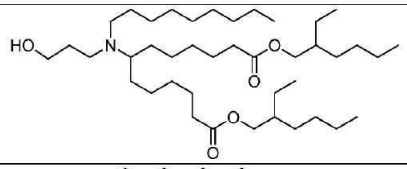
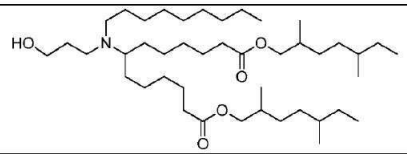
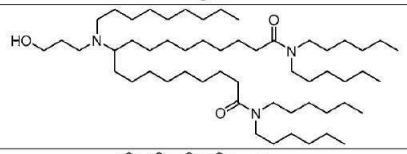
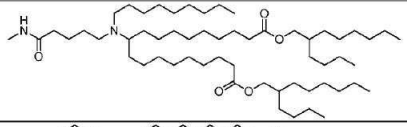
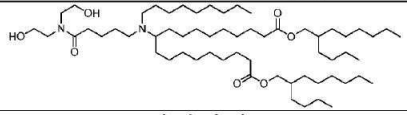
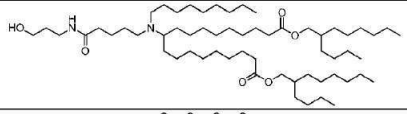
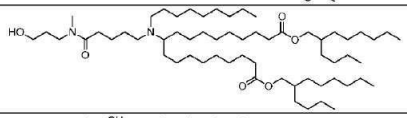
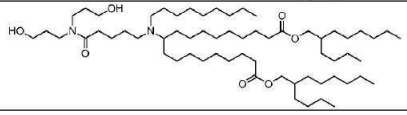
[0770]

번호	구조
VI-10	
VI-11	
VI-12	
VI-13	
VI-14	
VI-15	
VI-16	
VI-17	

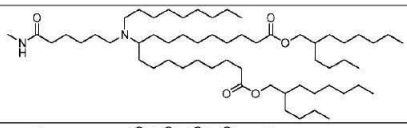
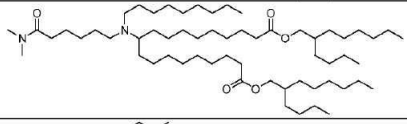
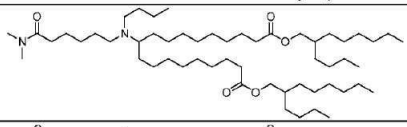
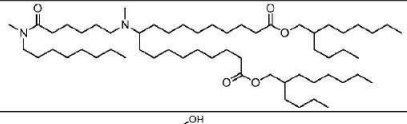
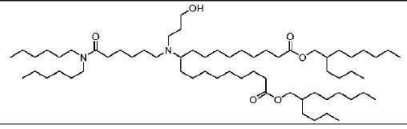
[0771]

번호	구조
VI-18	
VI-19	
VI-20	
VI-21	
VI-22	
VI-23	
VI-24	

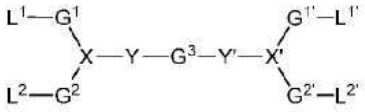
[0772]

번호	구조
VI-25	
VI-26	
VI-27	
VI-28	
VI-29	
VI-30	
VI-31	
VI-32	

[0773]

번호	구조
VI-33	
VI-34	
VI-35	
VI-36	
VI-37	

[0774] 일 실시형태에서, 양이온성 지질은 하기 구조식 (VII)을 갖는 화합물이다:



(VII)

[0775]

[0776] 또는 약제학적으로 허용 가능한 염, 전구약물 또는 입체이성질체, 식 중,

[0777] X 및 X^1은 각각 독립적으로 N 또는 CR이고;

[0778] Y 및 Y^1은 각각 독립적으로 존재하지 않거나, -O(C=O)-, -(C=O)O- 또는 NR이되,

[0779] 단,

[0780] a) X가 N인 경우 Y는 존재하지 않고;

[0781] b) X^1이 N인 경우 Y^1은 존재하지 않고;

[0782] c) X가 CR인 경우 Y는 -O(C=O)-, -(C=O)O- 또는 NR이고;

[0783] d) X'가 CR인 경우 Y'는 -O(C=O)-, -(C=O)O- 또는 NR이고;

[0784] L^1 및 $L^{1'}$는 각각 독립적으로 $-O(C=O)R^1$, $-(C=O)OR^1$, $-C(=O)R^1$, $-OR^1$, $-S(O)_zR^1$, $-S-SR^1$, $-C(=O)SR^x$, $-SC(=O)R^x$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ 또는 $-NR^aC(=O)OR^1$이고;

[0785] L^2 및 $L^{2'}$는 각각 독립적으로 $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_zR^2$, $-S-SR^2$, $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(-O)NR^cR^r$, $-NR^dC(=O)NR^eR^f$, $-OC(=O)NR^eR^f$; $-NR^dC(=O)OR^2$ 또는 R^2에 대한 직접 결합이고;

[0786] G^1, $G^{1'}$, G^2 및 $G^{2'}$는 각각 독립적으로 C_2-C_{12} 알킬렌 또는 C_2-C_{12} 알켄일렌이고;

[0787] G^3은 C_2-C_{24} 헤테로알킬렌 또는 C_2-C_{24} 헤테로알켄일렌이고;

[0788] R^a, R^b, R^d 및 R^e는, 각각의 경우에, 독립적으로 H, C_1-C_{12} 알킬 또는 C_2-C_{12} 알켄일이고;

[0789] R^c 및 R^f는, 각각의 경우에, 독립적으로 C_1-C_{12} 알킬 또는 C_2-C_{12} 알켄일이고;

[0790] R은, 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0791] R^1 및 R^2는, 각각의 경우에, 독립적으로 분지형 C_6-C_{24} 알킬 또는 분지형 C_6-C_{24} 알켄일이고;

[0792] z는 0, 1 또는 2이고,

[0793] 각각의 알킬, 알켄일, 알킬렌, 알켄일렌, 헤테로알킬렌 및 헤테로알켄일렌은 달리 명시되지 않는 한 독립적으로 치환되거나 치환되지 않는다.

[0794] 구조식 (VII)의 다른 상이한 실시형태에서:

[0795] X 및 X'는 각각 독립적으로 N 또는 CR이고;

[0796] Y 및 Y'는 각각 독립적으로 존재하지 않거나 NR이되, 단,

[0797] a) X가 N인 경우 Y는 존재하지 않고;

[0798] b) X'가 N인 경우 Y'는 존재하지 않고;

[0799] c) X가 CR인 경우 Y는 NR이고;

[0800] d) X'가 CR인 경우 Y'는 NR이고

[0801] L^1 및 $L^{1'}$는 각각 독립적으로 $-O(C=O)R^1$, $-(C=O)OR^1$, $-C(=O)R^1$, $-OR^1$, $-S(O)_zR^1$, $-S-SR^1$, $-C(=O)SR^x$, $-SC(=O)R^x$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ 또는 $-NR^aC(=O)OR^1$이고;

[0802] L^2 및 $L^{2'}$는 각각 독립적으로 $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_zR^2$, $-S-SR^2$, $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(-O)NR^cR^r$, $-NR^dC(=O)NR^eR^f$, $-OC(=O)NR^eR^f$; $-NR^dC(=O)OR^2$ 또는 R^2에 대한 직접 결합이고;

[0803] G^1, $G^{1'}$, G^2 및 $G^{2'}$는 각각 독립적으로 C_2-C_{12} 알킬렌 또는 C_2-C_{12} 알케닐렌이고;

[0804] G^3은 C_2-C_{24} 알킬렌옥사이드 또는 C_2-C_{24} 알케닐렌옥사이드이고;

[0805] R^a, R^b, R^d 및 R^e는, 각각의 경우에, 독립적으로 H, C_1-C_{12} 알킬 또는 C_2-C_{12} 알케닐이고;

[0806] R^c 및 R^f는, 각각의 경우에, 독립적으로 C_1-C_{12} 알킬 또는 C_2-C_{12} 알케닐이고;

[0807] R'는, 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0808] R^1 및 R^2는, 각각의 경우에, 독립적으로 분지형 C_6-C_{24} 알킬 또는 분지형 C_6-C_{24} 알케닐이고;

[0809] z는 0, 1 또는 2이고,

[0810] 각각의 알킬, 알케닐, 알킬렌, 알케닐렌, 알킬렌옥사이드 및 알케닐렌옥사이드는 달리 명시되지 않는 한 독립적으로 치환되거나 치환되지 않는다.

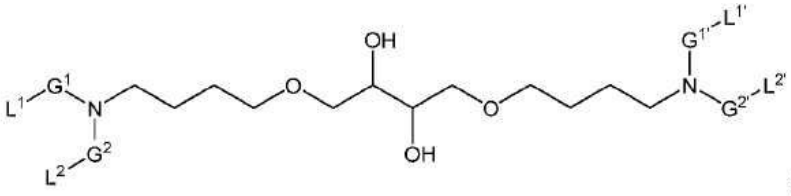
[0811] 구조식 (VII)의 일부 실시형태에서, G^3은 C_2-C_{24} 알킬렌옥사이드 또는 C_2-C_{24} 알케닐렌옥사이드이다. 특정 실시형태에서, G^3은 치환되지 않는다. 다른 실시형태에서, G^3은 치환되고, 예를 들어, 하이드록실로 치환된다. 보다 구체적인 실시형태에서 G^3은 C_2-C_{12} 알킬렌옥사이드이고, 예를 들어, 일부 실시형태에서 G^3은 C_3-C_7 알킬렌옥사이드이거나 다른 실시형태에서 G^3은 C_3-C_{12} 알킬렌옥사이드이다.

[0812] 구조식 (VII)의 다른 실시형태에서, G^3은 C_2-C_{24} 알킬렌아미닐 또는 C_2-C_{24} 알케닐렌아미닐, 예를 들어, C_6-C_{12} 알킬렌아미닐이다. 이들 실시형태 중 일부에서, G^3은 치환되지 않는다. 이들 실시형태 중 다른 것에서, G^3은 C_1-C_6 알킬로 치환된다.

[0813] 구조식 (VII)의 일부 실시형태에서, X 및 X'는 각각 N이고, Y 및 Y'는 각각 존재하지 않는다. 다른 실시형태에서, X 및 X'는 각각 CR이고, Y 및 Y'는 각각 NR이다. 이들 실시형태 중 일부에서, R은 H이다.

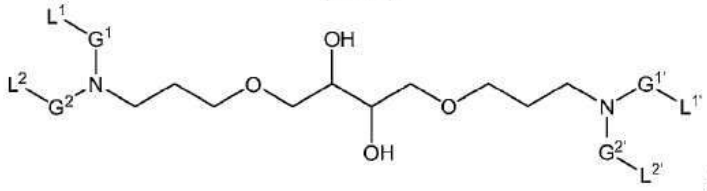
[0814] 구조식 (VII)의 특정 실시형태에서, X 및 X'는 각각 CR이고, Y 및 Y'는 각각 독립적으로 $-O(C=O)-$ 또는 $-(C=O)O-$이다.

[0815] 구조식 (VII)의 상기 실시형태 중 일부에서, 화합물은 하기 구조식 (VIIA), (VIIB), (VIIC), (VIID), (VIIE), (VIIF), (VIIG) 또는 (VIIH) 중 하나를 갖는다:



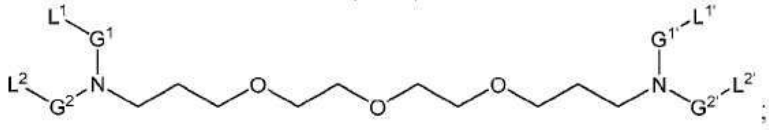
;

(VIIA)



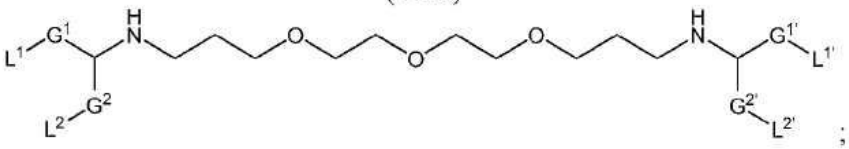
;

(VIIB)



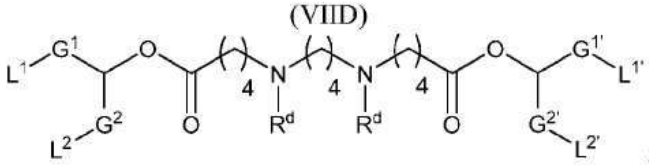
;

(VIIC)



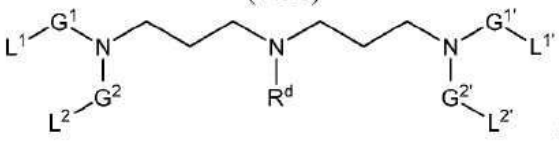
;

(VIID)



;

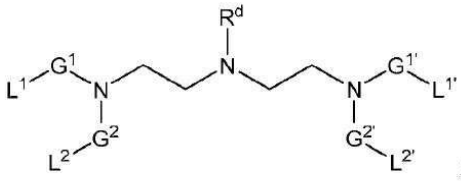
(VIIE)



;

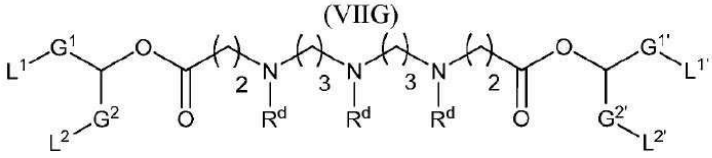
(VIIF)

[0816]



; 또는

(VIIG)



(VIIH)

[0817]

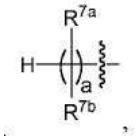
[0818] 식 중, R^d는, 각각의 경우에, 독립적으로 H 또는 선택적으로 치환된 C_1-C_6 알킬이다. 예를 들어, 일부 실시형태에서 R^d는 H이다. 다른 실시형태에서, R^d는 C_1-C_6 알킬, 예컨대, 메틸이다. 다른 실시형태에서, R^d는 치환된 C_1-C_6 알킬, 예컨대, -O(C=O)R, -(C=O)OR, -NRC(=O)R 또는 $-C(=O)N(R)_2$로 치환된 C_1-C_6 알킬이되,

[0819] R는, 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이다.

[0820] 구조식 (VII)의 상기 실시형태 중 일부에서, L^1 및 $L^{1'}$는 각각 독립적으로 $-O(C=O)R^1$, $-(C=O)OR^1$ 또는 $-C(=O)NR^bR^c$이고, L^2 및 $L^{2'}$는 각각 독립적으로 $-O(C=O)R^2$, $-(C=O)OR^2$ 또는 $-C(=O)NR^eR^f$이다. 예를 들어, 일부 실시형태에서 L^1 및 $L^{1'}$는 각각 $-(C=O)OR^1$이고, L^2 및 $L^{2'}$는 각각 $-(C=O)OR^2$이다. 다른 실시형태에서 L^1 및 $L^{1'}$는 각각 $-(C=O)OR^1$이고, L^2 및 $L^{2'}$는 각각 $-C(=O)NR^eR^f$이다. 다른 실시형태에서 L^1 및 $L^{1'}$는 각각 $-C(=O)NR^bR^c$이고, L^2 및 $L^{2'}$는 각각 $-C(=O)NR^eR^f$이다.

[0821] 상기 중 일부 실시형태에서, G^1, $G^{1'}$, G^2 및 $G^{2'}$는 각각 독립적으로 C_2-C_8 알킬렌, 예를 들어, C_4-C_8 알킬렌이다.

[0822] 구조식 (VII)의 상기 실시형태 중 일부에서, R^1 또는 R^2는 각각, 각각의 경우에, 독립적으로 분지형 C_6-C_{24} 알킬이다. 예를 들어, 일부 실시형태에서, R^1 및 R^2는 각각의 경우에, 독립적으로 하기 구조를 갖는다:

[0823] 

[0824] 식 중,

[0825] R^{7a} 및 R^{7b}는, 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0826] a는 2 내지 12이고,

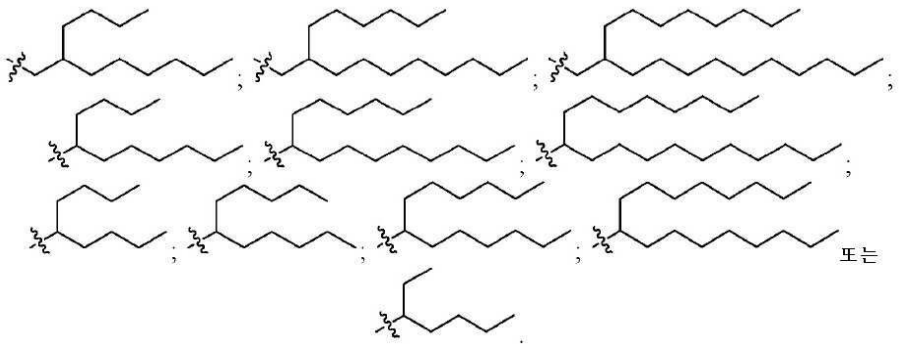
[0827] R^{7a}, R^{7b} 및 a는 R^1 및 R^2가 각각 독립적으로 6 내지 20개의 탄소 원자를 포함하도록 각각 선택된다.

[0828] 예를 들어, 일부 실시형태에서 a는 5 내지 9 또는 8 내지 12 범위의 정수이다.

[0829] 구조식 (VII)의 상기 실시형태 중 일부에서, R^{7a}의 적어도 하나의 경우는 H이다. 예를 들어, 일부 실시형태에서, R^{7a}는 각각의 경우에 H이다. 상기 중 다른 상이한 실시형태에서, R^{7b} 중 적어도 하나의 경우는 C_1-C_8 알킬이다.

[0830] 예를 들어, 일부 실시형태에서, C_1-C_8 알킬은 메틸, 에틸, n-프로필, 아이소프로필, n-부틸, 아이소-부틸, tert-부틸, n-헥실 또는 n-옥틸이다.

[0831] 구조식 (VII)의 상이한 실시형태에서, R^1 또는 R^2, 또는 이들 둘 다는 각각의 경우에 독립적으로 하기 구조 중 하나를 갖는다:

[0832] 

[0833] 구조식 (VII)의 상기 실시형태 중 일부에서, 존재하는 경우 R^b, R^c, R^e 및

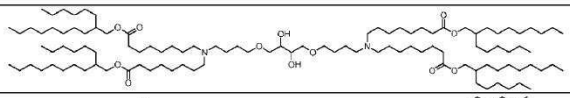
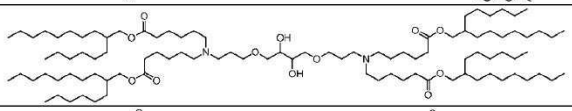
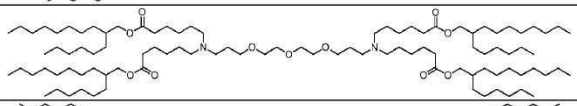
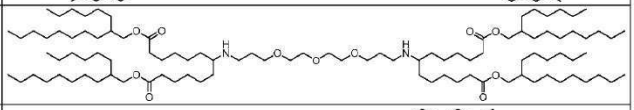
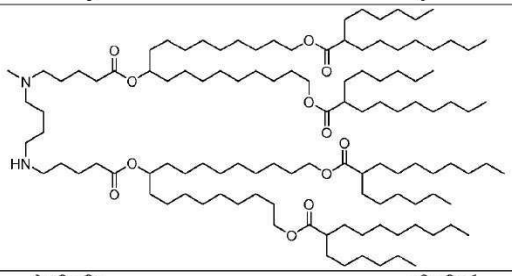
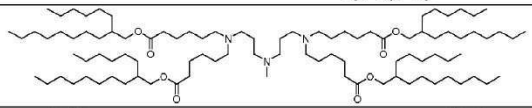
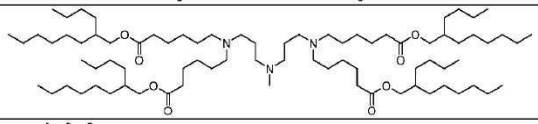
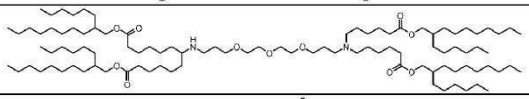
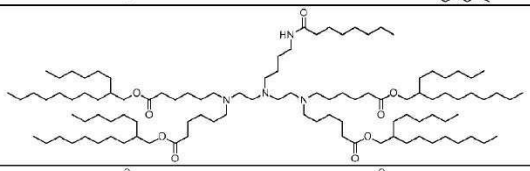
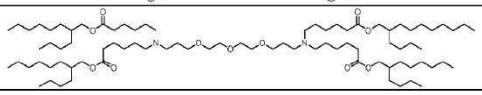
[0834] R^f는 각각 독립적으로 C_3-C_{12} 알킬이다. 예를 들어, 일부 실시형태에서, 존재하는 경우 R^b, R^c, R^e 및 R^f는 n-헥실

이고, 다른 실시형태에서, 존재하는 경우 R^b, R^c, R^e 및 R^f는 n-옥틸이다.

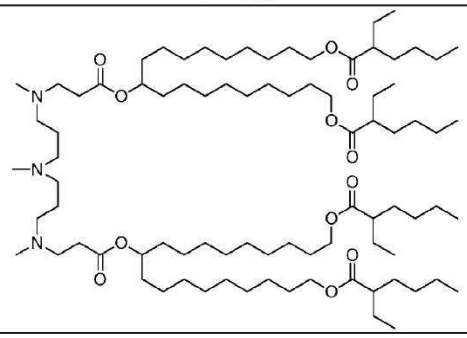
[0835] 구조식 (VII)의 다양한 상이한 실시형태에서, 양이온성 지질은 하기 표 6에 제시된 구조 중 하나를 갖는다.

표 6

구조식 (VII)의 대표적인 양이온성 지질

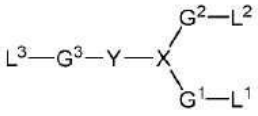
번호	구조
VII-1	
VII-2	
VII-3	
VII-4	
VII-5	
VII-6	
VII-7	
VII-8	
VII-9	
VII-10	

[0836]

번호	구조
VII-11	

[0837]

[0838] 일 실시형태에서, 양이온성 지질은 하기 구조식 (VIII)을 갖는 화합물, 또는 이의 약제학적으로 허용 가능한 염, 전구약물 또는 입체이성질체이다:



(VIII)

[0839]

[0840] 식 중,

[0841] X는 N이고, Y는 존재하지 않거나; X는 CR이고, Y는 NR이고;

[0842] L^1은 $-O(C=O)R^1$, $-(C=O)OR^1$ $-C(=O)R^1$, $-OR^1$, $-S(O)_xR^1$, $-S-SR^1$, $-C(=O)SR^1$, $-SC(=O)R^1$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ 또는 $-NR^aC(=O)OR^1$이고;

[0843] L^2는 $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_xR^2$, $-S-SR^2$, $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(=O)NR^eR^f$, $-NR^dC(=O)NR^eR^f$, $-OC(=O)NR^eR^f$; $-NR^dC(=O)OR^2$ 또는 R^2에 대한 직접 결합이고;

[0844] L^3은 $-O(C=O)R^3$ 또는 $-(C=O)OR^3$이고;

[0845] G^1 및 G^2는 각각 독립적으로 C_2-C_{12} 알킬렌 또는 C_2-C_{12} 알켄일렌이고;

[0846] G^3 은 C_1-C_{24} 알킬렌, C_2-C_{24} 알켄일렌, C_1-C_{24} 헤테로알킬렌 또는 C_2-

[0847] C_{24} 헤테로알켄일렌이고;

[0848] R^a, R^b, R^d 및 R^e는 각각 독립적으로 H 또는 C_1-C_{12} 알킬 또는 C_1-C_{12} 알켄일이고;

[0849] R^c 및 R^f는 각각 독립적으로 C_1-C_{12} 알킬 또는 C_2-C_{12} 알켄일이고;

[0850] 각각의 R은 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0851] R^1, R^2 및 R^3은 각각 독립적으로 C_1-C_{24} 알킬 또는 C_2-C_{24} 알켄일이고;

[0852] x는 0, 1 또는 2이고,

[0853] 각각의 알킬, 알켄일, 알킬렌, 알켄일렌, 헤테로알킬렌 및 헤테로알켄일렌은 달리 명시되지 않는 한 독립적으로 치환되거나 치환되지 않는다.

[0854] 구조식 (1)의 추가의 실시형태에서,

[0855] X는 N이고, Y는 존재하지 않거나; X는 CR이고, Y는 NR이고;

[0856] L^1은 $-O(C=O)R^1$, $-(C=O)OR^1$ $-C(=O)R^1$, $-OR^1$, $-S(O)_XR^1$, $-S-SR^1$, $-C(=O)SR^1$, $-SC(=O)R^1$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ 또는 $-NR^aC(=O)OR^1$이고;

[0857] L^2는 $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_XR^2$, $-S-SR^2$, $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(=O)NR^eR^f$, $-NR^dC(=O)NR^eR^f$, $-OC(=O)NR^eR^f$; $-NR^dC(=O)OR^2$ 또는 R^2에 대한 직접 결합이고;

[0858] L^3은 $-O(C=O)R^3$ 또는 $-(C=O)OR^3$이고;

[0859] G^1 및 G^2는 각각 독립적으로 C_2-C_{12} 알킬렌 또는 C_2-C_{12} 알켄일렌이고;

[0860] X가 CR이고, Y가 NR인 경우 G^3은 C_1-C_{24} 알킬렌, C_1-C_{24} 알켄일렌, C_1-C_{24} 헤테로알킬렌 또는 C_2-C_{24} 헤테로알켄일렌이고; X가 N이고 Y가 존재하지 않는 경우 G^3은 C_1-C_{24} 헤테로알킬렌 또는 C_2-C_{24} 헤테로알켄일렌이고;

[0861] R^a, R^b, R^d 및 R^e는 각각 독립적으로 H 또는 C_1-C_{12} 알킬 또는 C_1-C_{12} 알켄일이고;

[0862] R^c 및 R^f는 각각 독립적으로 C_1-C_{12} 알킬 또는 C_2-C_{12} 알켄일이고;

[0863] 각각의 R은 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0864] R^1, R^2 및 R^3은 각각 독립적으로 C_1-C_{24} 알킬 또는 C_2-C_{24} 알켄일이고;

[0865] x는 0, 1 또는 2이고,

[0866] 각각의 알킬, 알켄일, 알킬렌, 알켄일렌, 헤테로알킬렌 및 헤테로알켄일렌은 달리 명시되지 않는 한 독립적으로 치환되거나 치환되지 않는다.

[0867] 구조식 (I)의 다른 실시형태에서,

[0868] X는 N이고, Y는 존재하지 않거나, X는 CR이고, Y는 NR이고;

[0869] L^1은 $-O(C=O)R^1$, $-(C=O)OR^1$ $-C(=O)R^1$, $-OR^1$, $-S(O)_XR^1$, $-S-SR^1$, $-C(=O)SR^1$, $-SC(=O)R^1$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ 또는 $-NR^aC(=O)OR^1$이고;

[0870] L^2는 $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_XR^2$, $-S-SR^2$, $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(=O)NR^eR^f$, $-NR^dC(=O)NR^eR^f$, $-OC(=O)NR^eR^f$; $-NR^dC(=O)OR^2$ 또는 R^2에 대한 직접 결합이고;

[0871] L^3은 $-O(C=O)R^3$ 또는 $-(C=O)OR^3$이고;

[0872] G^1 및 G^2는 각각 독립적으로 C_2-C_{12} 알킬렌 또는 C_2-C_{12} 알켄일렌이고;

[0873] G^3은 C_1-C_{24} 알킬렌, C_1-C_{24} 알켄일렌, C_1-C_{24} 헤테로알킬렌 또는 C_2-C_{24} 헤테로알켄일렌이고;

[0874] R^a, R^b, R^d 및 R^e는 각각 독립적으로 H 또는 C_1-C_{12} 알킬 또는 C_1-C_{12} 알켄일이고;

[0875] R^c 및 R^f는 각각 독립적으로 C_1-C_{12} 알킬 또는 C_2-C_{12} 알켄일이고;

[0876] 각각의 R은 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0877] R^1, R^2 및 R^3은 각각 독립적으로 분지형 C_6-C_{24} 알킬 또는 분지형 C_6-C_{24} 알켄일이고;

[0878] x는 0, 1 또는 2이고,

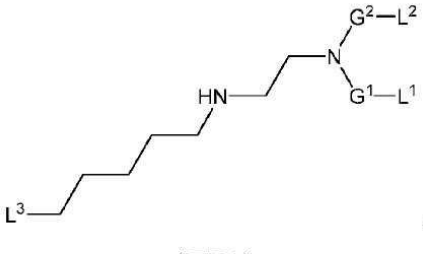
[0879] 각각의 알킬, 알켄일, 알킬렌, 알켄일렌, 헤테로알킬렌 및 헤테로알켄일렌은 달리 명시되지 않는 한 독립적으로 치환되거나 치환되지 않는다.

[0880] 구조식 (VIII)의 특정 실시형태에서, G^3은 치환되지 않는다. 보다 구체적인 실시형태에서 G^3은 C_2-C_{12} 알킬렌이고, 예를 들어, 일부 실시형태에서 G^3은 C_3-C_7 알킬렌이거나 다른 실시형태에서 G^3은 C_3-C_{12} 알킬렌이다. 일부 실시형태에서, G^3은 C_2 또는 C_3 알킬렌이다.

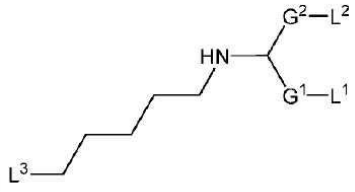
[0881] 구조식 (VIII)의 다른 실시형태에서, G^3은 C_1-C_{12} 헤테로알킬렌, 예를 들어, C_1-C_{12} 아미닐알킬렌이다.

[0882] 구조식 (VIII)의 특정 실시형태에서, X는 N이고, Y는 존재하지 않는다. 다른 실시형태에서, X는 CR이고, Y는 NR이며, 예를 들어, 이들 실시형태 중 일부에서 R은 H이다.

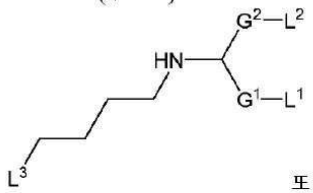
[0883] 구조식 (VIII)의 상기 실시형태 중 일부에서, 화합물은 하기 구조식 (VIIIA), (VIIIB), (VIIIC) 또는 (VIIID) 중 하나를 갖는다:



(VIIIA)

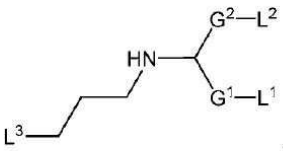


(VIIIB)



(VIIIC)

또는



(VIIID)

[0884]

[0885] 구조식 (VIII)의 상기 실시형태 중 일부에서, L^1은 $-O(C=O)R^1$, $-(C=O)OR^1$ 또는 $-C(=O)NR^bR^c$이고, L^2는 $-O(C=O)R^2$, $-(C=O)OR^2$ 또는 $-C(=O)NR^eR^f$이다.

[0886] 다른 구체적인 실시형태에서, L^1은 $-(C=O)OR^1$이고, L^2는 $-(C=O)OR^2$이다. 상기

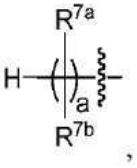
[0887] 실시형태 중 임의의 것에서, L^3은 $-(C=O)OR^3$이다.

[0888] 구조식 (VIII)의 상기 실시형태 중 일부에서, G^1 및 G^2는

[0889] 각각 독립적으로 C_2-C_{12} 알킬렌, 예를 들어, C_4-C_{10} 알킬렌이다.

[0890] 구조식 (VIII)의 상기 실시형태 중 일부에서, R^1, R^2 및 R^3

[0891] 은 각각 독립적으로 분지형 C_6-C_{24} 알킬이다. 예를 들어, 일부 실시형태에서, R^1, R^2 및 R^3은 각각 독립적으로 하기 구조를 갖는다:

[0892] 

[0893] 식 중,

[0894] R^{7a} 및 R^{7b}는, 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

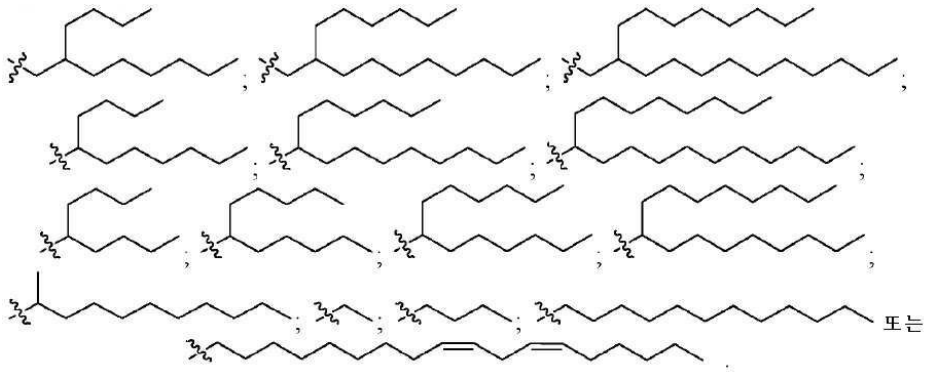
[0895] a는 2 내지 12이고,

[0896] R^{7a}, R^{7b} 및 a는 R^1 및 R^2가 각각 독립적으로 6 내지 20개의 탄소 원자를 포함하도록 각각 선택된다. 예를 들어, 일부 실시형태에서 a는 5 내지 9 또는 8 내지 12 범위의 정수이다.

[0897] 구조식 (VIII)의 상기 실시형태 중 일부에서, R^{7a}의 적어도 하나의 경우는 H이다. 예를 들어, 일부 실시형태에서, R^{7a}는 각각의 경우에 H이다. 상기 중 다른 상이한 실시형태에서, R^{7b} 중 적어도 하나의 경우는 C_1-C_8 알킬이다. 예를 들어, 일부 실시형태에서, C_1-C_8 알킬은 메틸, 에틸, n-프로필, 아이소프로필, n-부틸, 아이소-부틸, tert-부틸, n-헥실 또는 n-옥틸이다.

[0898] 구조식 (VIII)의 상기 실시형태 중 일부에서, X는 CR이고, Y는 NR이고, R^3은 C_1-C_{12} 알킬, 예컨대, 에틸, 프로필 또는 부틸이다. 이들 실시형태 중 일부에서, R^1 및 R^2는 각각 독립적으로 분지형 C_6-C_{24} 알킬이다.

[0899] 구조식 (VIII)의 상이한 실시형태에서, R^1, R^2 및 R^3은 각각 독립적으로 하기 구조 중 하나를 갖는다:



[0900]

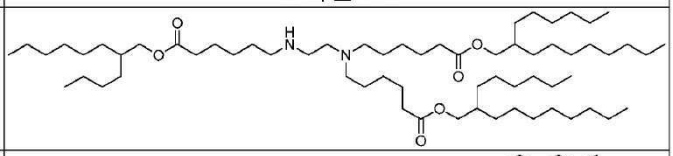
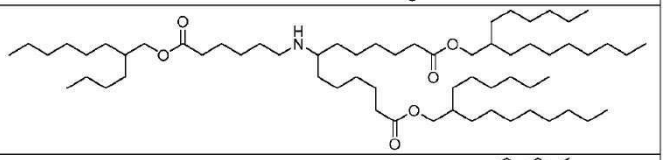
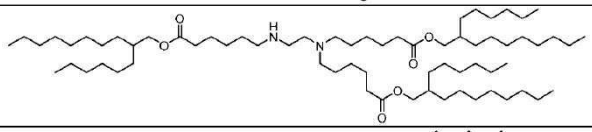
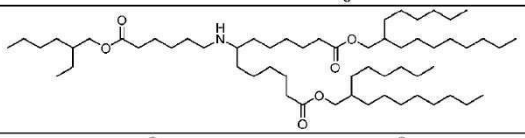
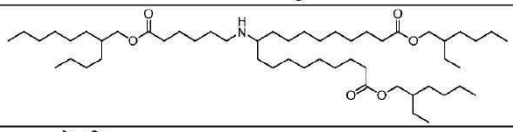
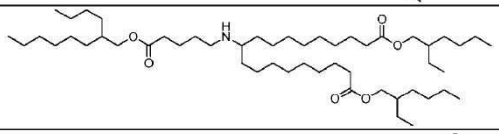
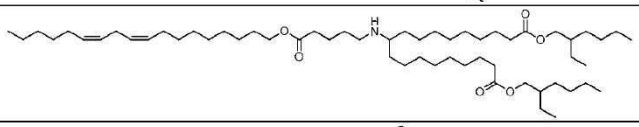
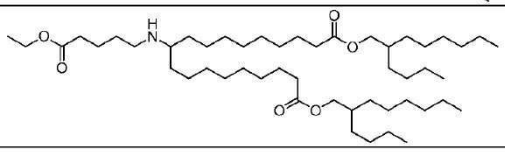
[0901] 구조식 (VIII)의 특정 실시형태에서, R^1 및 R^2 및 R^3은 각각 독립적으로, 분지형 C_6-C_{24} 알킬이고, R^3은 C_1-C_{24} 알킬 또는 C_2-C_{24} 알켄일이다.

[0902] 구조식 (VIII)의 상기 실시형태 중 일부에서, R^b, R^c, R^e 및 R^f는 각각 독립적으로 C_3-C_{12} 알킬이다. 예를 들어, 일부 실시형태에서 R^b, R^c, R^e 및 R^f는 n-헥실이고, 다른 실시형태에서 R^b, R^c, R^e 및 R^f는 n-옥틸이다.

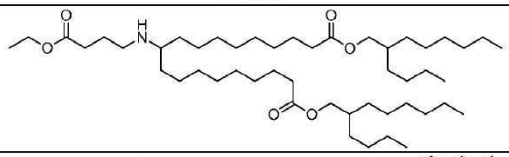
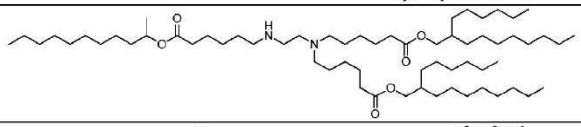
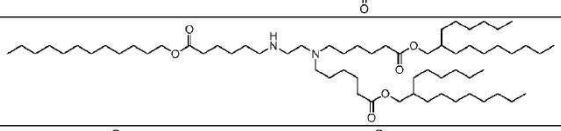
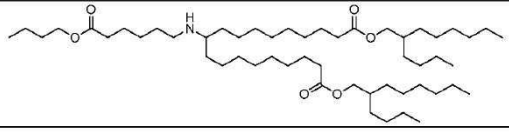
[0903] 구조식 (VIII)의 다양한 상이한 실시형태에서, 화합물은 하기 표 7에 제시된 구조 중 하나를 갖는다.

표 7

구조식 (VIII)의 대표적인 양이온성 지질

번호	구조
VIII-1	
VIII-2	
VIII-3	
VIII-4	
VIII-5	
VIII-6	
VIII-7	
VIII-8	

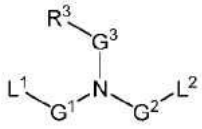
[0904]

번호	구조
VIII-9	
VIII-10	
VIII-11	
VIII-12	

[0905]

[0906] 일 실시형태에서, 양이온성 지질은 하기 구조식 (IX)를 갖는 화합물, 또는 이의 약제학적으로 허용 가능한 염,

전구약물 또는 입체이성질체이다:



(IX)

[0907]

[0908] 식 중,

[0909] L^1은 $-O(C=O)R^1$, $-(C=O)OR^1$ $-C(=O)R^1$, $-OR^1$, $-S(O)_xR^1$, $-S-SR^1$, $-C(=O)SR^1$, $-SC(=O)R^1$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ 또는 $-NR^aC(=O)OR^1$이고;

[0910] L^2는 $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_xR^2$, $-S-SR^2$, $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(=O)NR^eR^f$, $-NR^dC(=O)NR^eR^f$, $-OC(=O)NR^eR^f$; $-NR^dC(=O)OR^2$ 또는 R^2에 대한 직접 결합이고;

[0911] G^1 및 G^2는 각각 독립적으로 C_2-C_{12} 알킬렌 또는 C_2-C_{12} 알켄일렌이고;

[0912] G^3은 C_1-C_{24} 알킬렌, C_2-C_{24} 알켄일렌, C_3-C_8 사이클로알킬렌 또는 C_3-C_8 사이클로알켄일렌이고;

[0913] R^a, R^b, R^d 및 R^e는 각각 독립적으로 H 또는 C_1-C_{12} 알킬 또는 C_1-C_{12} 알켄일이고;

[0914] R^c 및 R^f는 각각 독립적으로 C_1-C_{12} 알킬 또는 C_2-C_{12} 알켄일이고;

[0915] R^1 및 R^2는 각각 독립적으로 분지형 C_6-C_{24} 알킬 또는 분지형 C_6-C_{24} 알켄일이고;

[0916] R^3은 $-N(R^4)R^5$이고;

[0917] R^4는 C_1-C_{12} 알킬이고;

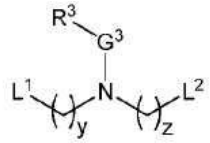
[0918] R^5는 치환된 C_1-C_{12} 알킬이고;

[0919] x는 0, 1 또는 2이고,

[0920] 각각의 알킬, 알켄일, 알킬렌, 알켄일렌, 사이클로알킬렌, 사이클로알켄일렌, 아릴 및 아르알킬은 달리 명시되지 않는 한 독립적으로 치환되거나 치환되지 않는다.

[0921] 구조식 (XI)의 특정 실시형태에서, G^3은 치환되지 않는다. 보다 구체적인 실시형태에서 G^3은 C_2-C_{12} 알킬렌이고, 예를 들어, 일부 실시형태에서 G^3은 C_3-C_7 알킬렌이거나 다른 실시형태에서 G^3은 C_3-C_{12} 알킬렌이다. 일부 실시형태에서, G^3은 C_2 또는 C_3 알킬렌이다.

[0922] 화학식 (IX)의 상기 실시형태 중 일부에서, 화합물은 하기 구조식 (IXA)를 갖는다:



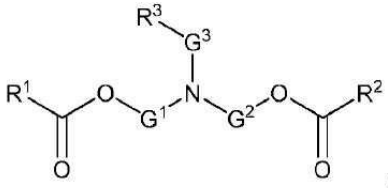
(IXA)

[0923]

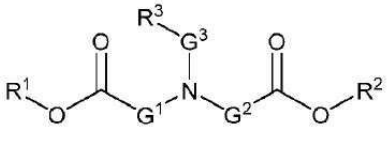
[0924] 식 중, y 및 z는 각각 독립적으로 2 내지 12의 정수, 예를 들어, 2 내지 6, 4 내지 10의 정수, 또는 예를 들어, 4 또는 5이다. 특정 실시형태에서, y 및 z는 각각 동일하고, 4, 5, 6, 7, 8 및 9로부터 선택된다.

[0925] 구조식 (VIII)의 상기 실시형태 중 일부에서, L^1은 $-O(C=O)R^1$, $-(C=O)OR^1$ 또는 $-C(=O)NR^bR^c$이고, L^2는 $-O(C=O)R^2$, $-(C=O)OR^2$ 또는 $-C(=O)NR^eR^f$이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2는 각각 $-(C=O)OR^1$ 및 $-(C=O)OR^2$이다. 다른 실시형태에서 L^1은 $-(C=O)OR^1$이고, L^2는 $-C(=O)NR^eR^f$이다. 다른 실시형태에서 L^1은 $-C(=O)NR^bR^c$이고, L^2는 $-C(=O)NR^eR^f$이다.

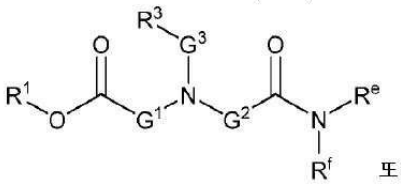
[0926] 상기의 다른 실시형태에서, 화합물은 하기 구조식 (IXB), (IXC), (IXD) 또는 (IXE) 중 하나를 갖는다:



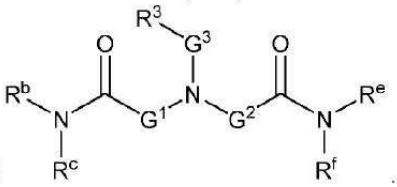
; (IXB)



(IXC)



또는

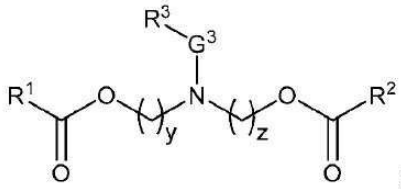


(IXD) (IXE)

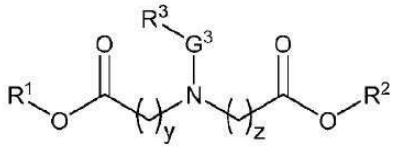
[0927]

[0928] 상기 실시형태 중 일부에서, 화합물은 구조식 (IXB)를 갖고, 다른 실시형태에서, 화합물은 구조식 (IXC)를 갖고, 추가의 다른 실시형태에서 화합물은 구조식 (IXD)를 갖는다. 다른 실시형태에서, 화합물은 구조식 (IXE)를 갖는다.

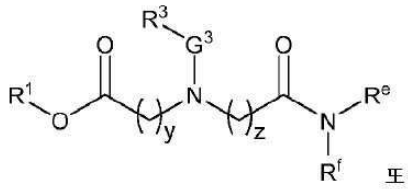
[0929] 상기의 일부 상이한 실시형태에서, 화합물은 하기 구조식 (IXF), (IXG), (IXH) 또는 (IXJ) 중 하나를 갖는다:



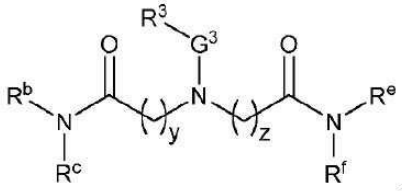
; (IXF)



(IXG)



또는



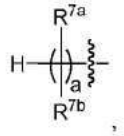
(IXH) (IXJ)

[0930]

[0931] 식 중, y 및 z는 각각 독립적으로 2 내지 12의 정수, 예를 들어, 2 내지 6, 예를 들어, 4이다.

[0932] 구조식 (IX)의 상기 실시형태 중 일부에서, y 및 z는 각각 독립적으로 2 내지 10, 2 내지 8, 4 내지 10 또는 4 내지 7 범위의 정수이다. 예를 들어, 일부 실시형태에서, y는 4, 5, 6, 7, 8, 9, 10, 11 또는 12이다. 일부 실시형태에서, z는 4, 5, 6, 7, 8, 9, 10, 11 또는 12이다. 일부 실시형태에서, y 및 z는 동일하지만, 다른 실시형태에서 y 및 z는 상이하다.

[0933] 구조식 (IX)의 상기 실시형태 중 일부에서, R^1 또는 R^2, 또는 둘 다는 분지형 C_6-C_{24} 알킬이다. 예를 들어, 일부 실시형태에서, R^1 및 R^2는 각각 독립적으로 하기 구조를 갖는다:

[0934] 

[0935] 식 중,

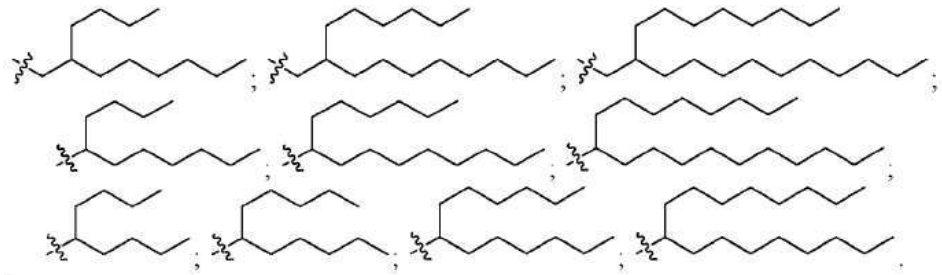
[0936] R^{7a} 및 R^{7b}는 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

[0937] a는 2 내지 12이고,

[0938] R^{7a}, R^{7b} 및 a는 R^1 및 R^2가 각각 독립적으로 6 내지 20개의 탄소 원자를 포함하도록 각각 선택된다. 예를 들어, 일부 실시형태에서 a는 5 내지 9 또는 8 내지 12 범위의 정수이다.

[0939] 구조식 (IX)의 상기 실시형태 중 일부에서, R^{7a}의 적어도 하나의 경우는 H이다. 예를 들어, 일부 실시형태에서, R^{7a}는 각각의 경우에 H이다. 상기 중 다른 상이한 실시형태에서, R^{7b} 중 적어도 하나의 경우는 C_1-C_8 알킬이다. 예를 들어, 일부 실시형태에서, C_1-C_8 알킬은 메틸, 에틸, n-프로필, 아이소프로필, n-부틸, 아이소-부틸, tert-부틸, n-헥실 또는 n-옥틸이다.

[0940] 구조식 (IX)의 상이한 실시형태에서, R^1 또는 R^2, 또는 이들 둘 다는 하기 구조 중 하나를 갖는다:

[0941] 

[0942] 구조식 (IX)의 상기 실시형태 중 일부에서, R^b, R^c, R^e 및 R^f는 각각 독립적으로 C_3-C_{12} 알킬이다. 예를 들어, 일부 실시형태에서 R^b, R^c, R^e 및 R^f는 n-헥실이고, 다른 실시형태에서 R^b, R^c, R^e 및 R^f는 n-옥틸이다.

[0943] 구조식 (IX)의 상기 실시형태 중 임의의 것에서, R^4는 치환된 또는 비치환된 메틸, 에틸, 프로필, n-부틸, n-헥실, n-옥틸 또는 n-노닐이다. 예를 들어, 일부 실시형태에서 R^4는 치환되지 않는다. 다른 실시형태에서, R^4는 $-OR^g$, $-NR^gC(=O)R^h$, $-C(=O)NR^gR^h$, $-C(=O)R^h$, $-OC(=O)R^h$, $-C(=O)OR^h$ 및 $-OR^iOH$로 이루어진 군으로부터 선택된 하나 이상의 치환체로 치환되고,

[0944] R^g는, 각각의 경우에 독립적으로 H 또는 C_1-C_6 알킬이고;

[0945] R^h는 각각의 경우에 독립적으로 C_1-C_6 알킬이고;

[0946] R^i는 각각의 경우에 독립적으로 C_1-C_6 알킬렌이다.

[0947] 구조식 (IX)의 상기 실시형태 중 임의의 것에서, R^5는 치환된 메틸, 에틸, 프로필, n-부틸, n-헥실, n-옥틸 또는 n-노닐이다. 일부 실시형태에서, R^5는 치환된 에틸 또는 치환된 프로필이다. 다른 상이한 실시형태에서, R^5는 하이드록실로 치환된다. 더 추가의 실시형태에서, R^5는 $-OR^g$, $-NR^gC(=O)R^h$, $-C(=O)NR^gR^h$, $-C(=O)R^h$, $-OC(=O)R^h$, $-C(=O)OR^h$ 및 $-OR^iOH$로 이루어진 군으로부터 선택된 하나 이상의 치환체로 치환되고,

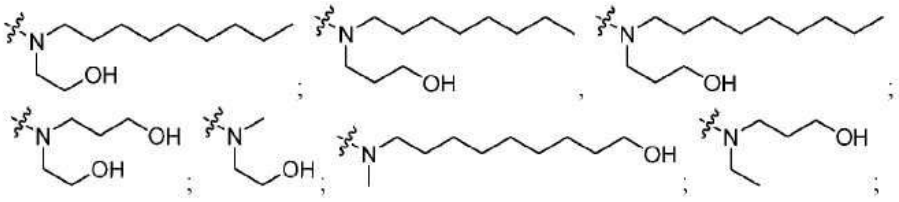
[0948] R^g는, 각각의 경우에 독립적으로 H 또는 C_1-C_6 알킬이고;

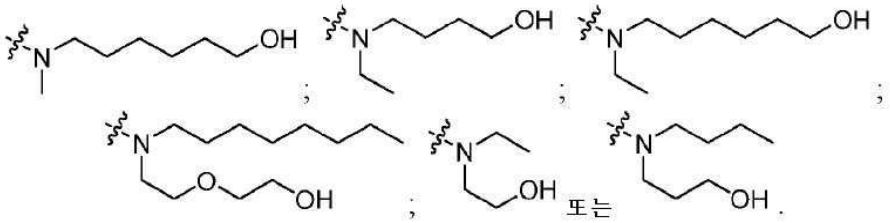
[0949] R^h는 각각의 경우에 독립적으로 C_1-C_6 알킬이고;

[0950] R^i는, 각각의 경우에 독립적으로 C_1-C_6 알킬렌이다.

[0951] 구조식 (IX)의 다른 실시형태에서, R^4는 비치환된 메틸이고, R^5는 치환된 메틸, 에틸, 프로필, n-부틸, n-헥실, n-옥틸 또는 n-노닐이다. 이들 실시형태 중 일부에서, R^5는 하이드록실로 치환된다.

[0952] 구조식 (IX)의 일부 다른 구체적인 실시형태에서, R^3은 하기 구조 중 하나를 갖는다:

[0953] 

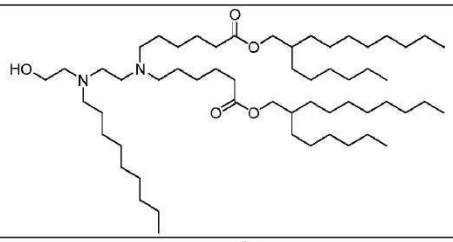
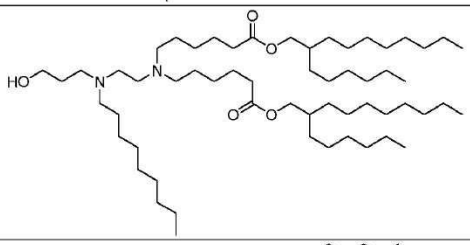
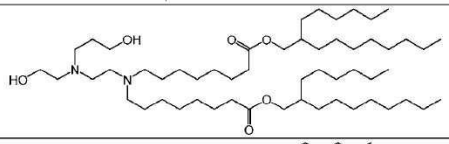
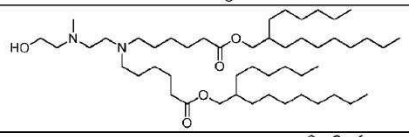
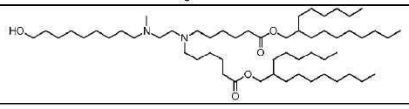


[0954] 또는

[0955] 구조식 (IX)의 다양한 상이한 실시형태에서, 양이온성 지질은 하기 표 8에 제시된 구조 중 하나를 갖는다.

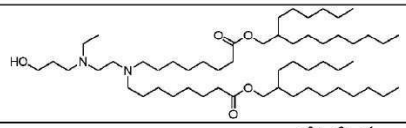
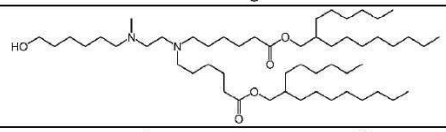
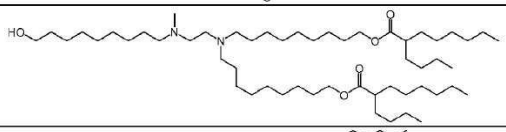
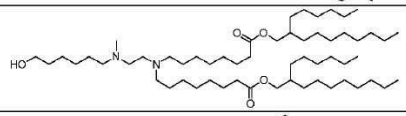
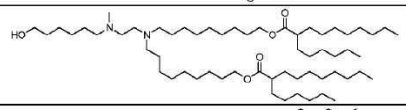
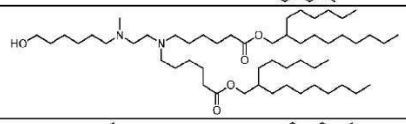
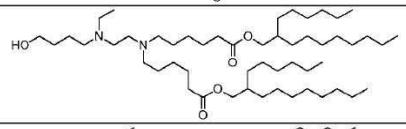
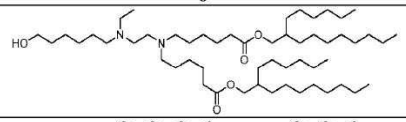
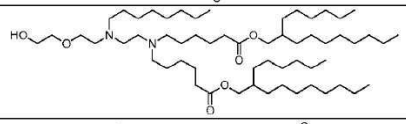
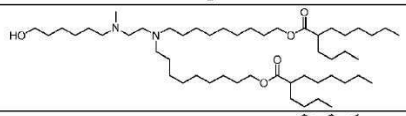
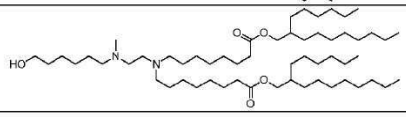
표 8

구조식 (IX)의 대표적인 양이온성 지질

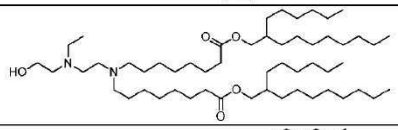
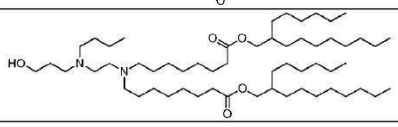
번호	구조
IX-1	
IX-2	
IX-3	
IX-4	
IX-5	

[0956]

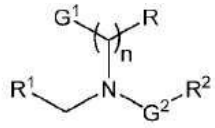
[0957]

번호	구조
IX-6	
IX-7	
IX-8	
IX-9	
IX-10	
IX-11	
IX-12	
IX-13	
IX-14	
IX-15	
IX-16	

[0958]

번호	구조
IX-17	
IX-18	

[0959] 일 실시형태에서, 양이온성 지질은 하기 구조식 (X)을 갖는 화합물, 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체이다:



(X)

[0960]

[0961] 식 중,

[0962] G^1은 -OH, $-NR^3R^4$, $-(C=O)NR^5$ 또는 $-NR^3(C=O)R^5$이고;

[0963] G^2는 $-CH_2-$ 또는 -(C=O)-이고;

[0964] R은, 각각의 경우에, 독립적으로 H 또는 OH이고;

[0965] R^1 및 R^2는 각각 독립적으로 분지형, 포화 또는 불포화 C_{12}-C_{36} 알킬이고;

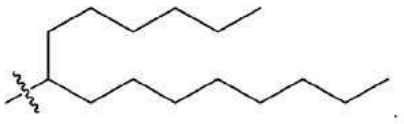
[0966] R^3 및 R^4는 각각 독립적으로 H 또는 선형 또는 분지형, 포화 또는 불포화 C_1-C_6 알킬이고;

[0967] R^5는 선형 또는 분지형, 포화 또는 불포화 C_1-C_6 알킬이고;

[0968] n은 2 내지 6의 정수이다.

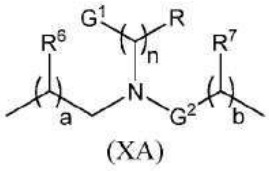
[0969] 일부 실시형태에서, R^1 및 R^2는 각각 독립적으로 분지형, 포화 또는 불포화 C_{12}-C_{30} 알킬, C_{12}-C_{20} 알킬 또는 C_{15}-C_{20} 알킬이다. 일부 구체적인 실시형태에서, R^1 및 R^2는 각각 포화된다. 특정 실시형태에서, R^1 및 R^2 중 적어도 하나는 불포화된다.

[0970] 구조식 (X)의 상기 실시형태 중 일부에서, R^1 및 R^2는 하기 구조를 갖는다:



[0971]

[0972] 화학식 (X)의 상기 실시형태 중 일부에서, 화합물은 하기 구조식 (XA)를 갖는다:



(XA)

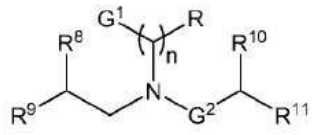
[0973]

[0974] 식 중,

[0975] R^6 및 R^7은, 각각의 경우에, 독립적으로 H 또는 선형 또는 분지형, 포화 또는 불포화 C_1-C_{14} 알킬이고;

[0976] a 및 b는 각각 독립적으로 1 내지 15 범위의 정수이되, 단, R^6 및 a, 및 R^7 및 b는, R^1 및 R^2 각각이 각각 독립적으로 분지형, 포화 또는 불포화 C_{12}-C_{36} 알킬이도록 각각 독립적으로 선택된다.

[0977] 상기 실시형태 중 일부에서, 화합물은 하기 구조식 (XB)를 갖는다:



(XB)

[0978]

[0979] 식 중,

[0980] R^8, R^9, R^{10} 및 R^{11}은 각각 독립적으로 선형 또는 분지형, 포화 또는 불포화 C_4-C_{12} 알킬이되, R^8 및 R^9, 및 R^{10} 및 R^{11}은 R^1 및 R^2 각각이 각각 독립적으로 분지형, 포화 또는 불포화 C_{12}-C_{36} 알킬이도록 각각 독립적으로 선택된다. (XB)의 일부 실시형태에서, R^8, R^9, R^{10} 및 R^{11}은 각각 독립적으로 선형 또는 분지형, 포화 또는 불포화 C_6-C_{10} 알킬이다. (XB)의 특정 실시형태에서, R^8, R^9, R^{10} 및 R^{11} 중 적어도 하나는 불포화된다. (XB)의 다른 특정 구체적인 실시형태에서, R^8, R^9, R^{10} 및 R^{11} 각각은 포화된다.

[0981] 상기 실시형태 중 일부에서, 화합물은 구조식 (XA)를 갖고, 다른 실시형태에서, 화합물은 구조식 (XB)를 갖는다.

[0982] 상기 실시형태 중 일부에서, G^1은 -OH이고, 일부 실시형태에서 G^1은 $-NR^3R^4$이다. 예를 들어, 일부 실시형태에서, G^1은 $-NH_2$, $-NHCH_3$ 또는 $-N(CH_3)_2$이다. 특정 실시형태에서, G^1은 $-(C=O)NR^5$이다. 특정 실시형태에서, G^1은 $-NR^3(C=O)R^5$이다. 예를 들어, 일부 실시형태에서 G^1은 $-NH(C=O)CH_3$ 또는 $-NH(C=O)CH_2CH_2CH_3$이다.

[0983] 구조식 (X)의 상기 실시형태 중 일부에서, G^2는 $-CH_2-$이다. 일부 상이한 실시형태에서, G^2는 -(C=O)-이다.

[0984] 구조식 (X)의 상기 실시형태 중 일부에서, n은 2 내지 6 범위의 정수이고, 예를 들어, 일부 실시형태에서 n은 2, 3, 4, 5 또는 6이다. 일부 실시형태에서, n은 2이다. 일부 실시형태에서, n은 3이다. 일부 실시형태에서, n은 4이다.

[0985] 구조식 (X)의 상기 실시형태 중 특별한 것에서, R^1, R^2, R^3, R^4 및 R^5 중 적어도 하나는 비치환된다. 예를 들어, 일부 실시형태에서, R^1, R^2, R^3, R^4 및 R^5는 각각 비치환된다. 일부 실시형태에서, R^3은 치환된다. 다른 실시형태에서 R^4는 치환된다. 보다 추가의 실시형태에서, R^5는 치환된다. 특정 구체적인 실시형태에서, R^3 및 R^4 각각은 치환된다. 일부 실시형태에서, R^3, R^4 또는 R^5 상의 치환체는 하이드록실이다. 특정 실시형태에서, R^3 및 R^4는 각각 하이드록실로 치환된다.

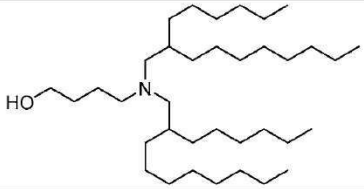
[0986] 구조식 (X)의 상기 실시형태 중 일부에서, 적어도 하나의 R은 OH이다. 다른 실시형태에서, 각각의 R은 H이다.

[0987] 구조식 (X)의 다양한 상이한 실시형태에서, 화합물은 하기 표 9에 제시된 구조 중 하나를 갖는다.

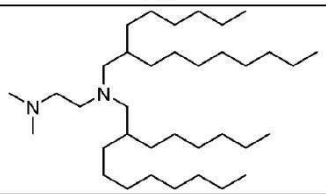
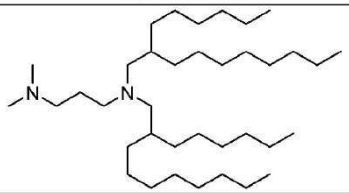
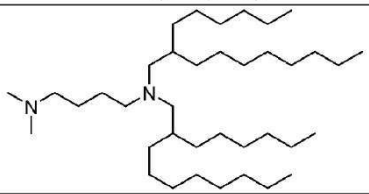
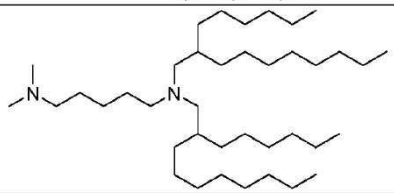
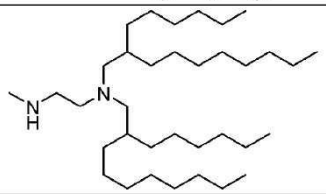
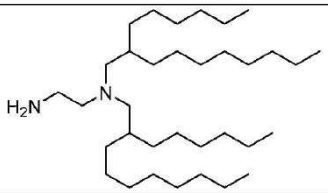
표 9

구조식 (X)의 대표적인 양이온성 지질

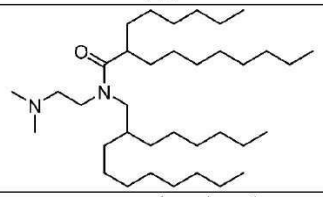
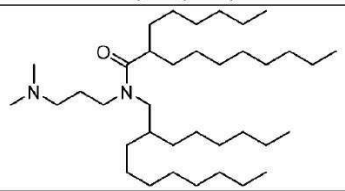
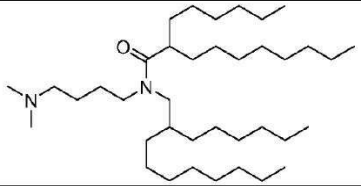
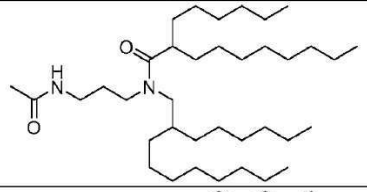
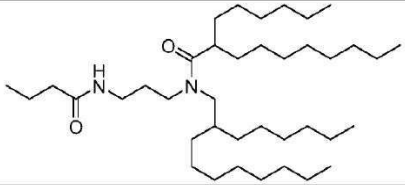
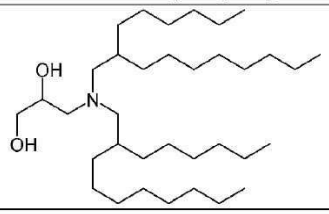
[0988]

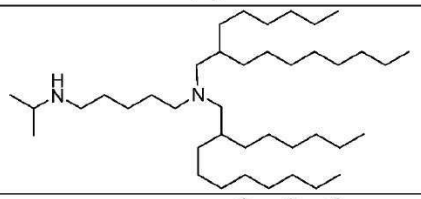
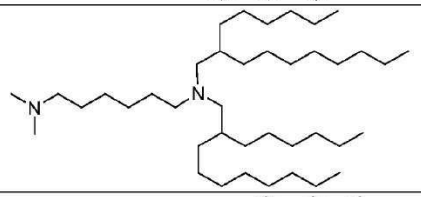
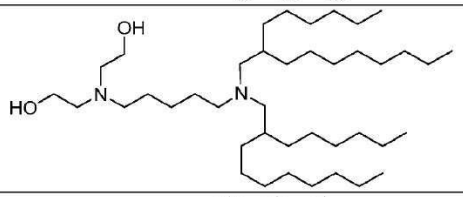
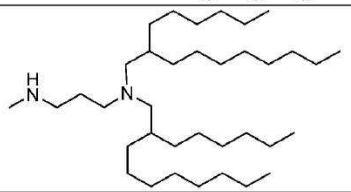
번호	구조
X-1	

[0989]

번호	구조
X-2	
X-3	
X-4	
X-5	
X-6	
X-7	

[0990]

번호	구조
X-8	
X-9	
X-10	
X-11	
X-12	
X-13	

번호	구조
X-14	
X-15	
X-16	
X-17	

[0991]

[0992] 실시형태 1, 2, 3, 4 또는 5 중 임의의 것에서, LNP는 중성 지질을 추가로 포함한다. 다양한 실시형태에서, 양이온성 지질 대 중성 지질의 몰비는 약 2:1 내지 약 8:1 범위이다. 특정 실시형태에서, 중성 지질은 5 내지 10 ㏖%, 5 내지 15㏖%, 7 내지 13㏖% 또는 9 내지 11㏖% 범위의 농도로 상기 LNP 중 임의의 것에 존재한다. 특정 구체적인 실시형태에서, 중성 지질은 약 9.5, 10 또는 10.5㏖%의 농도로 존재한다. 일부 실시형태에서, 양이온성 지질 대 중성 지질의 몰비는 약 4.1:1.0 내지 약 4.9:1.0, 약 4.5:1.0 내지 약 4.8:1.0 또는 약 4.7:1.0 내지 4.8:1.0 범위이다. 일부 실시형태에서, 총 양이온성 지질 대 중성 지질의 몰비는 약 4.1:1.0 내지 약 4.9:1.0, 약 4.5:1.0 내지 약 4.8:1.0 또는 약 4.7:1.0 내지 4.8:1.0 범위이다.

[0993] 실시형태 1, 2, 3, 4 또는 5 중 임의의 것에 사용하기 위한 예시적인 중성 지질은 예를 들어, 다이스테아로일포스파티딜콜린(DSPC), 다이올레오일포스파티딜콜린(DOPC), 다이팔미토일포스파티딜콜린(DPPC), 다이올레오일포스파티딜글리세롤(DOPG), 다이팔미토일포스파티딜글리세롤(DPPG), 다이올레오일포스파티딜에탄올아민(DOPE), 팔미토일올레오일포스파티딜콜린(POPC), 팔미토일올레오일-포스파티딜에탄올아민(POPE), 다이올레오일-포스파티딜에탄올아민 4-(N-말레이미도메틸)-사이클로헥산-1-카복실레이트(DOPE-mal), 다이팔미토일 포스파티딜 에탄올아민(DPPE), 다이미리스토일포스포에탄올아민(DMPE), 다이스테아로일-포스파티딜-에탄올아민(DSPE), 16-O-모노메틸 PE, 16-O-다이메틸 PE, 18-1-트랜스 PE, 1-스테아로일-2-올레오일-포스파티딜에탄올아민(SOPE) 및 1,2-다이엘라이도일-sn-글리세로-3-포스포에탄올아민(트랜스DOPE)을 포함한다. 일 실시형태에서, 중성 지질은 1,2-다이스테아로일-sn-글리세로-3포스포콜린(DSPC)이다. 일부 실시형태에서, 중성 지질은 DSPC, DPPC, DMPC, DOPC, POPC, DOPE 및 SM으로부터 선택된다. 일부 실시형태에서, 중성 지질은 DSPC이다.

[0994] 실시형태 1, 2, 3, 4 또는 5의 다양한 실시형태에서, 개시된 지질 나노입자 중 임의의 것은 스테로이드 또는 스테로이드 유사체를 포함한다. 특정 실시형태에서, 스테로이드 또는 스테로이드 유사체는 콜레스테롤이다. 일부 실시형태에서, 스테로이드는 39 내지 49몰%, 40 내지 46몰%, 40 내지 44몰%, 40 내지 42몰%, 42 내지 44몰% 또는 44 내지 46몰% 범위의 농도로 존재한다. 특정 구체적인 실시형태에서, 스테로이드는 40, 41, 42, 43, 44, 45 또는 46몰% 농도로 존재한다.

[0995] 특정 실시형태에서, 양이온성 지질 대 스테로이드의 몰비는 1.0:0.9 내지 1.0:1.2 또는 1.0:1.0 내지 1.0:1.2 범위이다. 이들 실시형태 중 일부에서, 양이온성 지질 대 콜레스테롤의 몰비는 약 5:1 내지 1:1 범위이다.

[0996] 특정 실시형태에서, 스테로이드는 스테로이드의 32 내지 40㏖% 범위의 농도로 존재한다.

[0997] 특정 실시형태에서, 총 양이온성 대 스테로이드의 몰비는 1.0:0.9 내지 1.0:1.2 또는 1.0:1.0 내지 1.0:1.2 범

위이다. 이들 실시형태 중 일부에서, 총 양이온성 지질 대 콜레스테롤의 몰비는 약 5:1 내지 1:1 범위이다.

[0998] 특정 실시형태에서, 스테로이드는 스테로이드의 32 내지 40㏖% 범위의 농도로 존재한다.

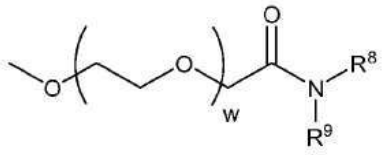
[0999] 실시형태 1, 2, 3 4 또는 5의 일부 실시형태에서, LNP는 중합체 접합된 지질을 추가로 포함한다. 실시형태 1, 2, 3 4 또는 5의 다양한 다른 실시형태에서, 중합체 접합된 지질은 페길화된 지질이다. 예를 들어, 일부 실시형태는 페길화된 다이아실글리세롤(PEG-DAG) 예컨대, 1-(모노메톡시-폴리에틸렌글리콜)-2,3-다이미리스토일글리세롤(PEG-DMG), 페길화된 포스파티딜에탄올로아민(PEG-PE), PEG 석시네이트 다이아실글리세롤(PEG- S-DAG), 예컨대, 4-O-(2',3'-다이(테트라데칸오일옥시)프로필-1-O-(ω-메톡시(폴리에톡시)에틸)부탄다이오에이트(PEG-S-DMG), 페길화된 세라마이드(PEG-cer) 또는 PEG 다이알콕시프로필카바메이트, 예컨대, 코-메톡시(폴리에톡시)에틸-N-(2,3-다이(테트라데칸옥시)프로필)카바메이트 또는 2,3-다이(테트라데칸옥시)프로필-N-(ω-메톡시(폴리에톡시)에틸)카바메이트를 포함한다

[1000] 다양한 실시형태에서, 중합체 접합된 지질은 1.0 내지 2.5몰% 범위의 농도로 존재한다. 특정 구체적인 실시형태에서, 중합체 접합된 지질은 약 1.7몰%의 농도로 존재한다. 일부 실시형태에서, 중합체 접합된 지질은 약 1.5몰% 농도로 존재한다.

[1001] 특정 실시형태에서, 양이온성 지질 대 중합체 접합된 지질의 몰비는 약 35:1 내지 약 25:1 범위이다. 일부 실시형태에서, 양이온성 지질 대 중합체 접합된 지질의 몰비는 약 100:1 내지 약 20:1 범위이다.

[1002] 특정 실시형태에서, 총 양이온성 지질(즉, 제1 양이온성 지질과 제2 양이온성 지질의 합) 대 중합체 접합된 지질의 몰비는 약 35:1 내지 약 25:1 범위이다. 일부 실시형태에서, 총 양이온성 지질 대 중합체 접합된 지질의 몰비는 약 100:1 내지 약 20:1 범위이다.

[1003] 실시형태 1, 2, 3 4 또는 5의 일부 실시형태에서, 페길화된 지질은 존재하는 경우 하기 화학식 (XI), 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체 또는 입체이성질체를 갖는다:



(XI)

[1004]

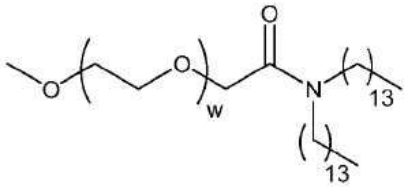
[1005] 식 중,

[1006] R^{12} 및 R^{13}은 각각 독립적으로 10 내지 30개 탄소 원자를 함유하는 선형 또는 분지형, 포화 또는 불포화 알킬 쇄이되, 알킬 쇄는 선택적으로 하나 이상의 에스터 결합에 의해서 개재되고;

[1007] w는 30 내지 60 범위의 평균 값을 갖는다.

[1008] 일부 실시형태에서, R^{12} 및 R^{13}은 각각 독립적으로 12 내지 16개 탄소 원자를 함유하는 선형, 포화 알킬 쇄이다. 다른 실시형태에서, 평균 w는 42 내지 55 범위이고, 예를 들어, 평균 w는 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54 또는 55이다. 일부 구체적인 실시형태에서, 평균 w는 약 49이다.

[1009] 일부 실시형태에서, 페길화된 지질은 하기 화학식을 갖는다:



(XIa)

[1010]

[1011] 식 중, 평균 w는 약 49이다.

[1012] 실시형태 1, 2, 3 4 또는 5의 일부 실시형태에서, 핵산은 안티센스 및 메신저 RNA로부터 선택된다. 예를 들어,

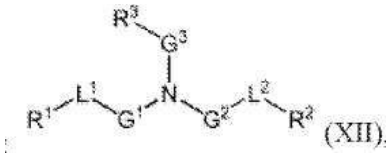
메신저 RNA는 예를 들어, 면역원성 단백질의 번역에 의해서 면역 반응(예를 들어, 백신으로서)을 유도하기 위해서 사용될 수 있다.

[1013] 실시형태 1, 2, 3 4 또는 5의 다른 실시형태에서, 핵산은 mRNA이고, LNP에서 mRNA 대 지질 비율(즉, N/P)은 N이 양이온성 지질의 몰을 나타내고, P가 핵산의 일부로서 존재하는 인산염의 몰을 나타낸다.

[1014] 실시형태에서, 전달 비히클은 미국 특허 공개 제20190314524호에 기재된 지질 또는 이온화 가능한 지질을 포함한다.

[1015] 본 발명의 일부 실시형태는 생체내에서 핵산의 증가된 활성 및 조성물의 개선된 내약성을 제공하는, 표 10에 열거된 구조로서 본 명세서에 기재된 신규한 양이온성 지질 중 하나 이상을 포함하는 핵산-지질 나노입자 조성물을 제공한다.

[1016] 일 실시형태에서, 이온화 가능한 지질은 하기 구조식 (XII), 또는 이의 약제학적으로 허용 가능한 염, 호변이성질체, 전구약물 또는 입체이성질체를 갖는다:

[1017] 
(XII)

[1018] 식 중,

[1019] L^1 또는 L^2 중 하나는 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_x-$, -S-S-, -C(=O)S-, SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=)NR^a-$, $-OC(=O)NR^a-$ 또는 $-NR^aC(=O)O-$이고, L^1 또는 L^2 중 나머지는 -O(C=O)-, -(C=O)O-, -C(=O)-, -O-, $-S(O)_x-$, -S-S-, -C(=O)S-, SC(=O)-, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 또는 $-NR^aC(=O)O-$ 또는 직접 결합이고;

[1020] G^1 및 G^2는 각각 독립적으로 비치환된 C_1-C_{12} 알킬렌 또는 C_1-C_{12} 알켄일렌이고;

[1021] G^3은 C_1-C_{24} 알킬렌, C_1-C_{24} 알켄일렌, C_3-C_8 사이클로알킬렌, C_3-C_8 사이클로알켄일렌이고;

[1022] R^a는 H 또는 C_1-C_{12} 알킬이고;

[1023] R^1 및 R^2는 각각 독립적으로 C_6-C_{24} 알킬 또는 C_6-C_{24} 알켄일이고;

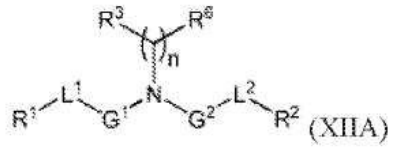
[1024] R^3은 H, OR^5, CN, $-C(=O)OR^4$, $-OC(=O)R^4$ 또는 $-NR^5C(=O)R^4$이고;

[1025] R^4는 C_1-C_{12} 알킬이고;

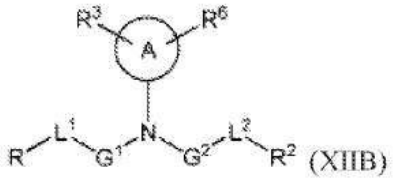
[1026] R^5는 H 또는 C_1-C_6 알킬이고;

[1027] x는 0, 1 또는 2이다.

[1028] 일부 실시형태에서, 이온화 가능한 지질은 하기 구조식 (XIIA) 또는 (XIIB) 중 하나를 갖는다:



(XIIA)



(XIIB)

[1029]

[1030] 식 중,

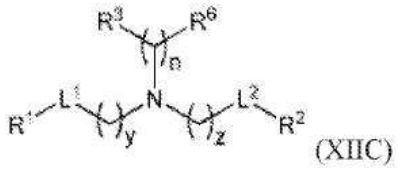
[1031] A는 3 내지 8-원 사이클로알킬 또는 사이클로알킬렌 고리이고;

[1032] R^6은, 각각의 경우에, 독립적으로 H, OH 또는 C_1-C_{24} 알킬이고;

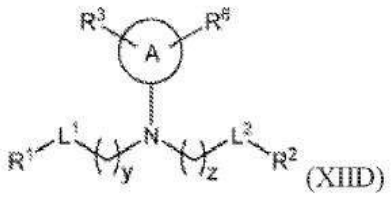
[1033] n은 1 내지 15 범위의 정수이다.

[1034] 일부 실시형태에서, 이온화 가능한 지질은 구조식 (XIIA)를 갖고, 다른 실시형태에서, 이온화 가능한 지질은 구조식 (XIIB)를 갖는다.

[1035] 다른 실시형태에서, 이온화 가능한 지질은 하기 구조식 (XIIC) 또는 (XIID) 중 하나를 갖는다:



(XIIC)



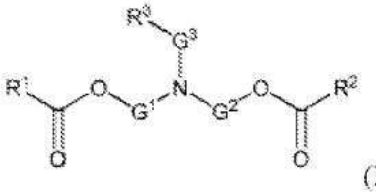
(XIID)

[1036]

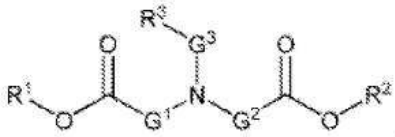
[1037] 식 중, y 및 z는 각각 독립적으로 1 내지 12 범위의 정수이다.

[1038] 일부 실시형태에서, L^1 또는 L^2 중 하나는 -O(C=O)-이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 -O(C=O)-이다. 상기 중 임의의 것의 일부 상이한 실시형태에서, L^1 및 L^2는 각각 독립적으로 -(C=O)O- 또는 -O(C=O)-이다. 예를 들어, 일부 실시형태에서 L^1 및 L^2 각각은 -(C=O)O-이다.

[1039] 일부 실시형태에서, 이온화 가능한 지질은 하기 구조식 (XIIE) 또는 (XIIF) 중 하나를 갖는다:



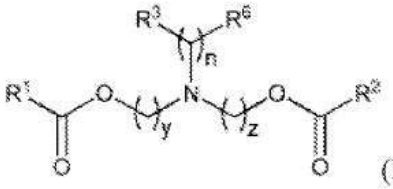
(XIIE)



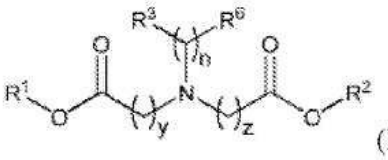
(XIIF)

[1040]

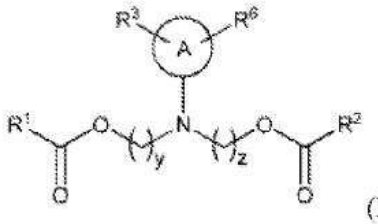
[1041] 일부 실시형태에서, 이온화 가능한 지질은 하기 구조식 (XIIG), (XIIH), (XIII) 또는 (XIIJ 중 하나를 갖는다:



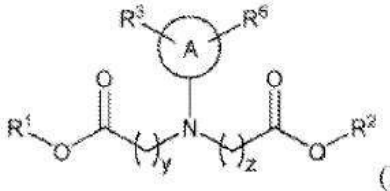
(XIIG)



(XIIH)



(XIII)



(XIIJ)

[1042]

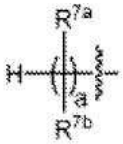
[1043] 일부 실시형태에서, n은 2 내지 12, 예를 들어, 2 내지 8 또는 2 내지 4 범위의 정수이다. 예를 들어, 일부 실시형태에서, n은 3, 4, 5 또는 6이다. 일부 실시형태에서, n은 3이다. 일부 실시형태에서, n은 4이다. 일부 실시형태에서, n은 5이다. 일부 실시형태에서, n은 6이다.

[1044] 일부 실시형태에서, y 및 z는 각각 독립적으로 2 내지 10 범위의 정수이다. 예를 들어, 일부 실시형태에서, y 및 z는 각각 독립적으로 4 내지 9 또는 4 내지 6 범위의 정수이다.

[1045] 일부 실시형태에서, R^6은 H이다. 다른 실시형태에서, R^6은 C_1-C_{24} 알킬이다. 다른 실시형태에서, R^6은 OH이다.

[1046] 일부 실시형태에서, G^3은 치환되지 않는다. 다른 실시형태에서, G3은 치환된다. 다양한 상이한 실시형태에서, G^3은 선형 C_1-C_{24} 알킬렌 또는 선형 C_1-C_{24} 알켄일렌이다.

[1047] 일부 실시형태에서, R^1 또는 R^2, 또는 둘 다는 C_6-C_{24} 알켄일이다. 예를 들어, 일부 실시형태에서, R^1 및 R^2는 각각 독립적으로 하기 구조를 갖는다:

[1048] 

[1049] 식 중,

[1050] R^{7a} 및 R^{7b}는 각각의 경우에, 독립적으로 H 또는 C_1-C_{12} 알킬이고;

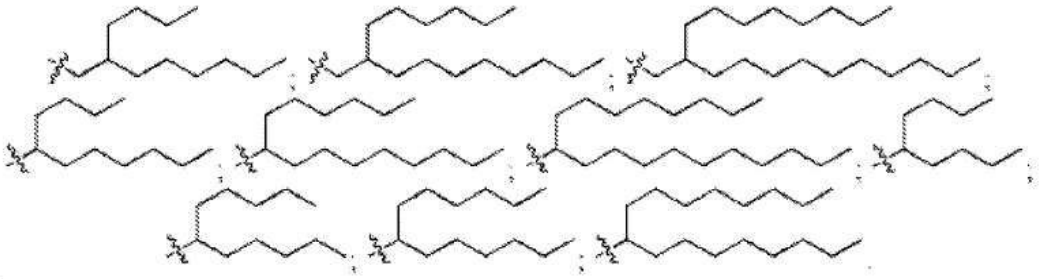
[1051] a는 2 내지 12이고,

[1052] R^{7a}, R^{7b} 및 a는 R^1 및 R^2가 각각 독립적으로 6 내지 20개의 탄소 원자를 포함하도록 각각 선택된다.

[1053] 일부 실시형태에서 a는 5 내지 9 또는 8 내지 12 범위의 정수이다.

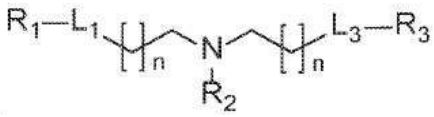
[1054] 일부 실시형태에서, R^{7a}의 적어도 하나의 경우는 H이다. 예를 들어, 일부 실시형태에서, R^{7a}는 각각의 경우에 H이다. 다른 상이한 실시형태에서, R^{7b} 중 적어도 하나의 경우는 C_1-C_8 알킬이다. 예를 들어, 일부 실시형태에서, C_1-C_8 알킬은 메틸, 에틸, n-프로필, 아이소프로필, n-부틸, 아이소-부틸, tert-부틸, n-헥실 또는 n-옥틸이다.

[1055] 상이한 실시형태에서, R^1 또는 R^2, 또는 둘 다는 하기 구조 중 하나를 갖는다:

[1056] 

[1057] 일부 실시형태에서, R^3은 -OH, -CN, -C(=O)OR^4, -OC(=O)R^4 또는 -NHC(=O)R^4이다. 일부 실시형태에서, R^4는 메틸 또는 에틸이다.

[1058] 일부 실시형태에서, 이온화 가능한 지질은 화학식 (1)의 화합물이다:

[1059] 

[1060] 화학식 (1)

[1061] 식 중,

[1062] 각각의 n은 독립적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 또는 15이고;

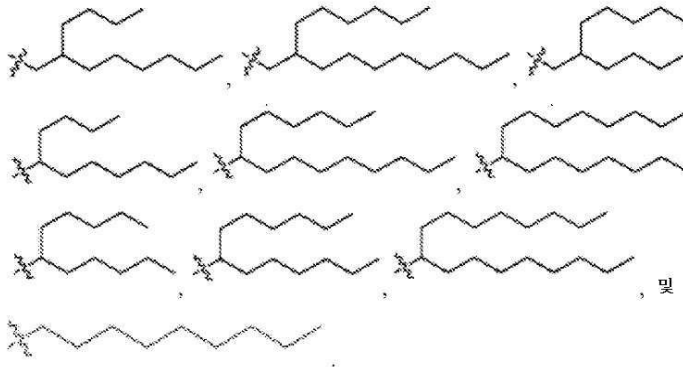
[1063] L_1 및 L_3은 각각 독립적으로 -OC(O)-* 또는 -C(O)O-*이되, "*"는 R_1 또는 R_3에 대한 부착점을 나타내고;

[1064] R_1 및 R_3은 각각 독립적으로 옥소, 할로, 하이드록시, 사이아노, 알킬, 알켄일, 알데하이드, 헤테로사이클릴알킬, 하이드록시알킬, 다이하이드록시알킬, 하이드록시알킬아미노알킬, 아미노알킬, 알킬아미노알킬, 다이알킬아미노알킬, (헤테로사이클릴)(알킬)아미노알킬, 헤테로사이클릴, 헤테로아릴, 알킬헤테로아릴, 알킨일, 알콕시, 아미노, 다이알킬아미노, 아미노알킬카보닐아미노, 아미노카보닐알킬아미노, (아미노카보닐알킬)(알킬)아미노, 알켄일카보닐아미노, 하이드록시카보닐, 알킬옥시카보닐, 아미노카보닐, 아미노알킬아미노카보닐, 알킬아미노알킬아미노카보닐, 다이알킬아미노알킬아미노카보닐, 헤테로사이클릴알킬아미노카보닐, (알킬아미노알킬)(알킬)아미노카보닐, 알킬아미노알킬카보닐, 다이알킬아미노알킬카보닐, 헤테로사이클릴카보닐, 알켄일카보닐, 알킨일카보닐, 알킬설폭사이드, 알킬설폭사이드알킬, 알킬설포닐 및 알킬설폰알킬로부터 선택된 하

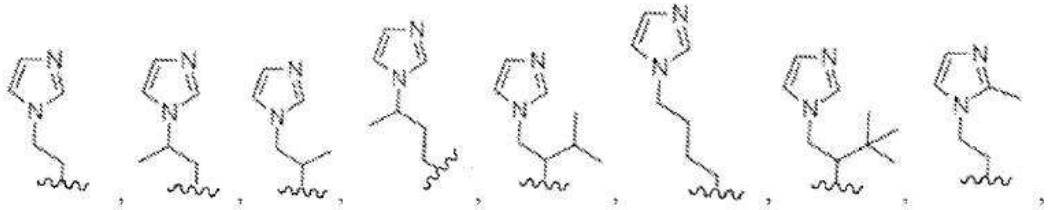
나 이상의 치환체에 의해서 선택적으로 치환된 선형 또는 분지형 C_9-C_{20} 알킬 또는 C_9-C_{20} 알켄일이다.

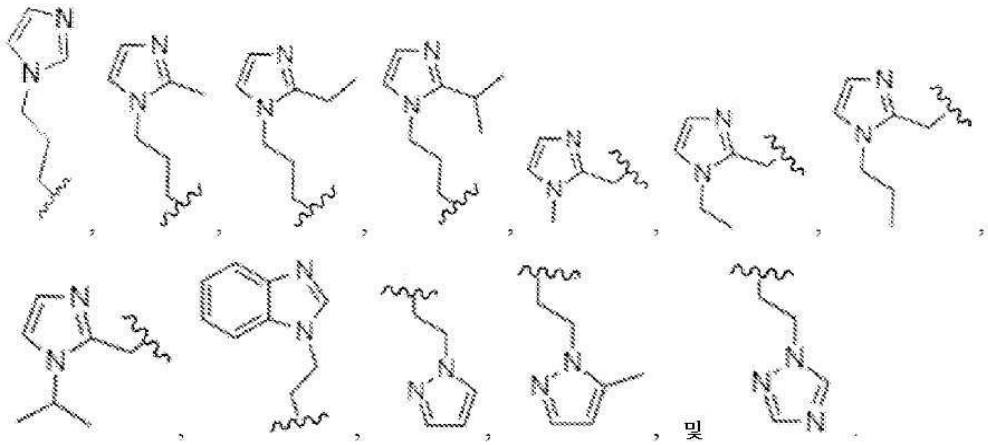
[1065] 일부 실시형태에서, R_1과 R_3은 동일하다. 일부 실시형태에서, R_1과 R_3은 상이하다.

[1066] 일부 실시형태에서, R_1 및 R_3은 각각 독립적으로 분지형 포화 C_9-C_{20} 알킬이다. 일부 실시형태에서, R_1 및 R_3 중 하나는 분지형 포화 C_9-C_{20} 알킬이고, 나머지는 부분지형 포화 C_9-C_{20} 알킬이다. 일부 실시형태에서, R_1 및 R_3은 각각 독립적으로 하기로 이루어진 군으로부터 선택된다:

[1067] 

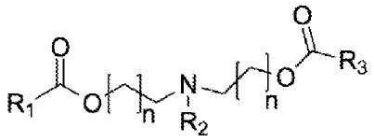
[1068] 다양한 실시형태에서, R_2는 하기로 이루어진 군으로부터 선택된다:

[1069] 

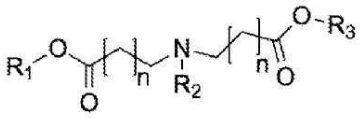
[1070] 

[1071] 일부 실시형태에서, R_2는 전문이 본 명세서에 참조에 의해 포함된 국제 특허 공개 제WO2019/152848 A1호에 기재된 바와 같을 수 있다.

[1072] 일부 실시형태에서, 이온화 가능한 지질은 하기 화학식 (1-1) 또는 화학식 (1-2)의 화합물이다:



화학식 (I-1)



화학식 (I-2)

[1073]

[1074] 식 중, n, R_1, R_2 및 R_3은 화학식 (1)에 정의된 바와 같다.

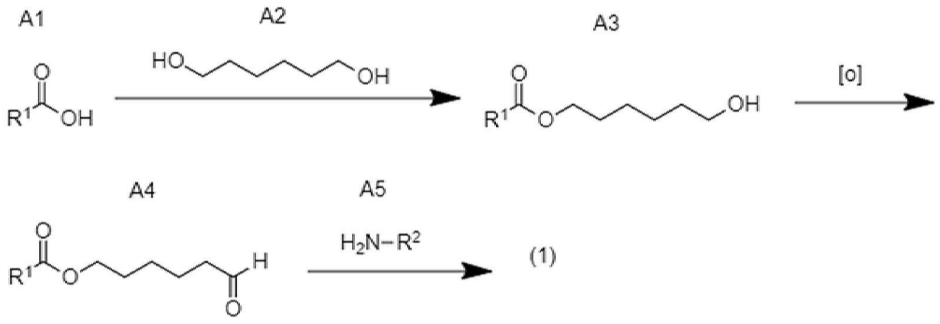
[1075] 상기 화합물 및 조성물에 대한 제조 방법은 본 명세서 하기에 기재되어 있고/있거나 당업계에 공지되어 있다.

[1076] 본 명세서에 기재된 방법에서 중간체 화합물의 작용기는 적절한 보호기에 의해 보호될 필요가 있을 수 있다는 것이 당업자에 의해 이해될 것이다. 이러한 작용기는 예를 들어, 하이드록실, 아미노, 머캅토 및 카복실산을 포함한다. 하이드록실에 적합한 보호기는 예를 들어, 트라이알킬실릴 또는 다이아릴알킬실릴 (예를 들어, t-부틸다이메틸실릴, t-부틸다이페닐실릴 또는 트라이메틸실릴), 테트라하이드로피란일, 벤질 등을 포함한다 아미노, 아미딘오 및 구아니딘오에 적합한 보호기는, 예를 들어, t-부톡시카보닐, 벤질옥시카보닐 등을 포함한다. 머캅토에 적합한 보호기는 예를 들어, -C(O)-R"(식 중, R"는 알킬, 아릴 또는 아릴알킬임), p-메톡시벤질, 트리틸 등을 포함한다. 카복실산에 적합한 보호기는 예를 들어, 알킬, 아릴 또는 아릴알킬 에스터를 포함한다. 보호기는 당업자에게 공지되어 있고, 본 명세서에 기재된 바와 같은 표준 기술에 따라 첨가 또는 제거될 수 있다. 보호기의 사용은 예를 들어, 문헌[Green, T. W. and P. G. M. Wutz, Protective Groups in Organic Synthesis (1999), 3rd Ed., Wiley]에 상세하게 기재되어 있다. 당업자가 이해하는 바와 같이, 보호기는 또한 중합체 수지, 예컨대, 왕 수지(Wang resin), 링크 수지(Rink resin,) 또는 2-클로로트리틸-클로라이드 수지일 수 있다.

[1077] 본 발명의 화합물의 보호된 유도체가 그 자체로 약리학적 활성을 갖지 않을 수 있지만, 이들은 포유동물에게 투여된 후 체내에서 대사되어 약리학적으로 활성인 본 발명의 화합물을 형성할 수 있다는 것이 또한 당업자에게 인식될 것이다. 따라서 이러한 유도체는 전구약물로 설명될 수 있다. 본 발명의 화합물의 모든 전구약물은 본 발명의 범위 내에 포함된다.

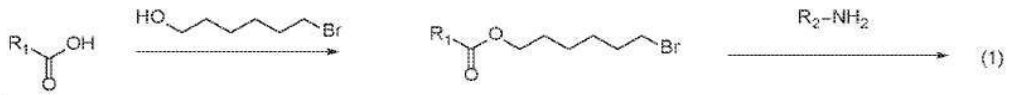
[1078] 또한, 유리 염기 또는 산 형태로 존재하는 본 발명의 모든 화합물은 당업자에게 공지된 방법에 의해 적절한 무기 또는 유기 염기 또는 산으로 처리함으로써 이의 약제학적으로 허용 가능한 염으로 전환될 수 있다. 본 발명의 화합물의 염은 또한 표준 기술에 의해 이의 유리 염기 또는 산 형태로 전환될 수 있다.

[1079] 하기 반응식은 화학식 (1)의 화합물을 제조하는 예시적인 방법을 예시한다:



[1080] A1은 당업계에 공지된 방법에 따라 구입하거나 제조한다. 적절한 축합 조건(예를 들어, DCC) 하에서 A1과 다이올 A2의 반응은 에스터/알코올 A3를 생성하고, 이어서 이것을 알데하이드 A4로 산화(예를 들어, PCC 사용)시킬 수 있다. 환원성 아민화 조건 하에서 A4와 아민 A5의 반응은 화학식 (1)의 화합물을 생성한다.

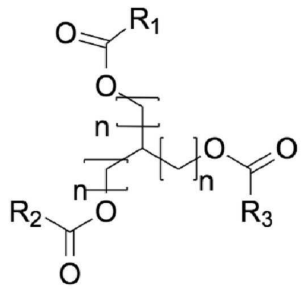
[1081] 하기 반응식은 화학식 1의 화합물을 제조하는 두 번째 예시적인 방법을 예시하며, 여기서 R_1과 R_3은 동일하다:

[1082] 

[1083] 보호기를 사용하는 것과 같은 상기 반응식에 대한 변형은 R_1과 R_3이 상이한 화합물을 생성할 수 있다. 상기 반응식에 대한 보호기의 사용뿐만 아니라 다른 변형 방법은 당업자에게 용이하게 명백할 것이다.

[1084] 당업자는 유사한 방법에 의해 또는 당업자에게 공지된 다른 방법을 조합함으로써 이들 화합물을 제조할 수 있다는 것을 이해한다. 또한, 당업자는 적절한 출발 물질을 사용하고, 합성 매개변수를 변형시킴으로써 본 명세서에 구체적으로 예시되지 않은 화학식 (1)의 다른 화합물을 제조할 수 있다는 것을 이해한다. 일반적으로, 출발 물질은 시그마 알드리치사(Sigma Aldrich), 랑카스터 신쎄시스사(Lancaster Synthesis, Inc.), 메이브리지사(Maybridge), 매트릭스 사이언티픽사(Matrix Scientific), TCI, 및 플루오로켐사(Fluorochem USA) 등과 같은 공급원으로부터 얻거나 당업자에게 공지된 출처에 따라 합성되거나(예를 들어, 문헌[Reactions, Mechanisms, and Structure, 5th edition (Wiley, December 2000)]), 본 명세서에 기재된 바와 같이 제조될 수 있다.

[1085] 일부 실시형태에서, 이온화 가능한 지질은 화학식 (2)의 화합물이다:

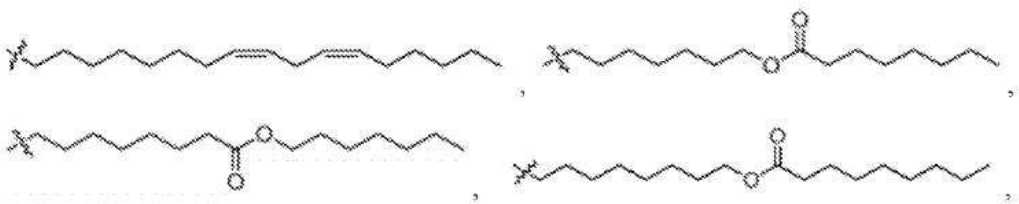
[1086] 

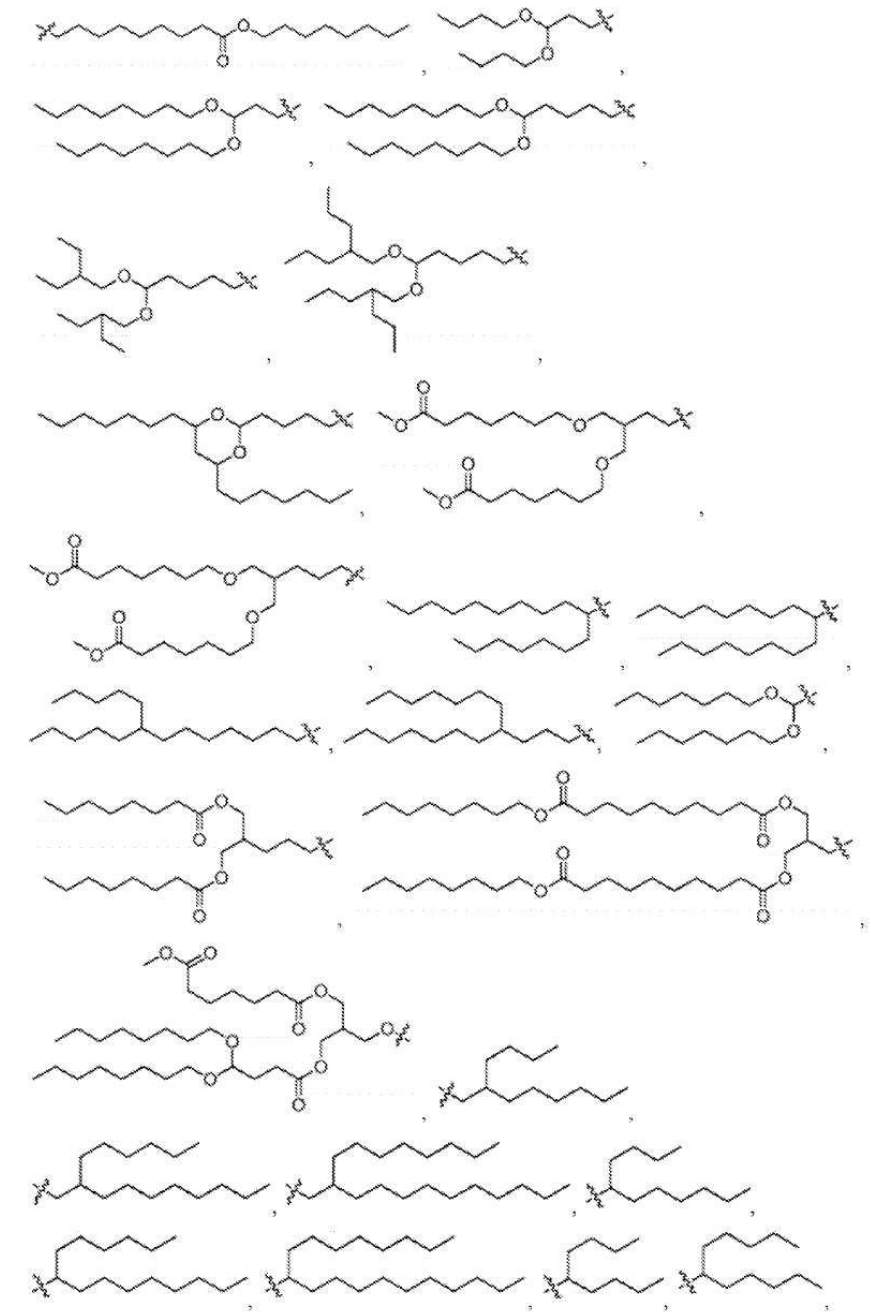
[1087] 화학식 (2),

[1088] 각각의 n은 독립적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 또는 15이다.

[1089] 일부 실시형태에서, 화학식 (2)에서 사용된 바와 같이, R_1 및 R_2는 화학식 (1)에 정의된 바와 같다.

[1090] 일부 실시형태에서, 화학식 (2)에서 사용된 바와 같이, R_1 및 R_2는 각각 독립적으로 하기로 이루어진 군으로부터 선택된다:

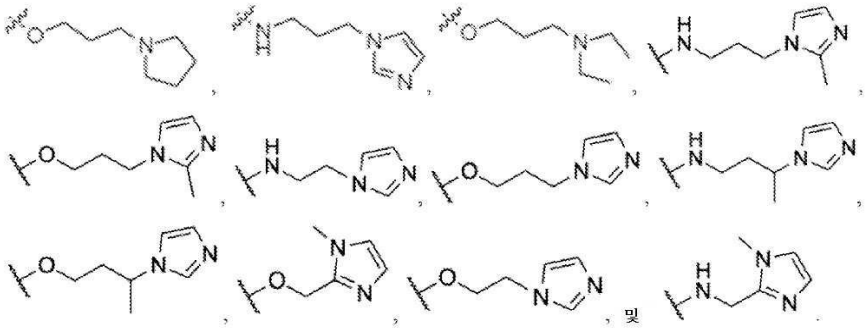
[1091] 

[1092] 

[1093] 및

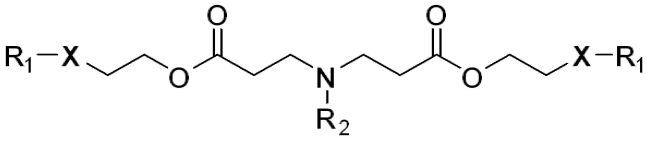
[1094] 일부 실시형태에서, 화학식 (2)에서 사용된 바와 같은 R_1 및/또는 R_2는 전문이 본 명세서에 참조에 의해 포함된 국제 특허 공개 제WO2015/095340 A1호에 기재된 바와 같을 수 있다. 일부 실시형태에서, 화학식 (2)에서 사용된 바와 같은 R_1은 전문이 본 명세서에 참조에 의해 포함된 국제 특허 공개 제WO2019/152557 A1호에 기재된 바와 같을 수 있다.

[1095] 일부 실시형태에서, 화학식 (2)에서 사용된 바와 같이, R_3은 하기로 이루어진 군으로부터 선택된다:



[1096]

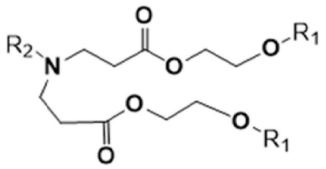
[1097] 일부 실시형태에서, 이온화 가능한 지질은 화학식 (3)의 화합물이다:



[1098]

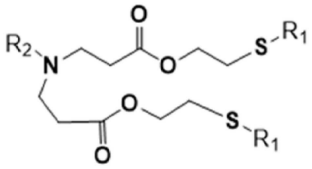
[1099] 식 중, X는 -O-, -S- 또는 -OC(O)-*로부터 선택되고, 식 중, *는 R_1에 대한 부착점을 나타낸다.

[1100] 일부 실시형태에서, 이온화 가능한 지질은 화학식 (3-1)의 화합물이다:



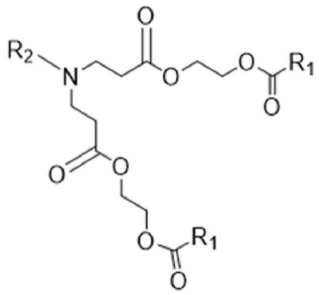
[1101] (3-1).

[1102] 일부 실시형태에서, 이온화 가능한 지질은 화학식 (3-2)의 화합물이다:



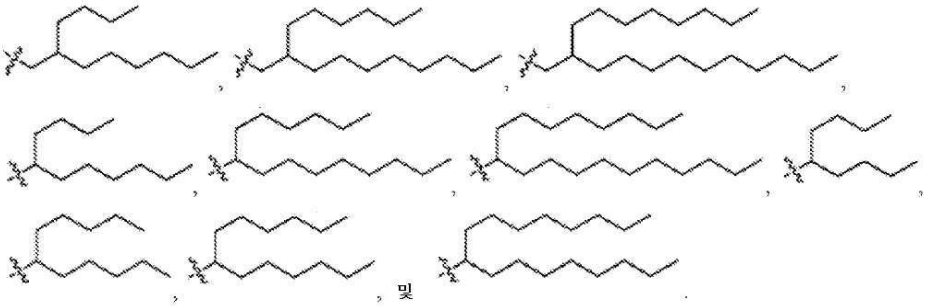
[1103] (3-2).

[1104] 일부 실시형태에서, 이온화 가능한 지질은 화학식 (3-3)의 화합물이다:



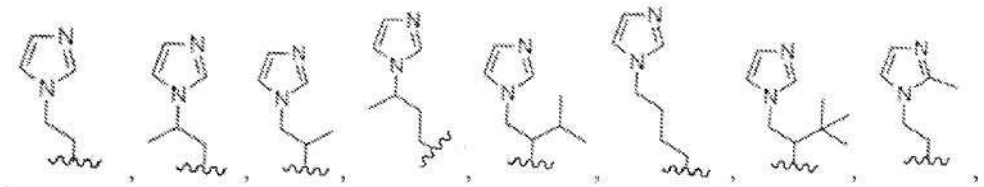
[1105] (3-3).

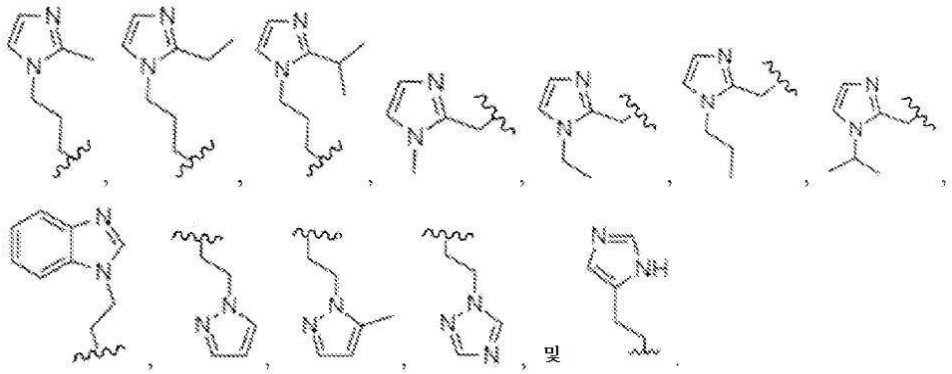
[1106] 일부 실시형태에서, 화학식 (3-1), (3-2) 또는 (3-3)에서 사용된 바와 같이, 각각의 R_1은 독립적으로 분지형 포화 C_9-C_{20} 알킬이다. 일부 실시형태에서, 각각의 R_1은 하기로 이루어진 군으로부터 독립적으로 선택된다:

[1107] 

[1108] 일부 실시형태에서, 화학식 (3-1), (3-2) 또는 (3-3)에서 각각의 R_1은 동일하다.

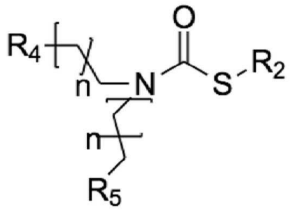
[1109] 일부 실시형태에서, 화학식 (3-1), (3-2) 또는 (3-3)에서 사용된 바와 같이, R_2는 하기로 이루어진 군으로부터 선택된다:

[1110] 

[1111] 

[1112] 일부 실시형태에서, 화학식 (3-1), (3-2) 또는 (3-3)에서 사용된 바와 같은 R_2는 전문이 본 명세서에 참조에 의해 포함된 국제 특허 공개 제WO2019/152848A1호에 기재된 바와 같을 수 있다.

[1113] 일부 실시형태에서, 이온화 가능한 지질은 화학식 (5)의 화합물이다:



[1114] (5),

[1115] 식 중,

[1116] 각각의 n은 독립적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 또는 15이고;

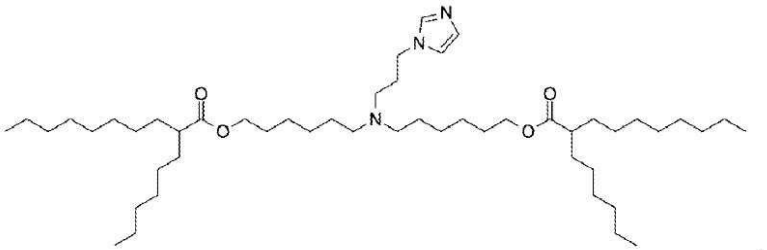
[1117] R_2는 화학식 (1)에 정의된 바와 같다.

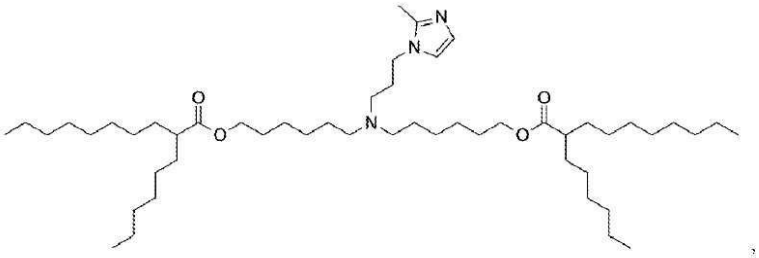
[1118] 일부 실시형태에서, 화학식 (5)에서 사용된 바와 같이, R_4 및 R_5는 각각 화학식 (1)에서 R_1 및 R_3으로서 정의된

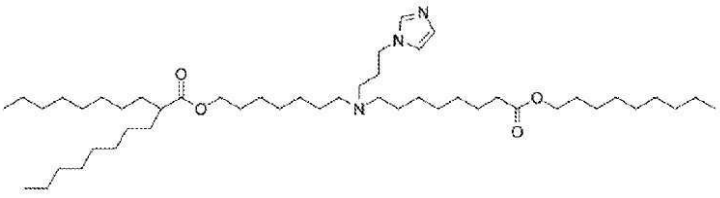
다. 일부 실시형태에서, 화학식 (5)에서 사용된 바와 같이, R_4 및 R_5는 전문이 본 명세서에 참조에 의해 포함된 국제 특허 공개 제WO2019/191780 A1호에 기재된 바와 같을 수 있다.

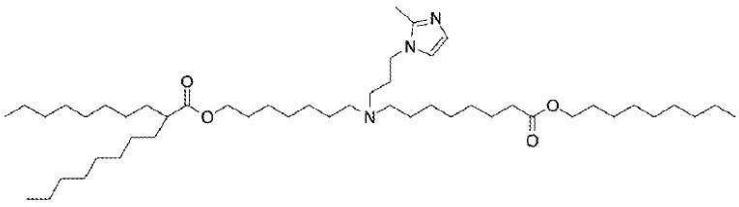
[1119] 일부 실시형태에서, 본 개시내용의 이온화 가능한 지질은 표 10a로부터 선택된다. 일부 실시형태에서, 이온화 가능한 지질은 표 10a의 지질 26이다. 일부 실시형태에서, 이온화 가능한 지질은 표 10a의 지질 27이다. 일부 실시형태에서, 이온화 가능한 지질은 표 10a의 지질 53이다. 일부 실시형태에서, 이온화 가능한 지질은 표 10a의 지질 54이다.

[1120] 일부 실시형태에서, 본 개시내용의 이온화 가능한 지질은 하기로 이루어진 군으로부터 선택된다:


,

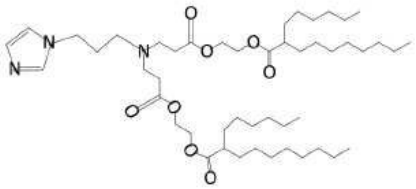
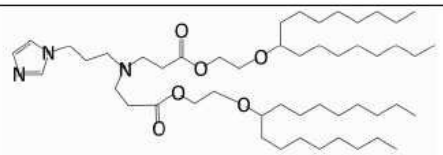
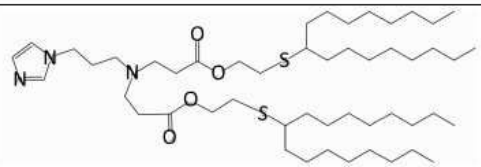
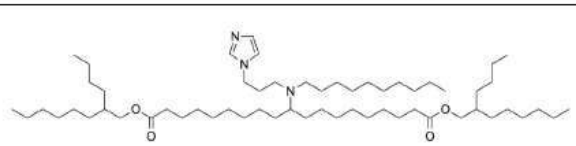
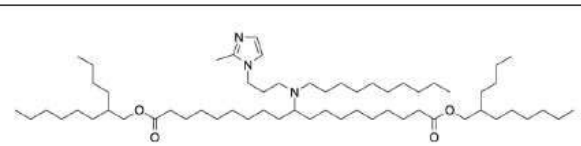

,


, 및


.

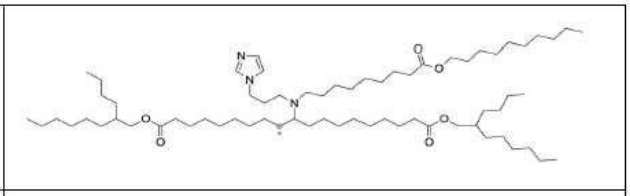
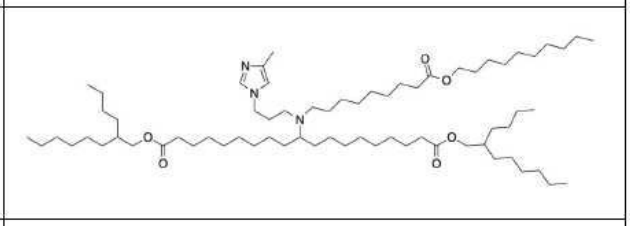
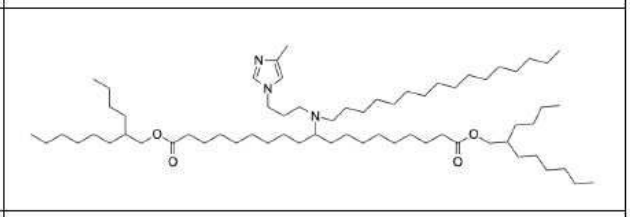
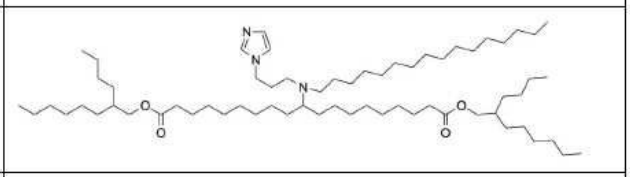
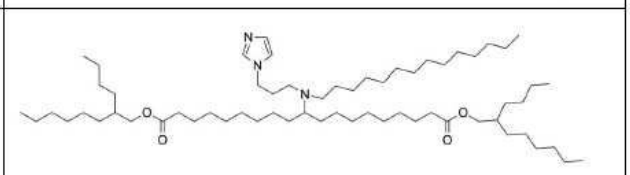
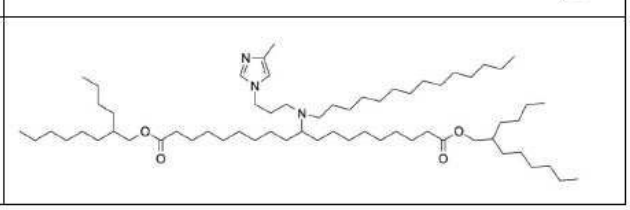
[1121]

[1122] [표 10a]

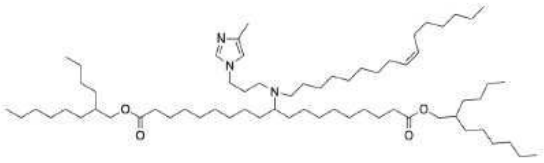
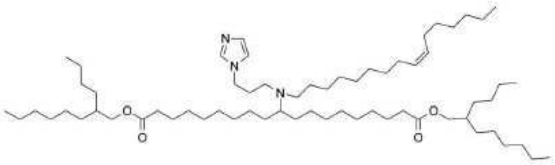
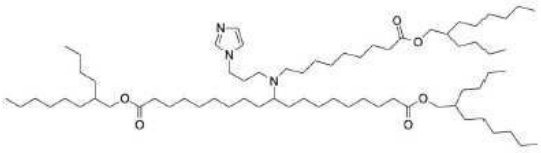
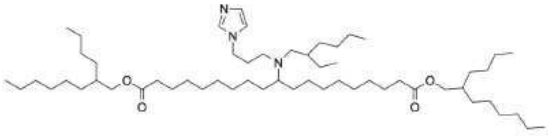
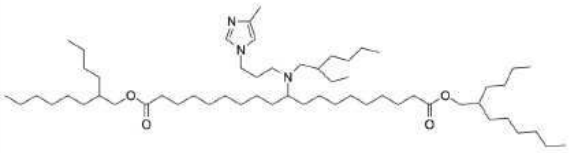
이온화 가능한 지질 번호	구조
1	
2	
3	
4	
5	

[1123]

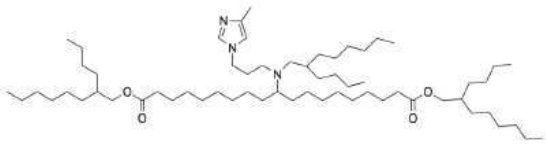
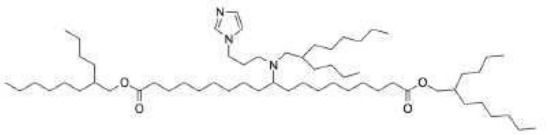
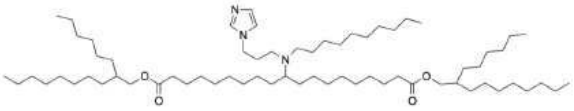
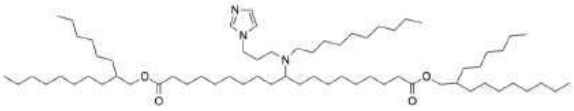
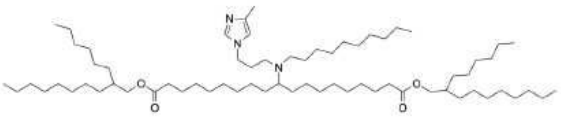
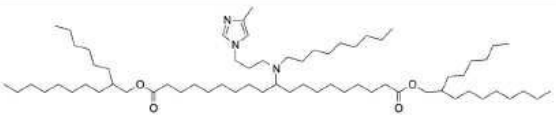
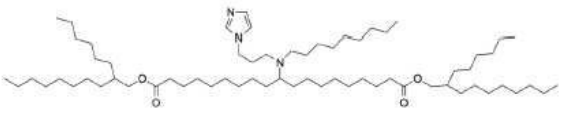
[1124]

6	
7	
8	
9	
10	
11	

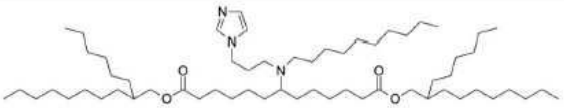
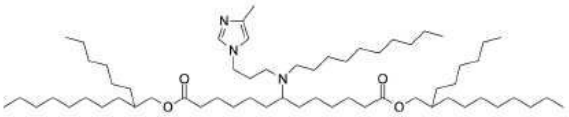
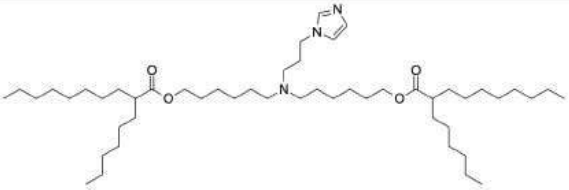
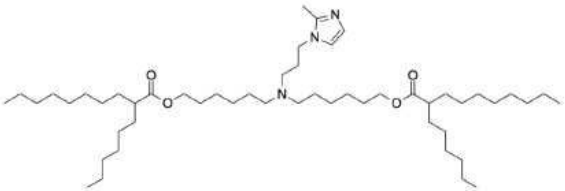
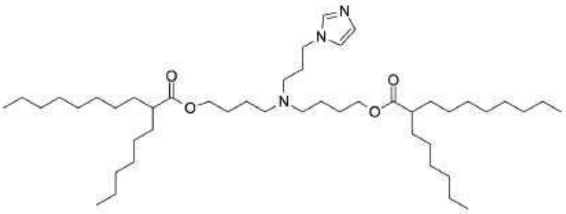
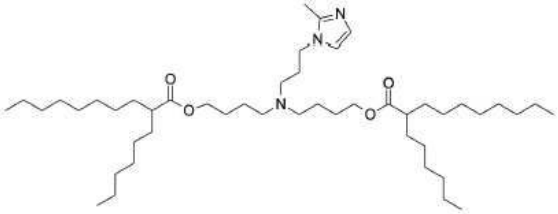
[1125]

12	
13	
14	
15	
16	

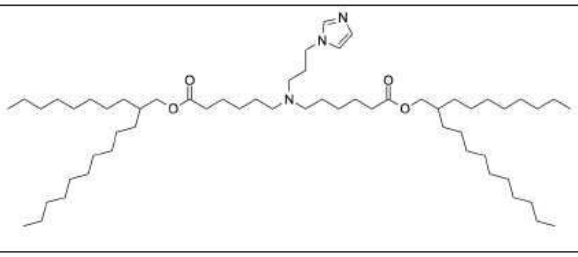
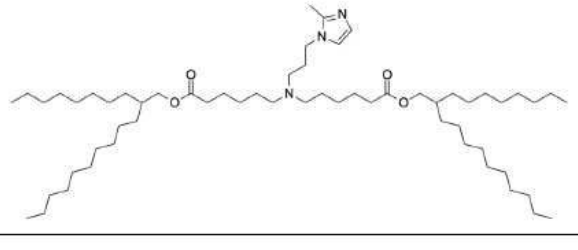
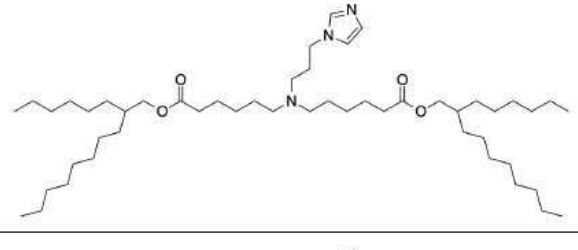
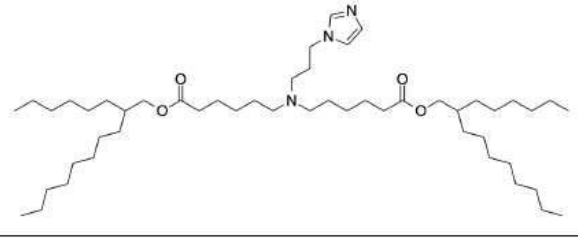
[1126]

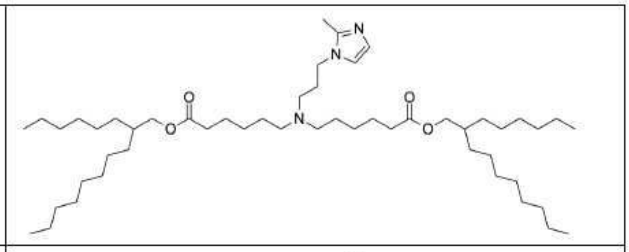
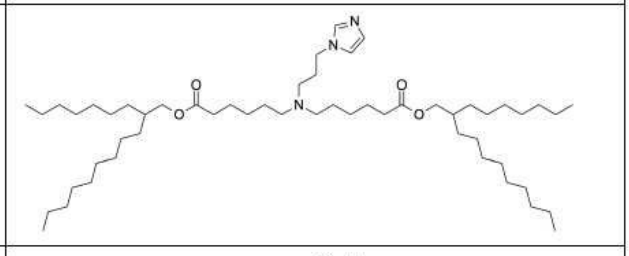
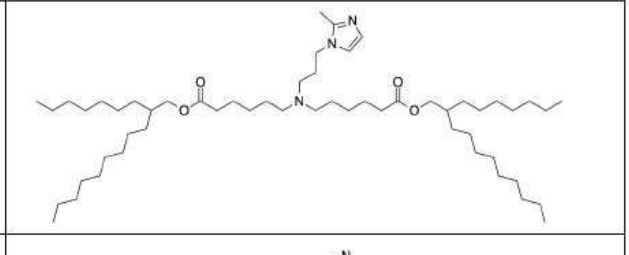
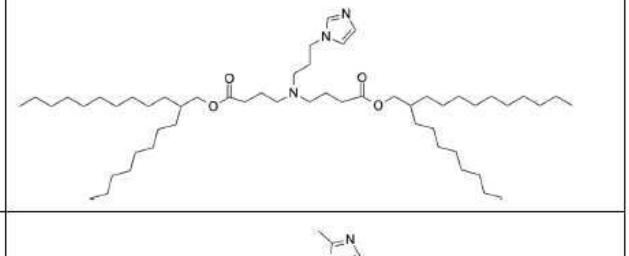
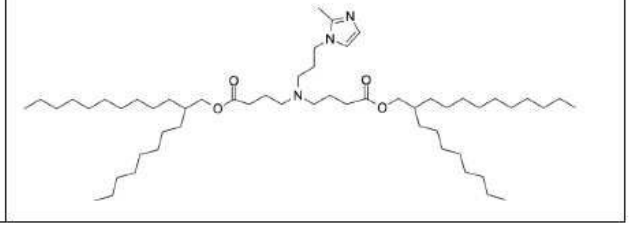
17	
18	
19	
20	
21	
22	
23	

[1127]

24	
25	
26	
27	
28	
29	

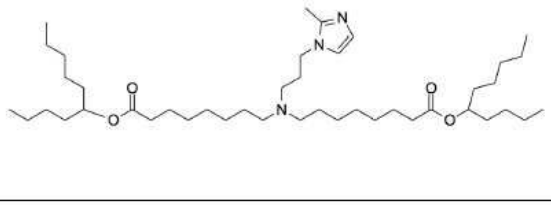
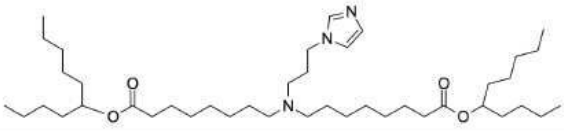
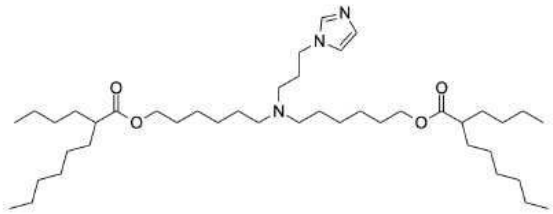
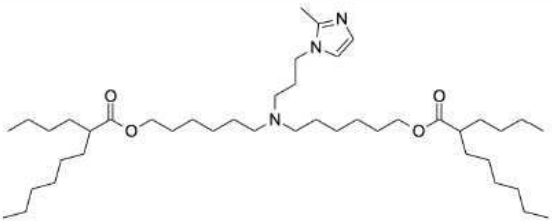
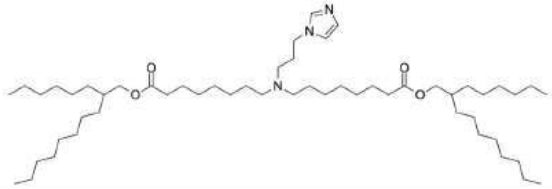
[1128]

30	
31	
32	
33	

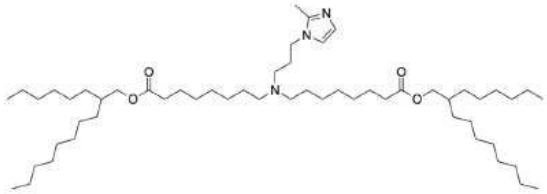
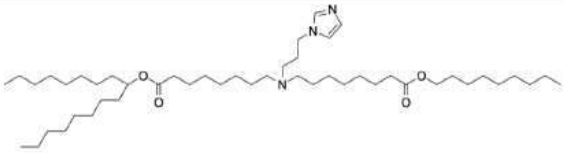
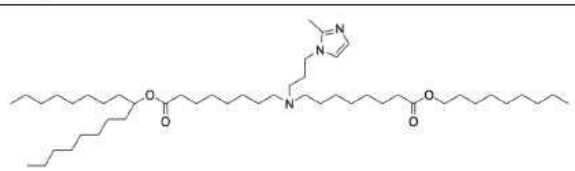
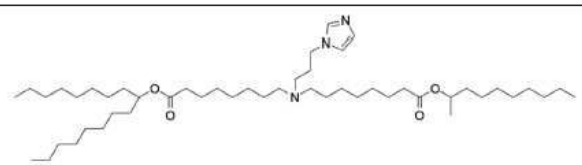
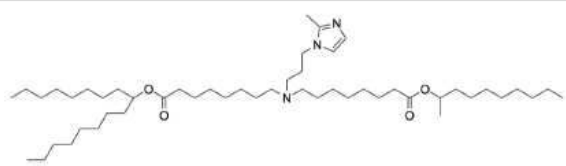
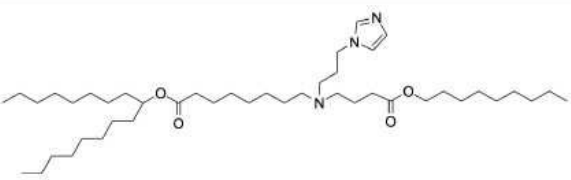
34	
35	
36	
37	
38	

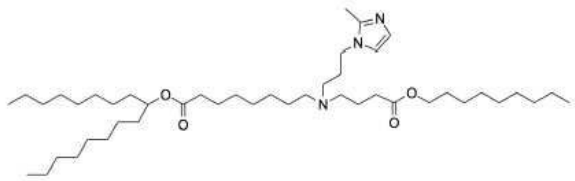
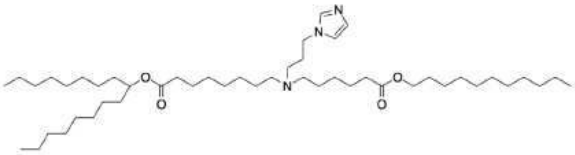
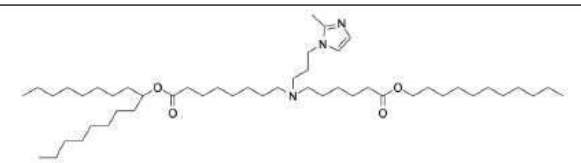
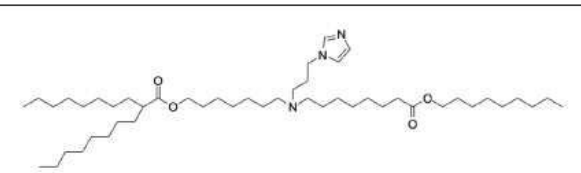
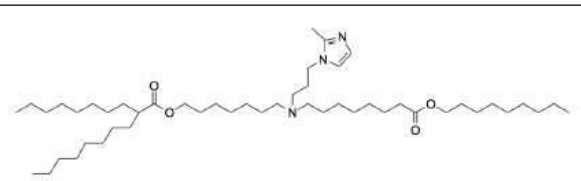
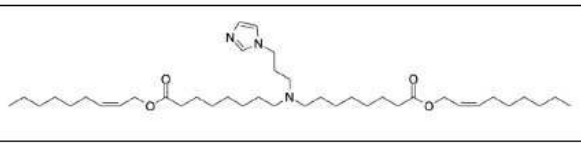
[1129]

[1130]

39	
40	
41	
42	
43	

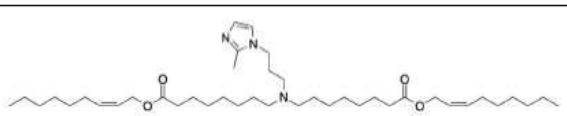
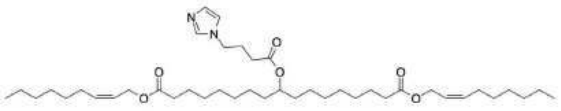
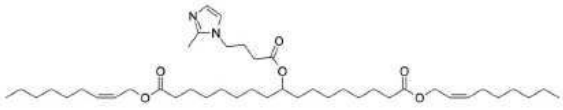
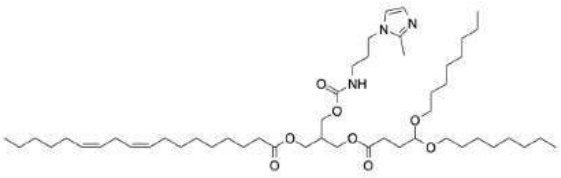
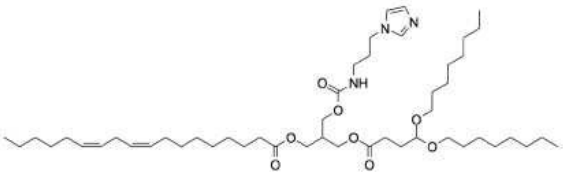
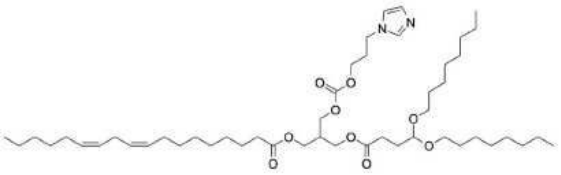
[1131]

44	
45	
46	
47	
48	
49	

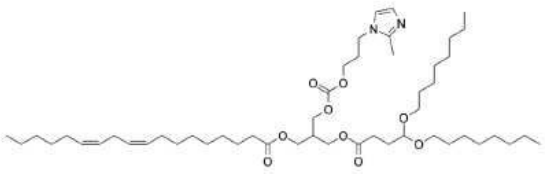
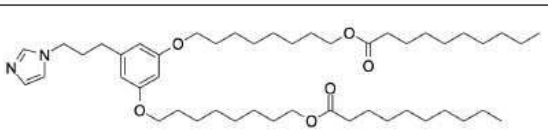
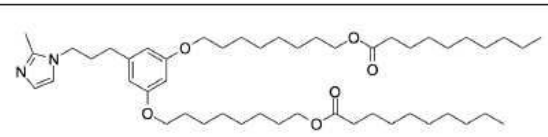
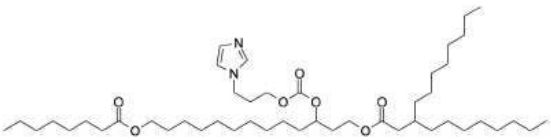
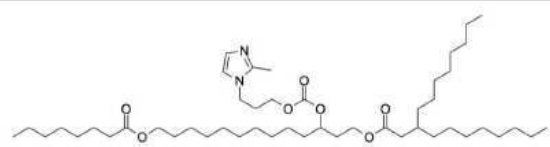
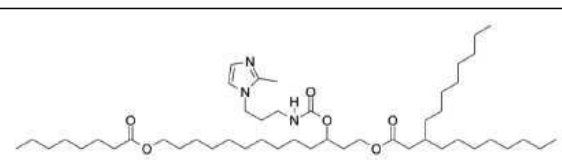
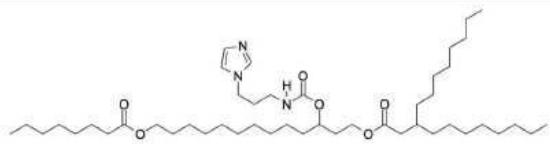
50	
51	
52	
53	
54	
55	

[1132]

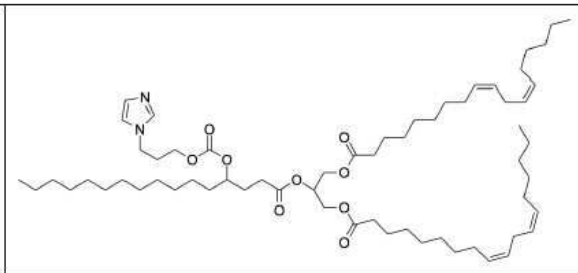
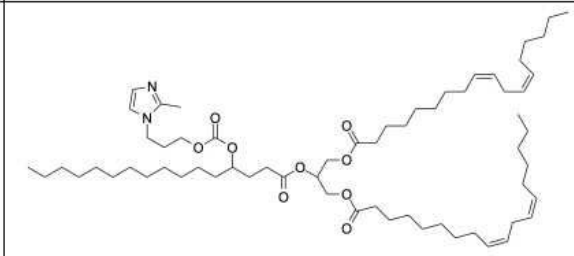
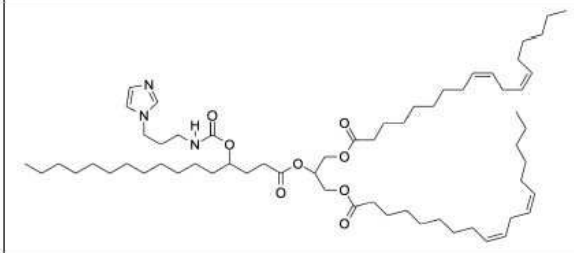
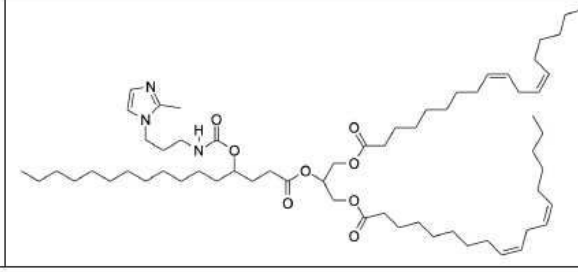
[1133]

56	
57	
58	
59	
60	
61	

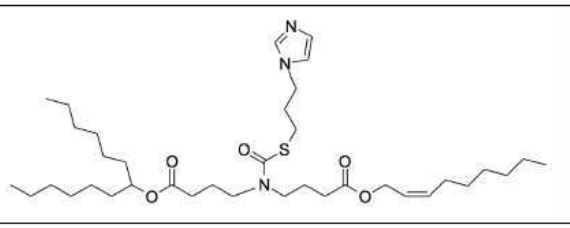
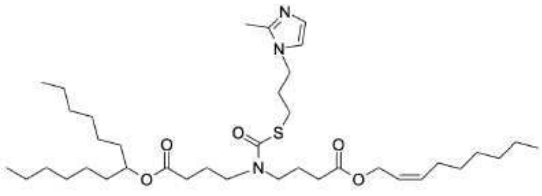
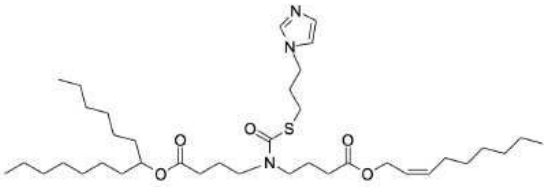
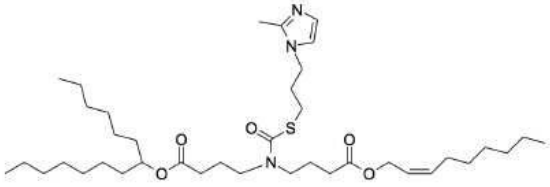
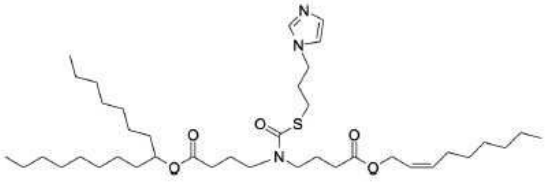
[1134]

62	
63	
64	
65	
66	
67	
68	

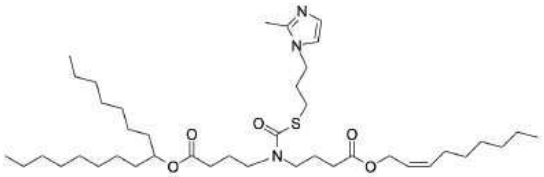
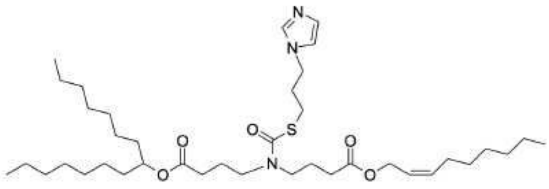
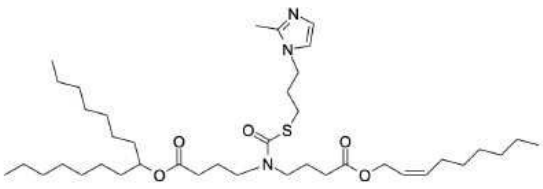
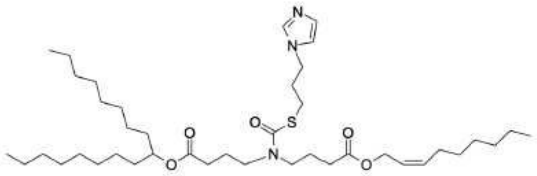
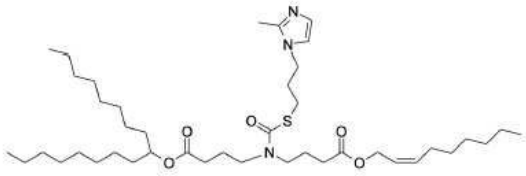
[1135]

69	
70	
71	
72	

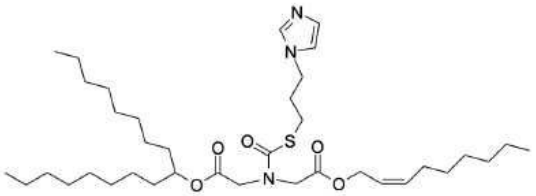
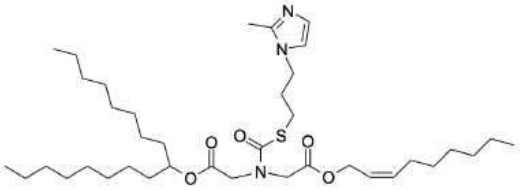
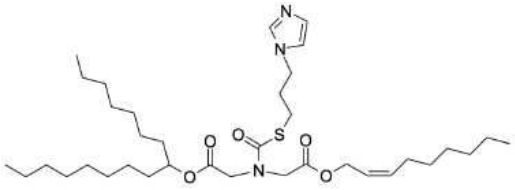
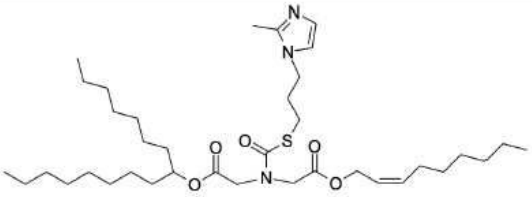
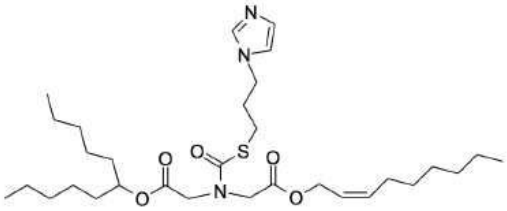
[1136]

73	
74	
75	
76	
77	

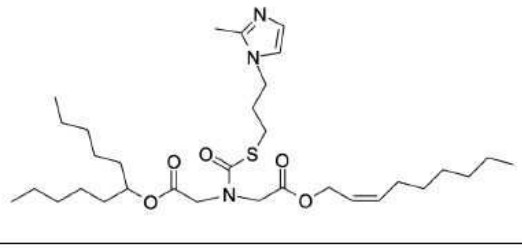
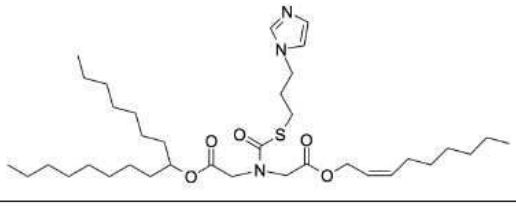
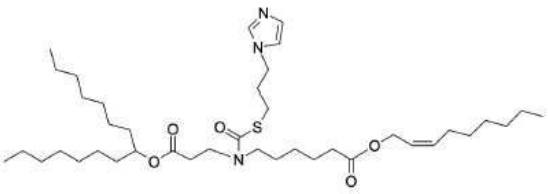
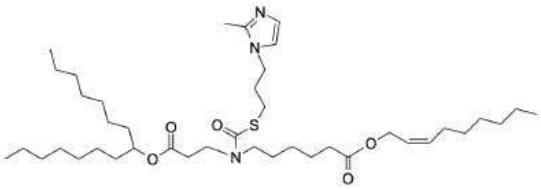
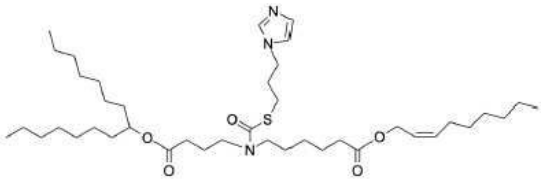
[1137]

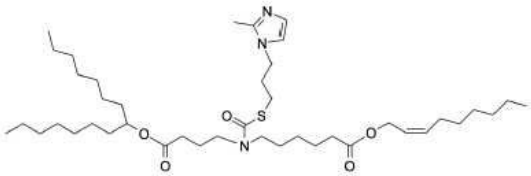
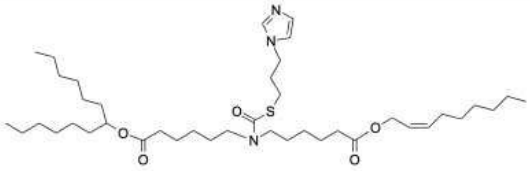
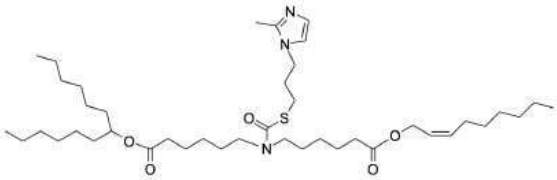
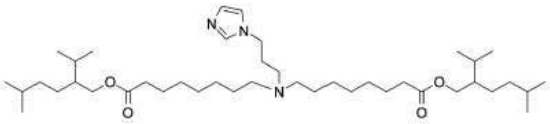
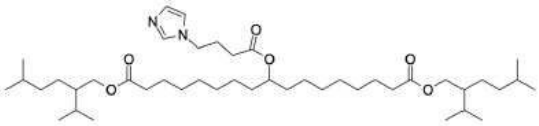
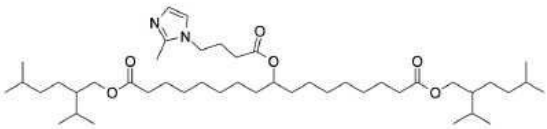
78	
79	
80	
81	
82	

[1138]

83	
84	
85	
86	
87	

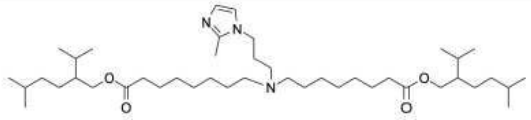
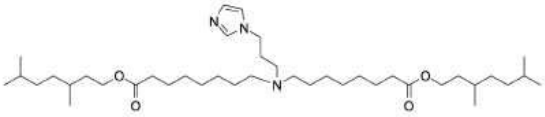
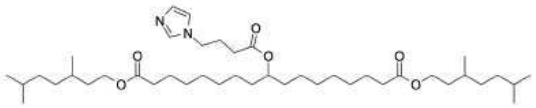
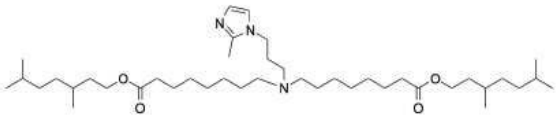
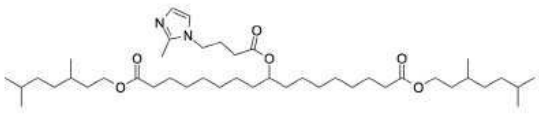
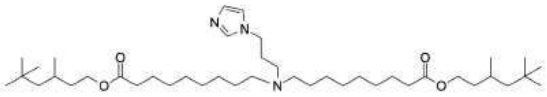
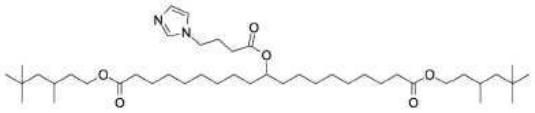
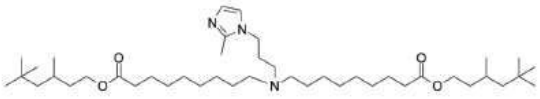
[1139]

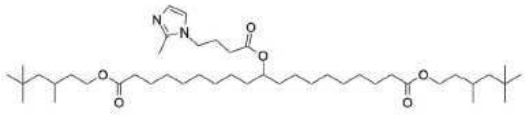
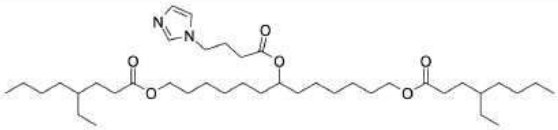
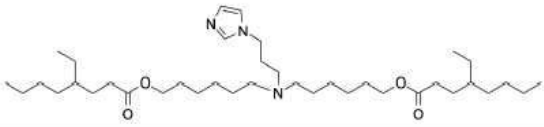
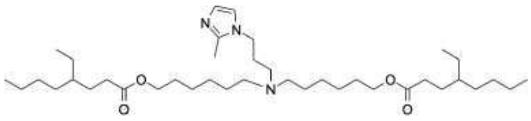
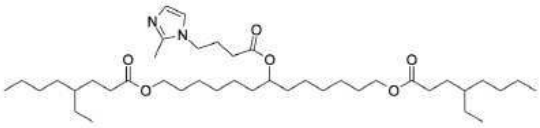
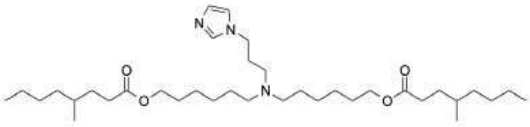
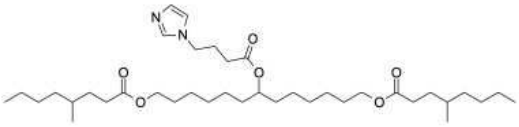
88	
89	
90	
91	
92	

93	
94	
95	
96	
97	
98	

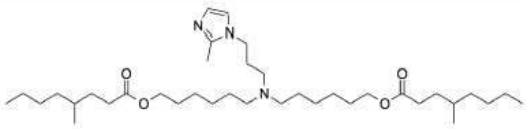
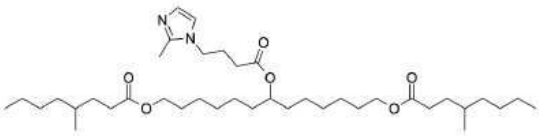
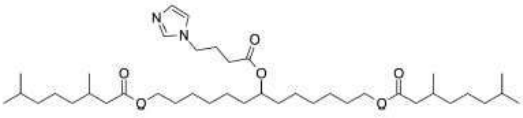
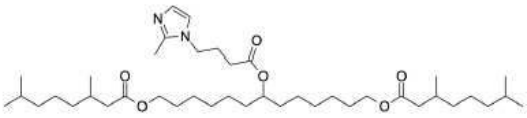
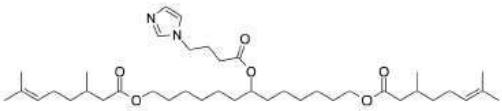
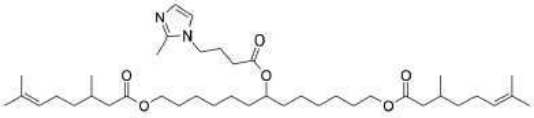
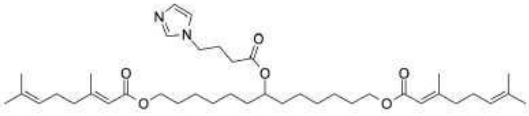
[1140]

[1141]

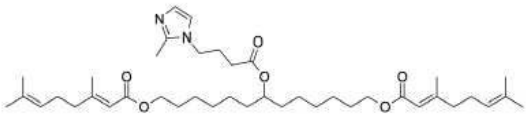
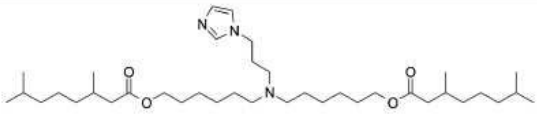
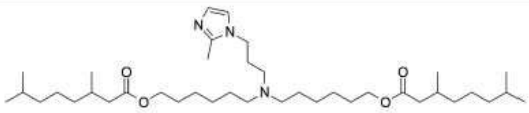
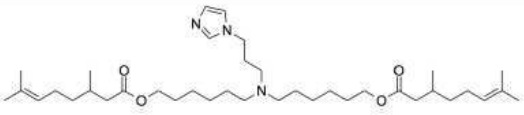
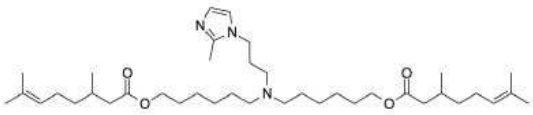
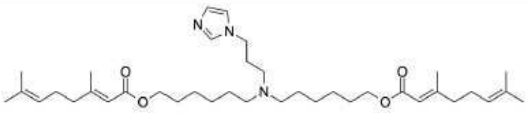
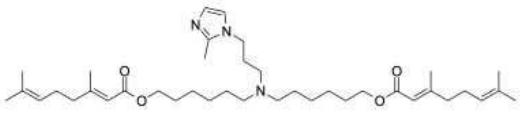
99	
100	
101	
102	
103	
104	
105	
106	

107	
108	
109	
110	
111	
112	
113	

[1142]

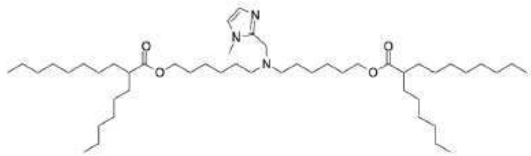
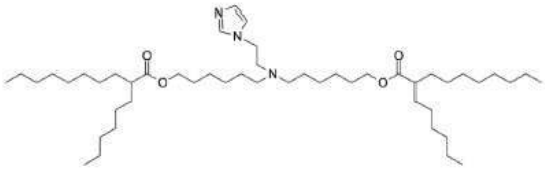
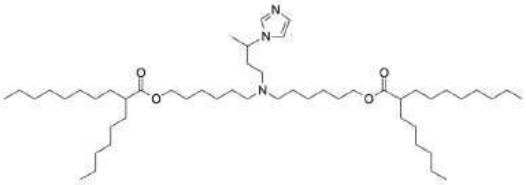
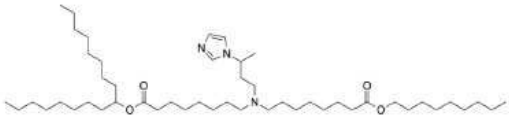
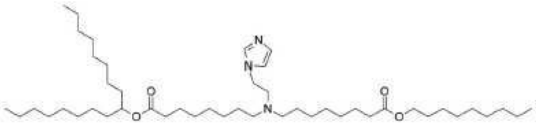
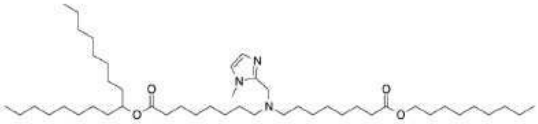
114	
115	
116	
117	
118	
119	
120	

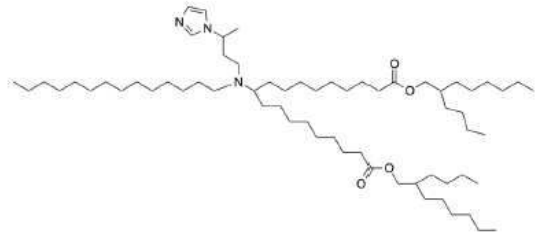
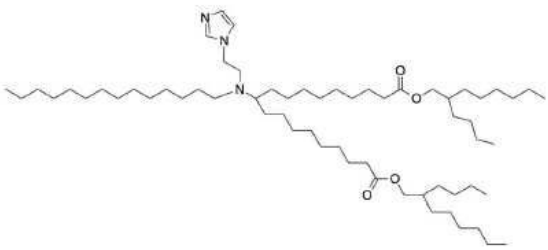
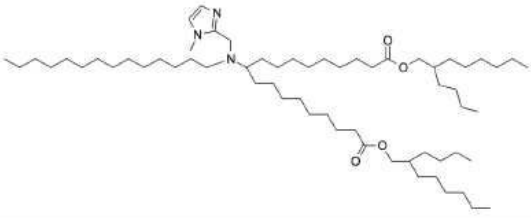
[1143]

121	
122	
123	
124	
125	
126	
127	

[1144]

[1145]

128	
129	
130	
131	
132	
133	

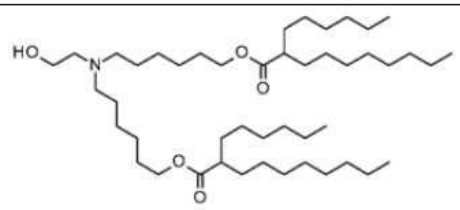
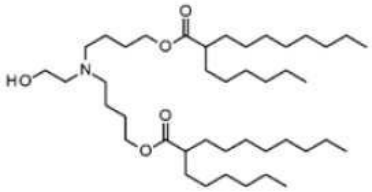
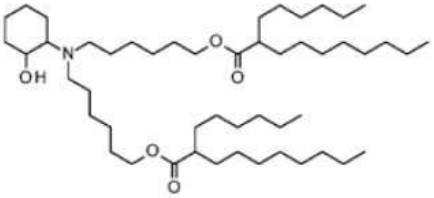
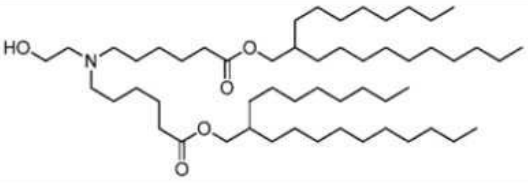
134	
135	
136	

[1146]

[1147] 일부 실시형태에서, 이온화 가능한 지질은 베타-하이드록실 아민 헤드기를 갖는다. 일부 실시형태에서, 이온화 가능한 지질은 감마-하이드록실 아민 헤드기를 갖는다.

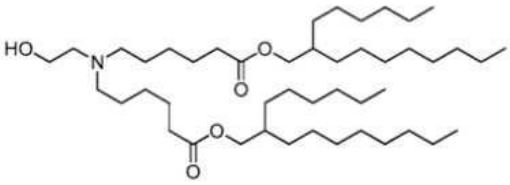
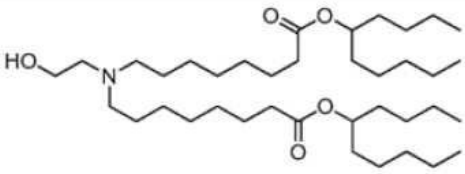
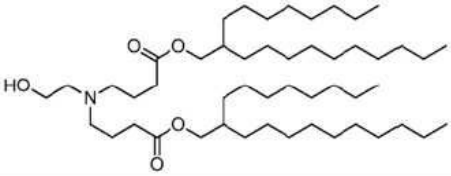
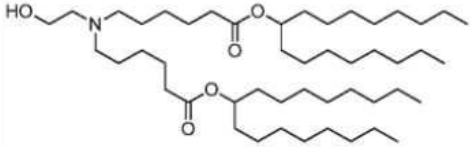
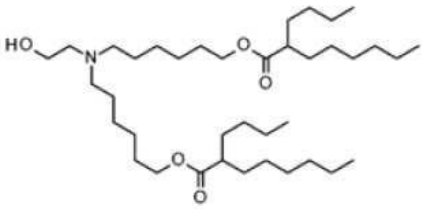
[1148] 일부 실시형태에서, 본 개시내용의 이온화 가능한 지질은 표 10b로부터 선택된 지질이다. 일부 실시형태에서, 본 개시내용의 이온화 가능한 지질은 표 10b로부터의 지질 15이다. 실시형태에서, 이온화 가능한 지질은 미국 출원 공보 제US20170210697A1호에 기재되어 있다. 실시형태에서, 이온화 가능한 지질은 미국 출원 공보 제US20170119904A1호에 기재되어 있다.

[1149] [표 10b]

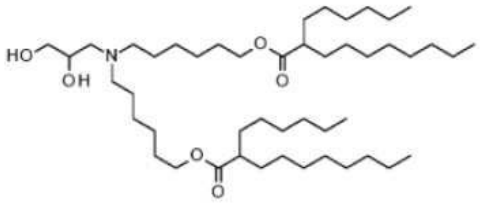
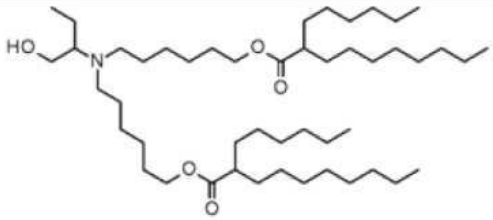
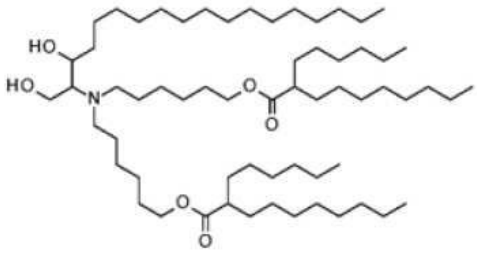
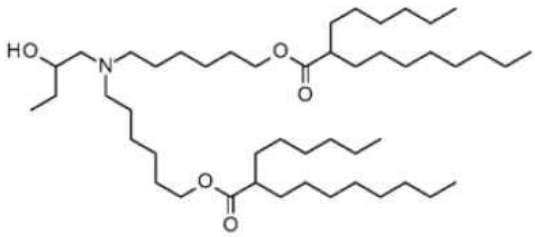
이온화 가능한 지질 번호	구조
1	
2	
3	
4	

[1150]

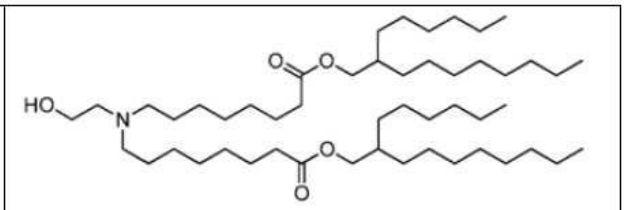
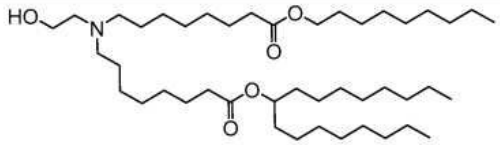
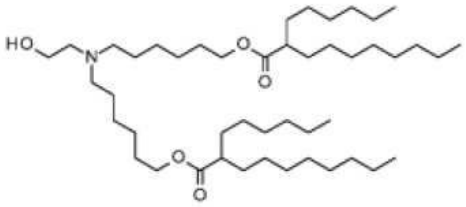
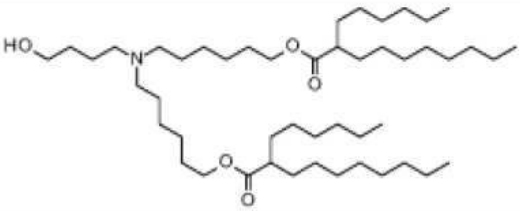
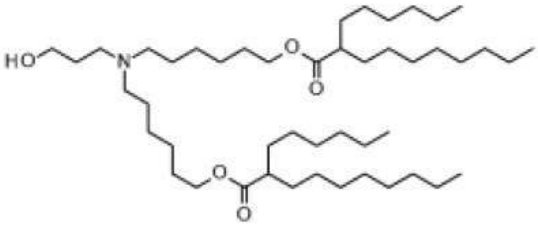
[1151]

5	
6	
7	
8	
9	

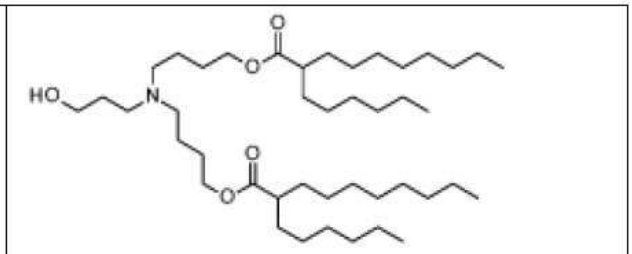
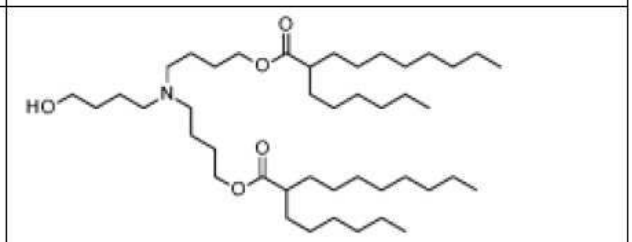
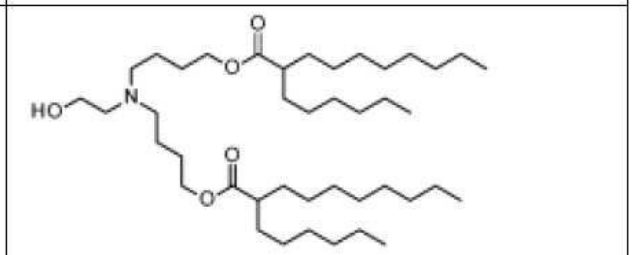
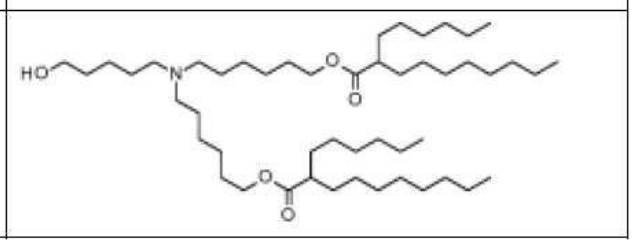
[1152]

10	
11	
12	
13	

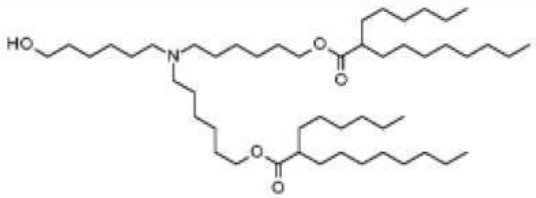
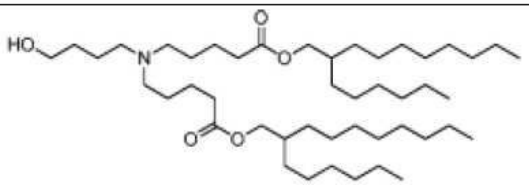
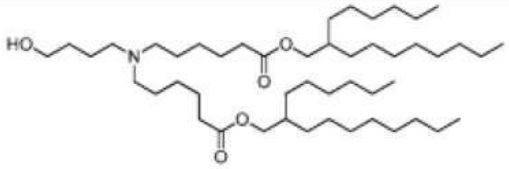
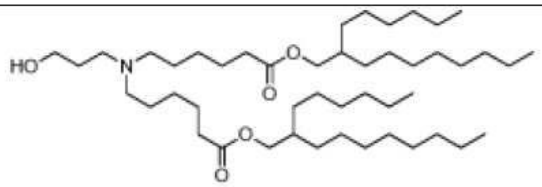
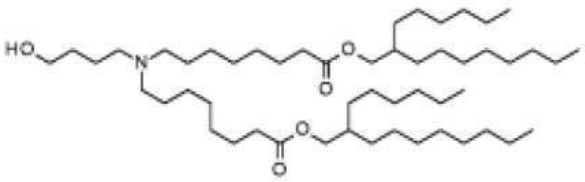
[1153]

14	
15	
16	
17	
18	

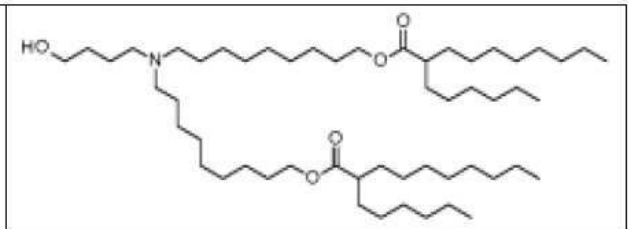
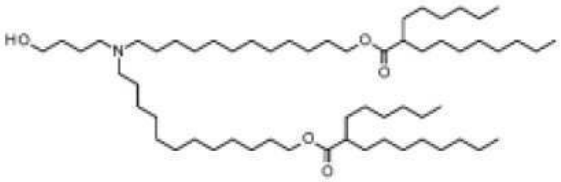
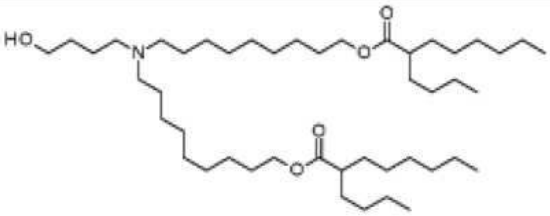
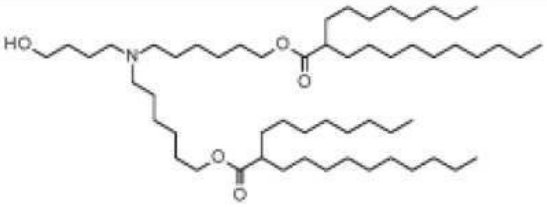
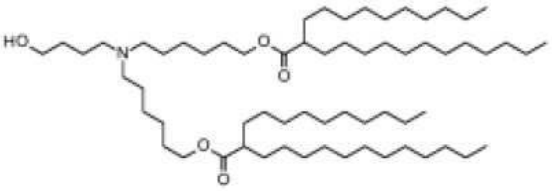
[1154]

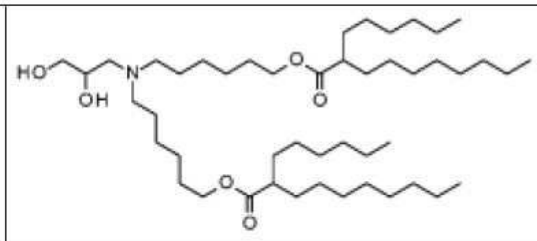
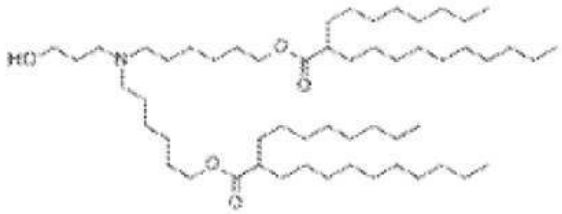
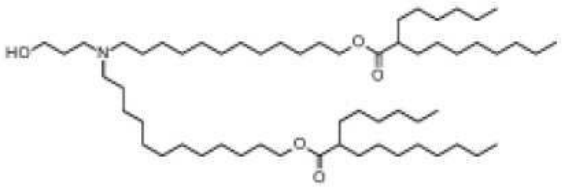
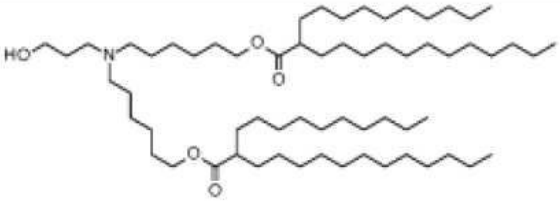
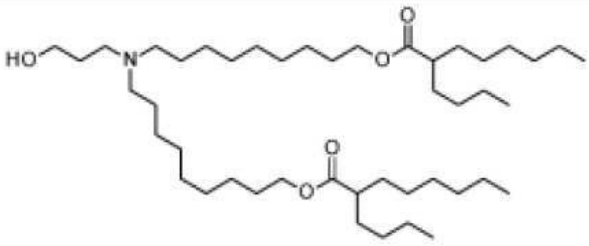
19	
20	
21	
22	

[1155]

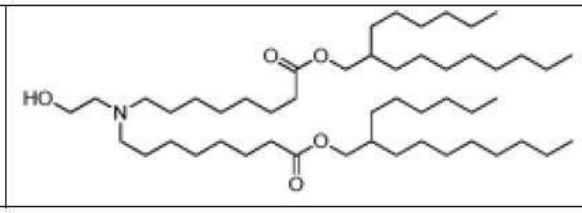
23	
24	
25	
26	
27	

[1156]

28	
29	
30	
31	
32	

33	
34	
35	
36	
37	

[1157]

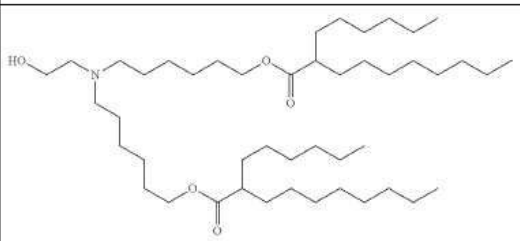
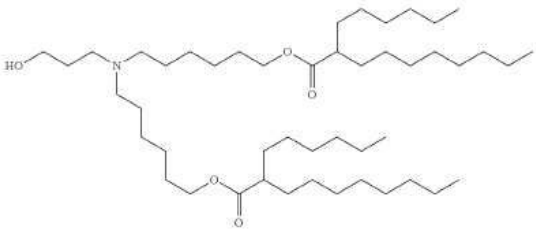
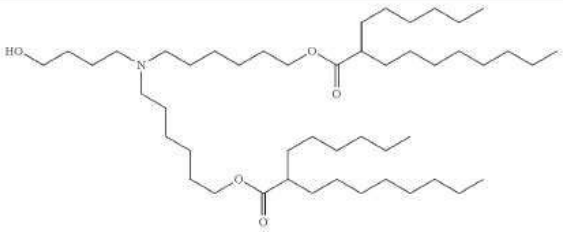
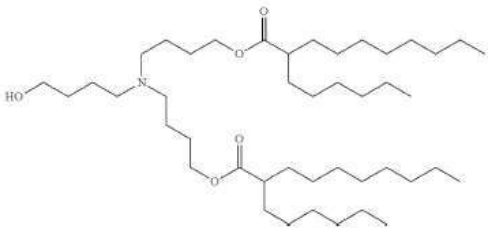
38	

[1158]

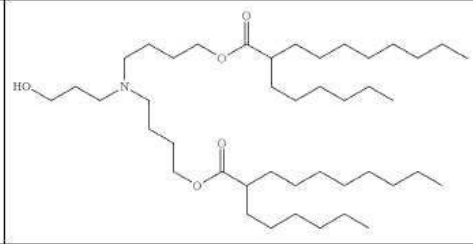
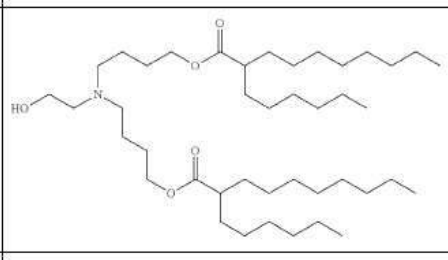
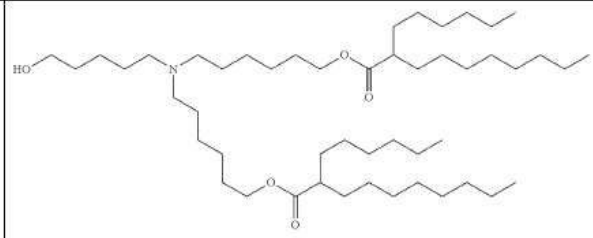
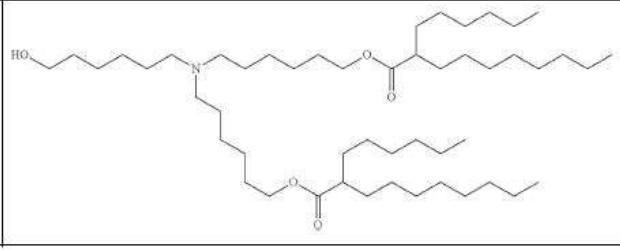
[1159] 일부 실시형태에서, 이온화 가능한 지질은 하기 표 10에 제시된 구조 중 하나를 갖는다.

표 10

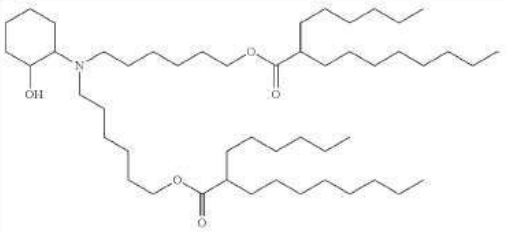
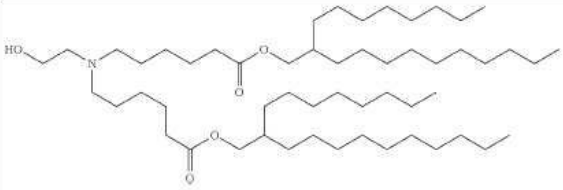
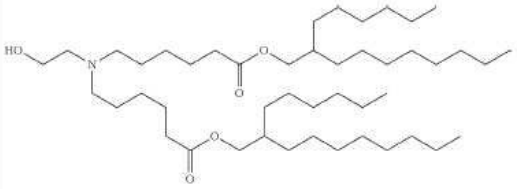
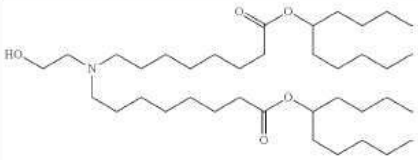
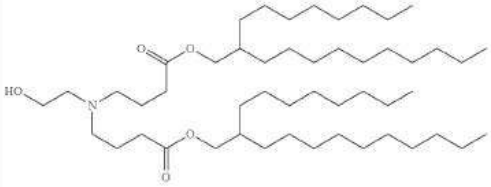
[1160]

번호	구조
1	
2	
3	
4	

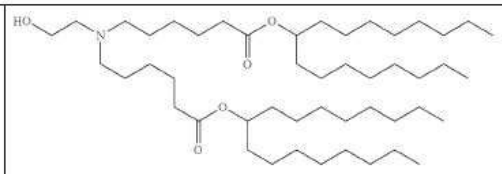
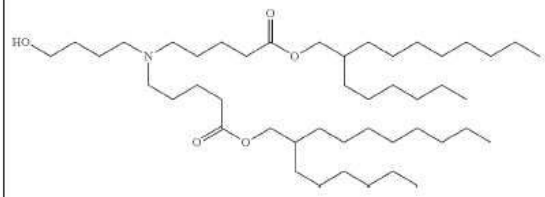
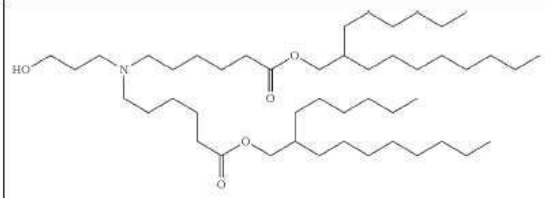
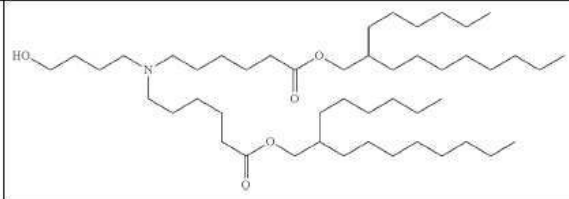
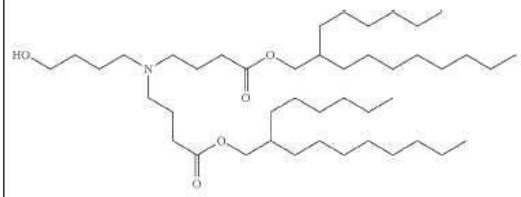
[1161]

5	
6	
7	
8	

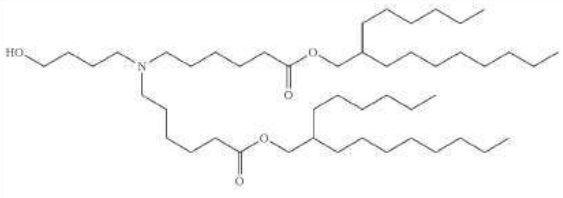
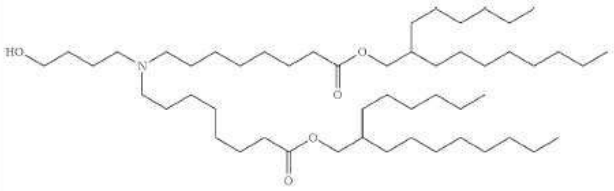
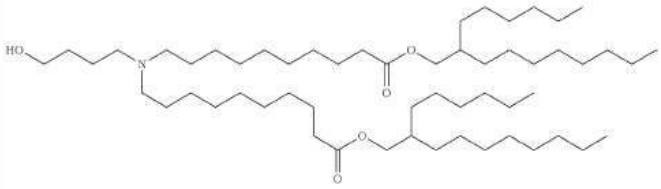
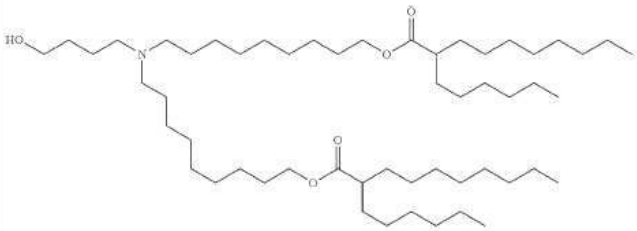
[1162]

9	
10	
11	
12	
13	

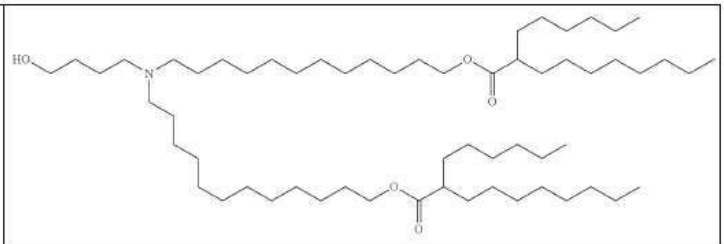
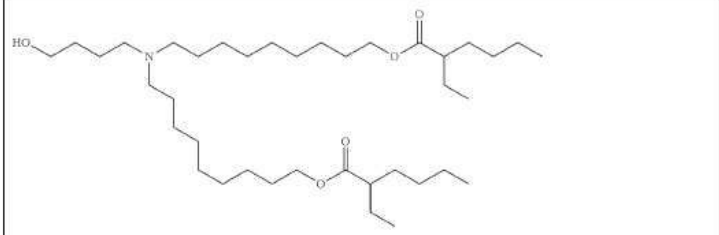
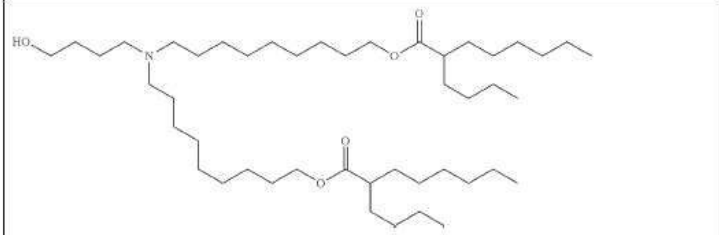
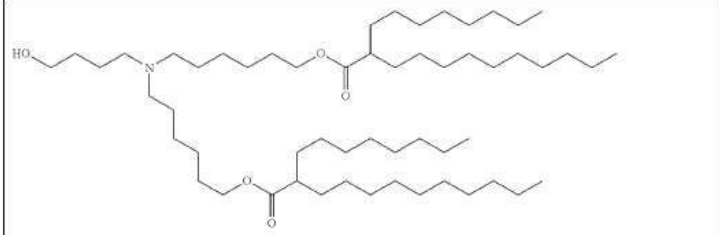
[1163]

14	
15	
16	
17	
18	

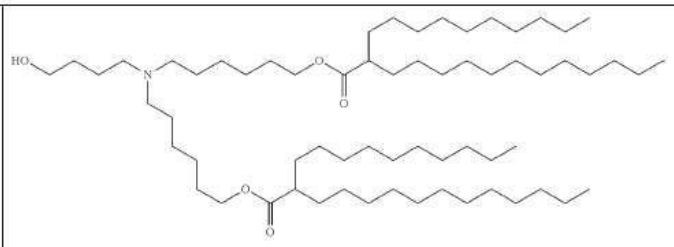
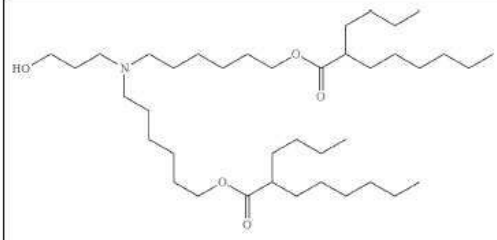
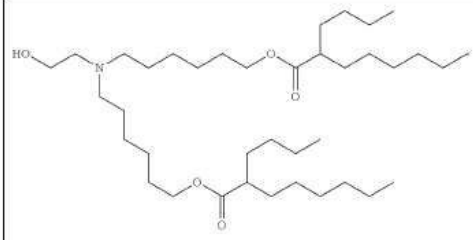
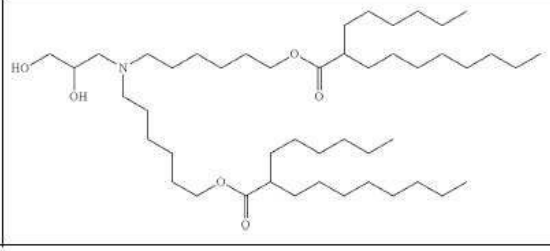
[1164]

19	
20	
21	
22	

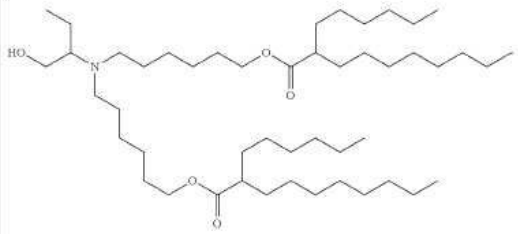
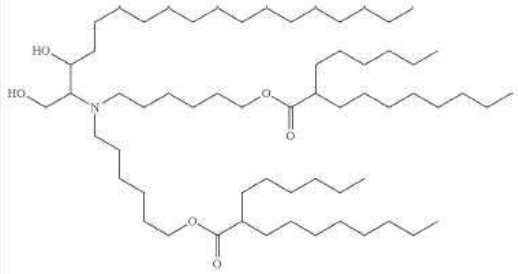
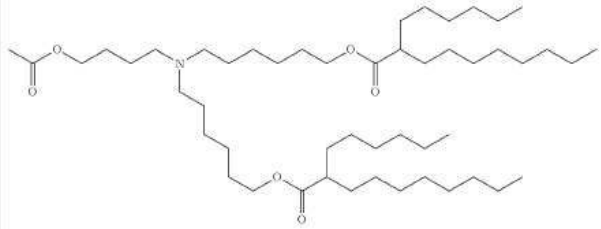
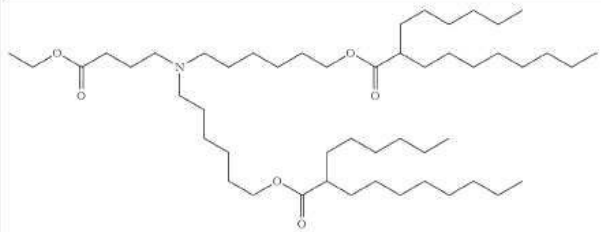
[1165]

23	
24	
25	
26	

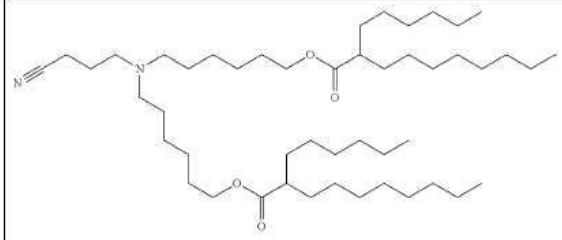
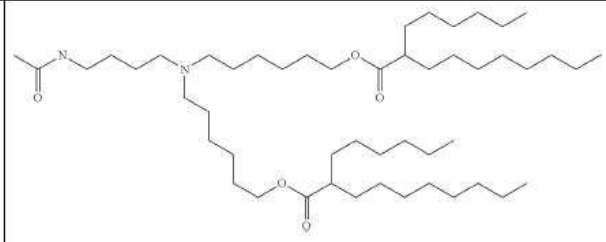
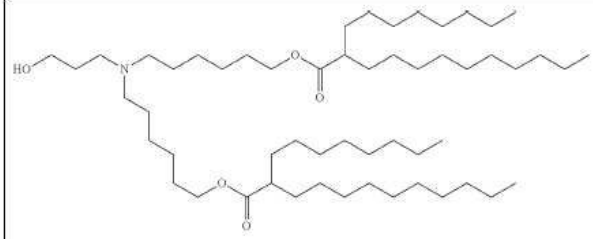
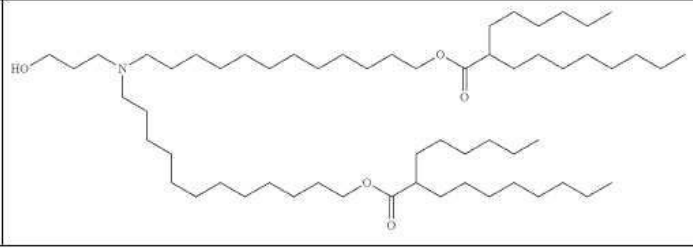
[1166]

27	
28	
29	
30	

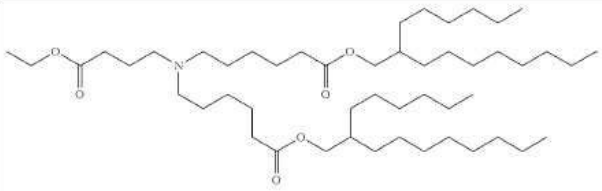
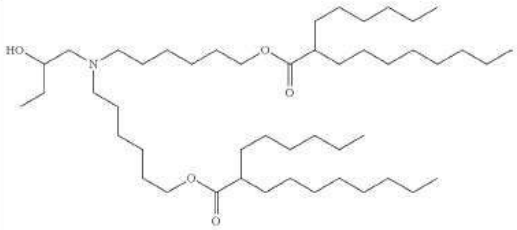
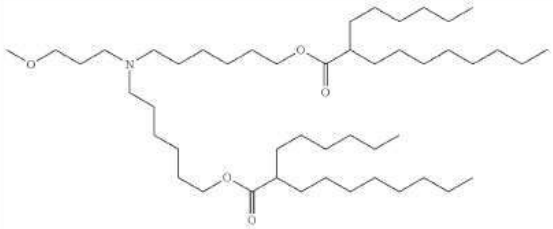
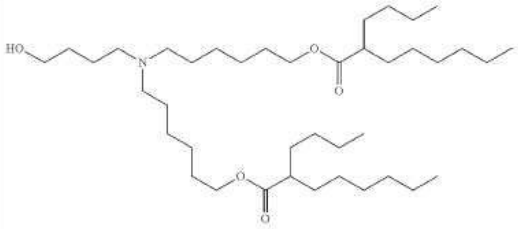
[1167]

31	
32	
33	
34	

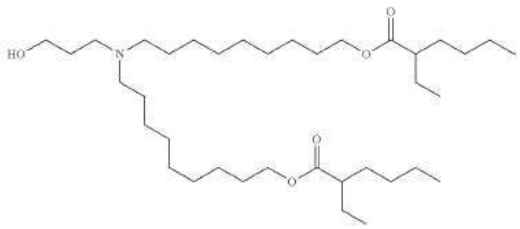
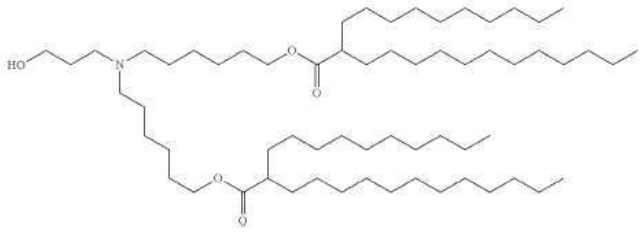
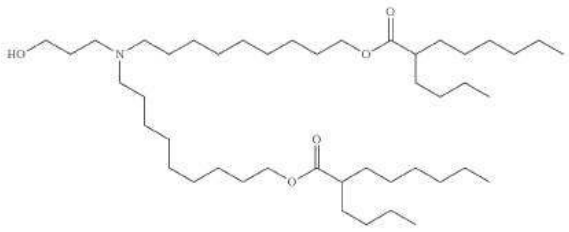
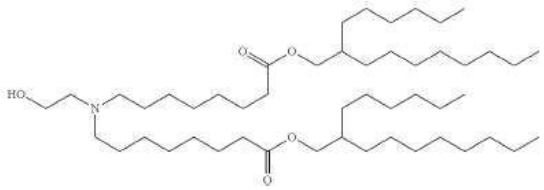
[1168]

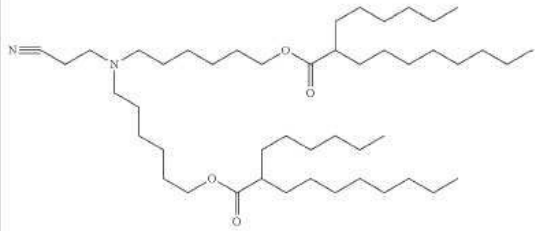
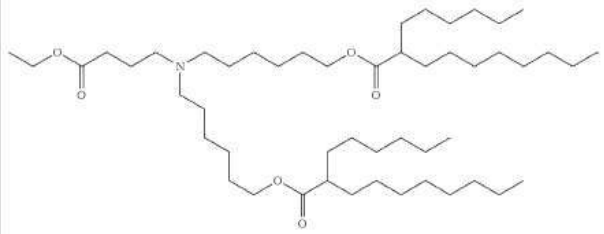
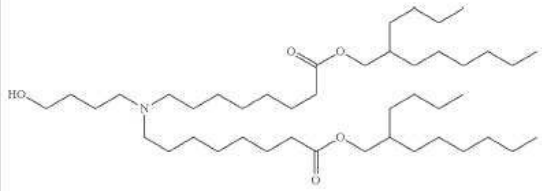
35	
36	
37	
38	

[1169]

39	
40	
41	
42	

[1170]

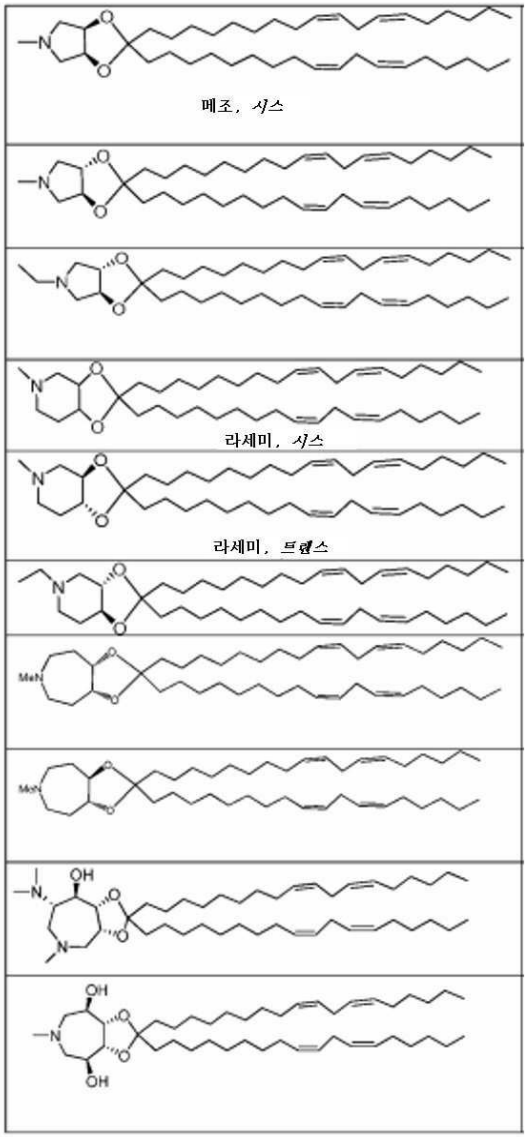
43	
44	
45	
46	

47	
48	
49	

[1171]

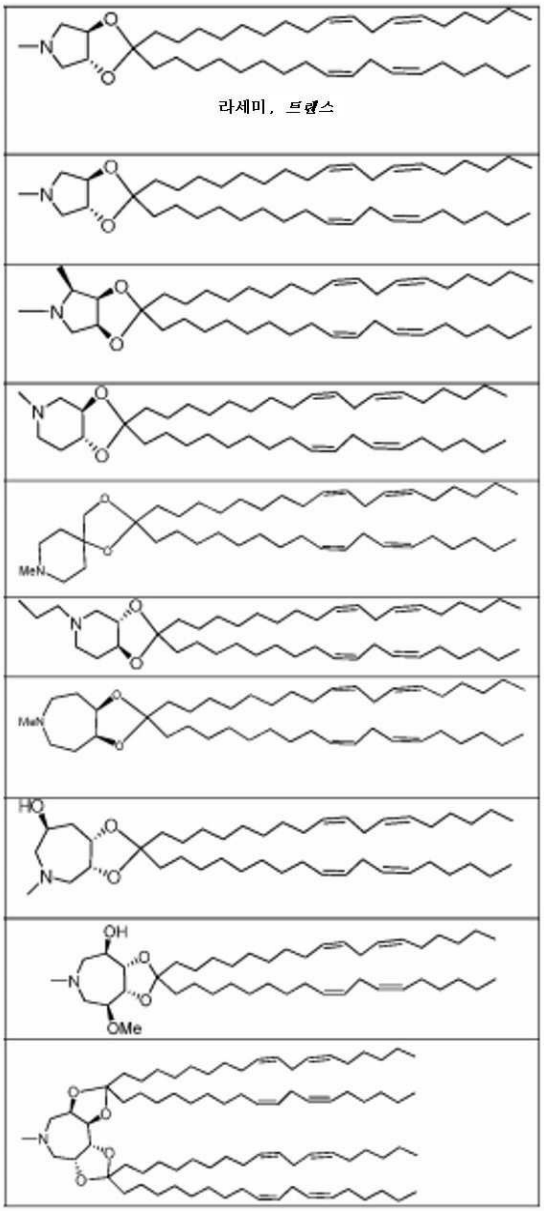
[1172] 일부 실시형태에서, 이온화 가능한 지질은 하기 표 11에 제시된 구조 중 하나를 갖는다. 일부 실시형태에서, 표 11에 제시된 바와 같은 이온화 가능한 지질은 국제 특허 출원 제PCT/US2010/061058호에 기재된 바와 같다.

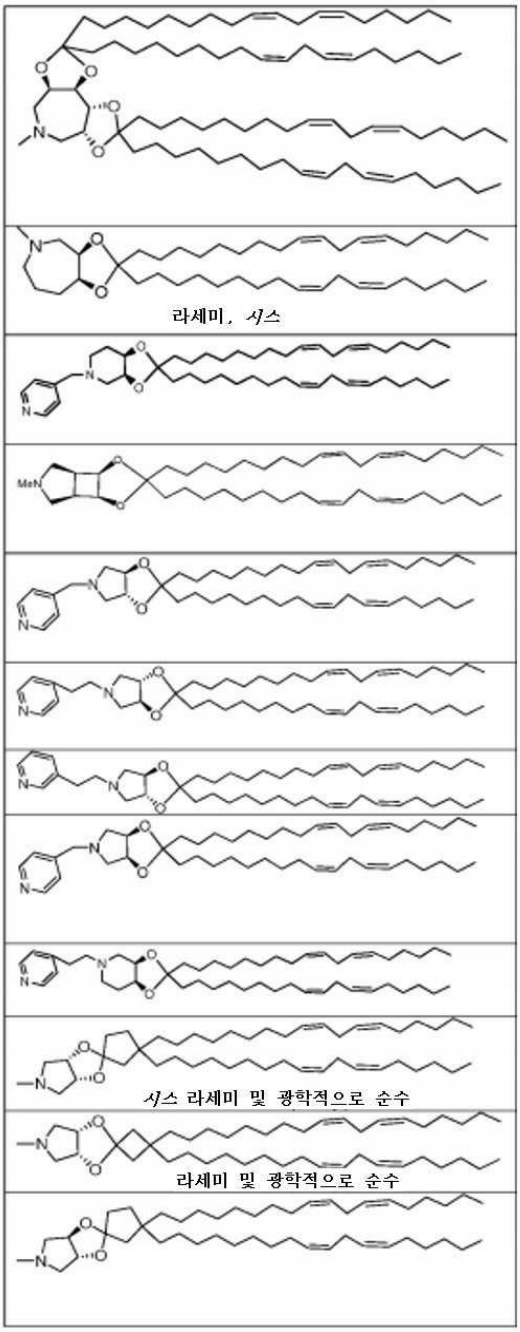
표 11



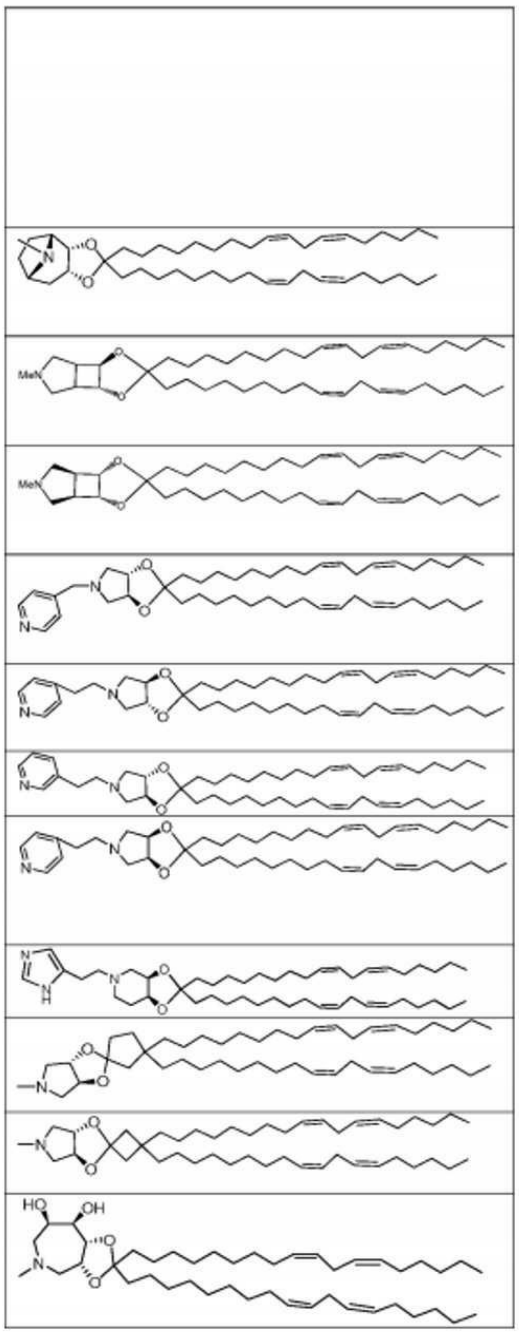
[1173]

[1174]

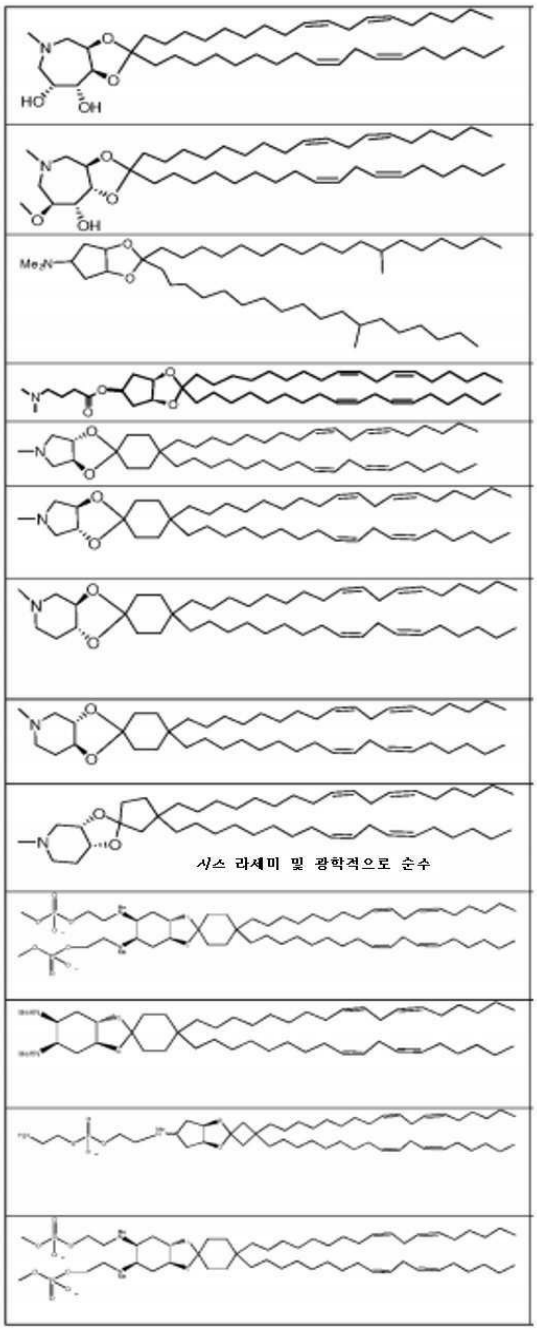




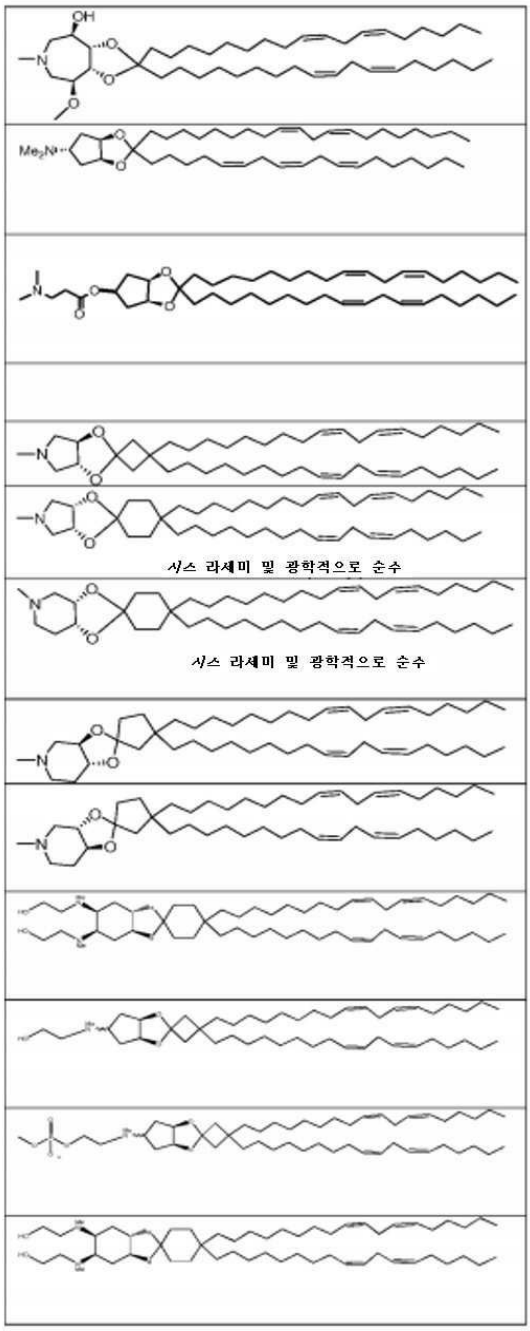
[1175]



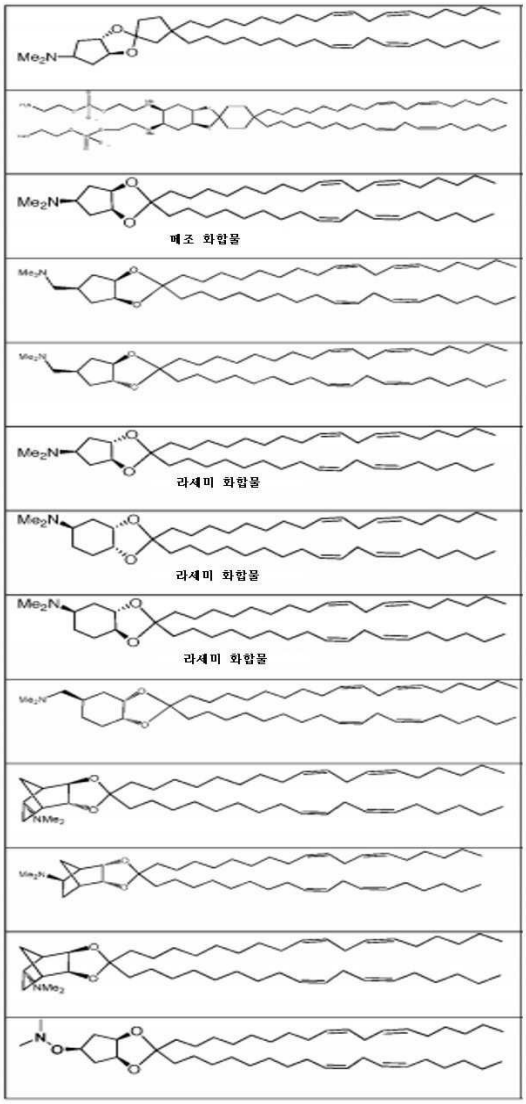
[1176]



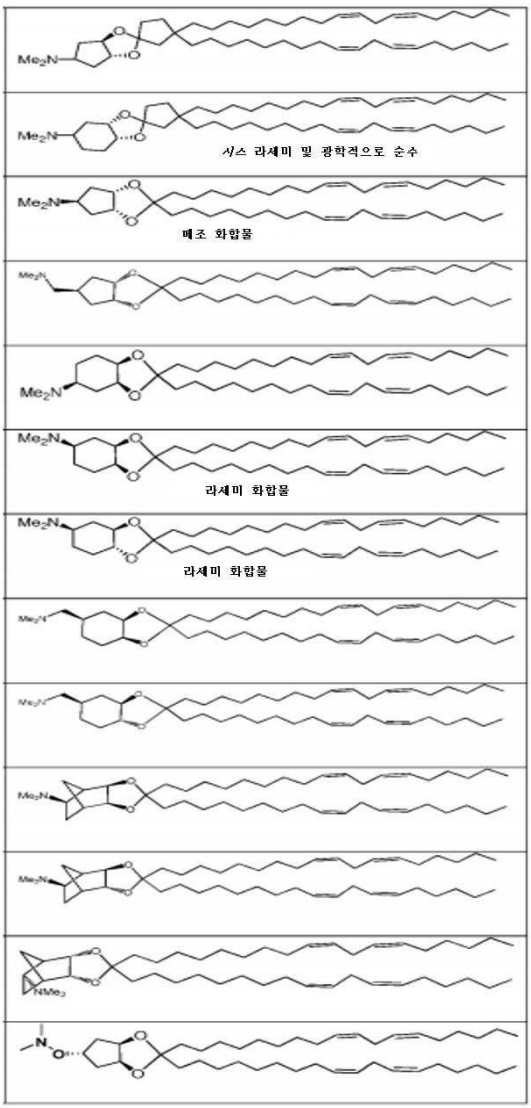
[1177]



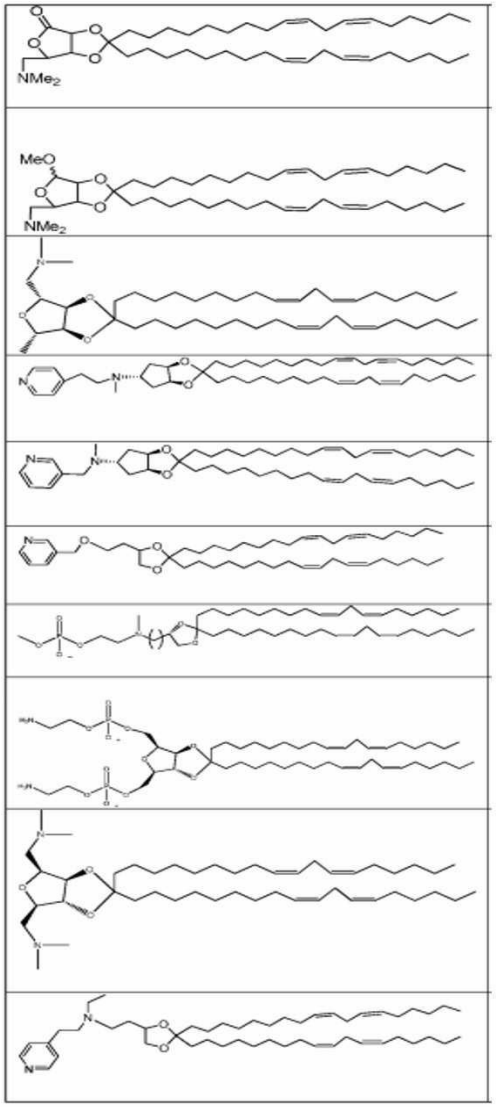
[1178]



[1179]

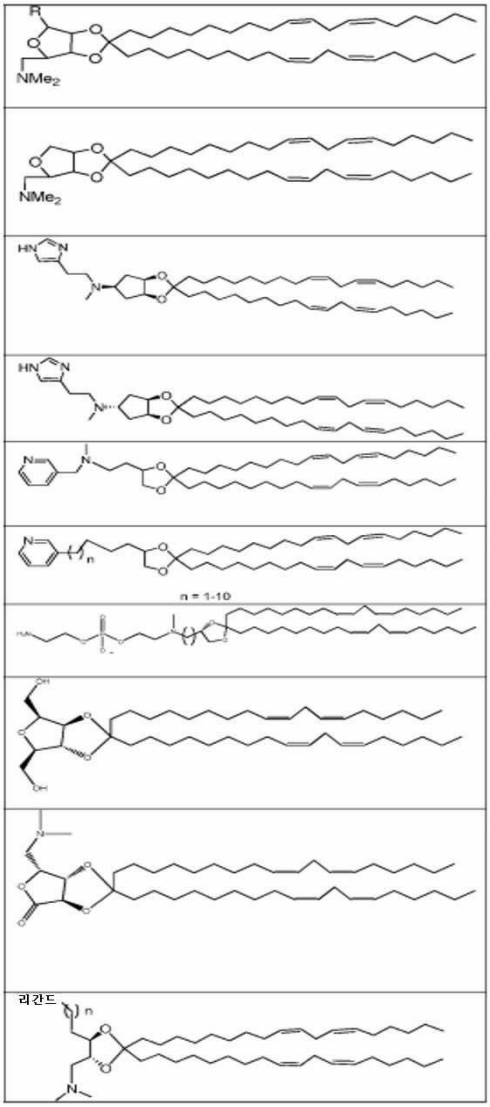


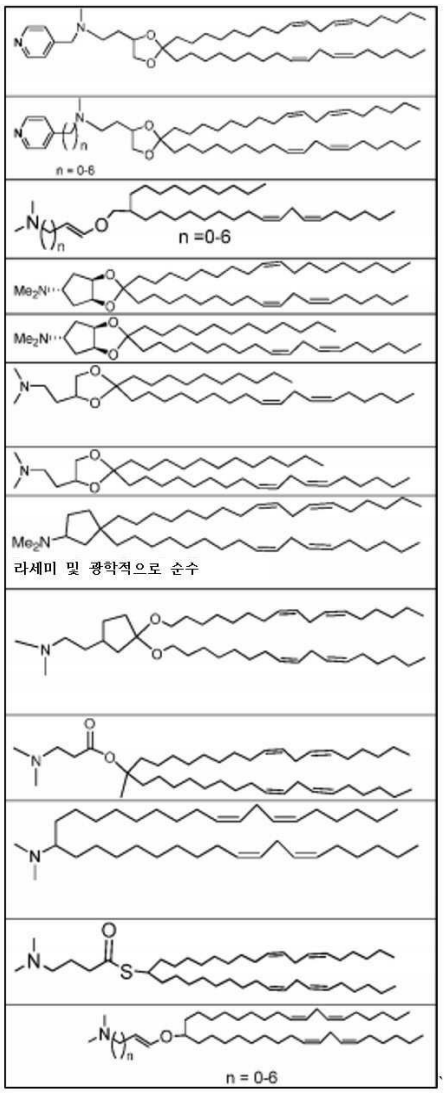
[1180]



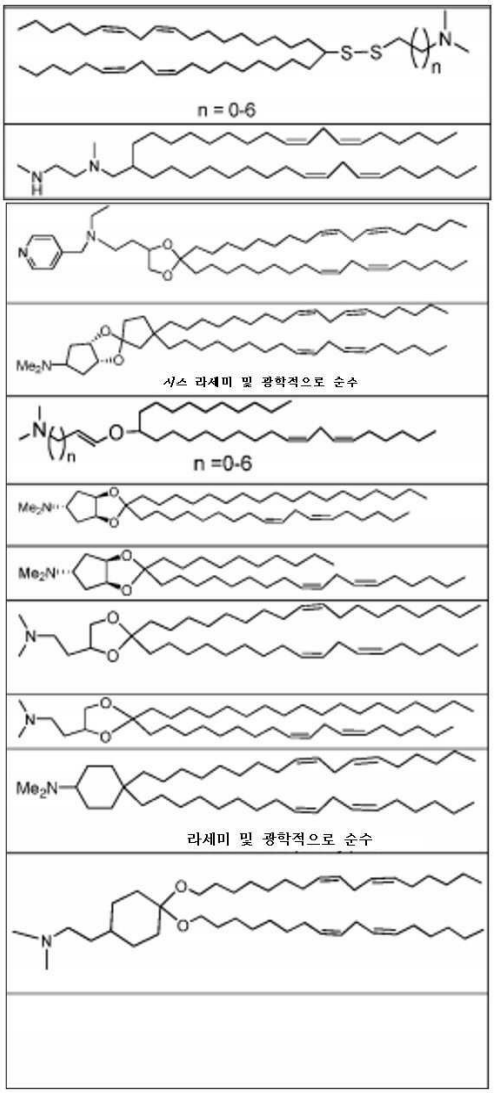
[1181]

[1182]

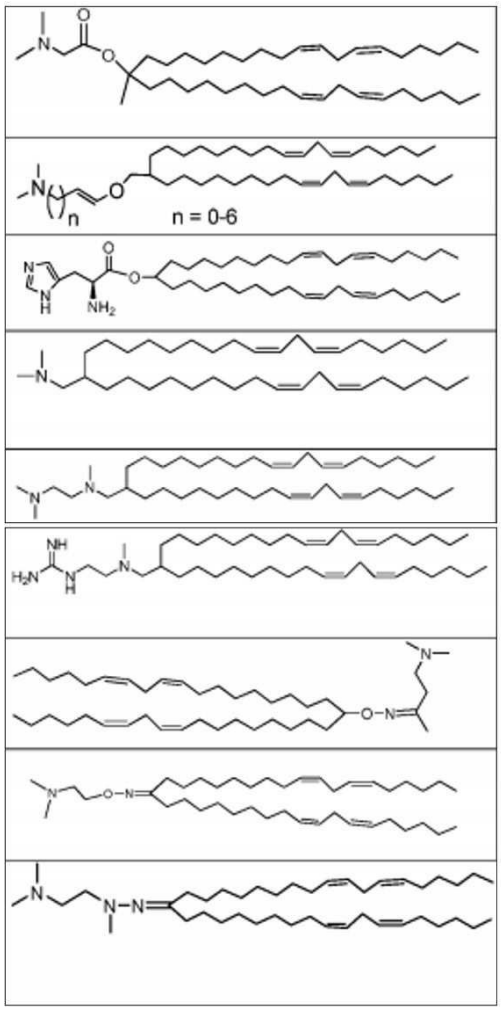




[1183]

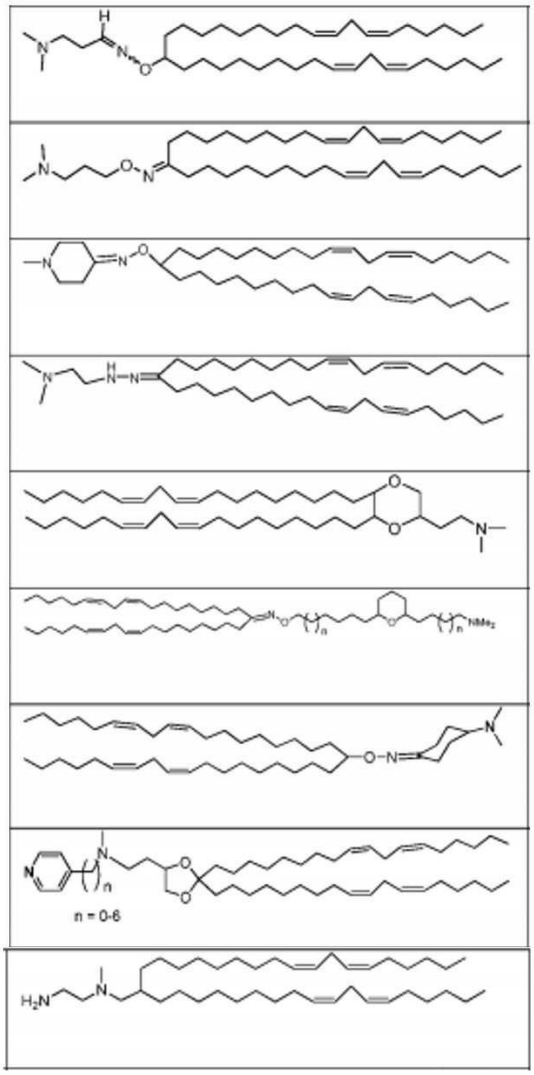


[1184]

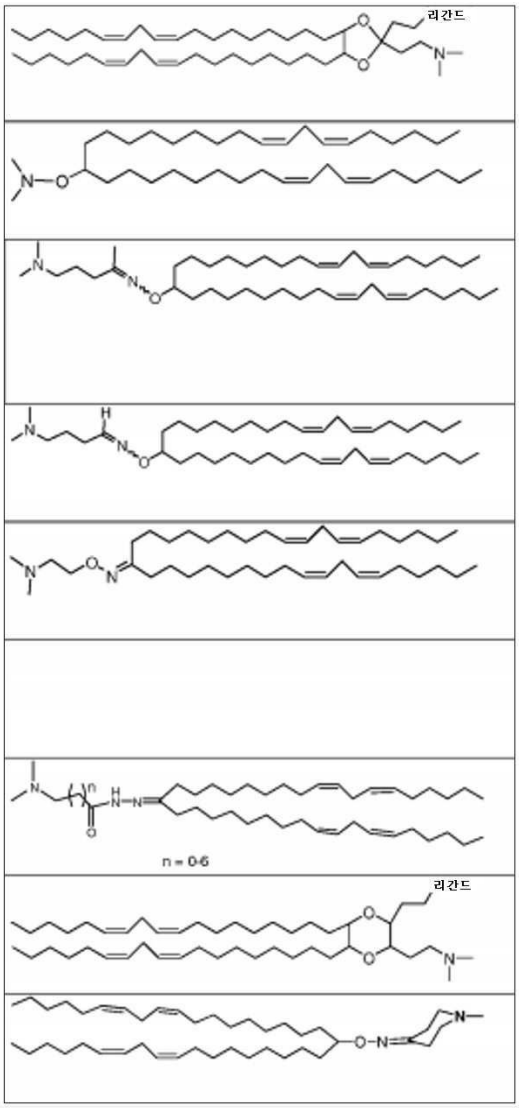


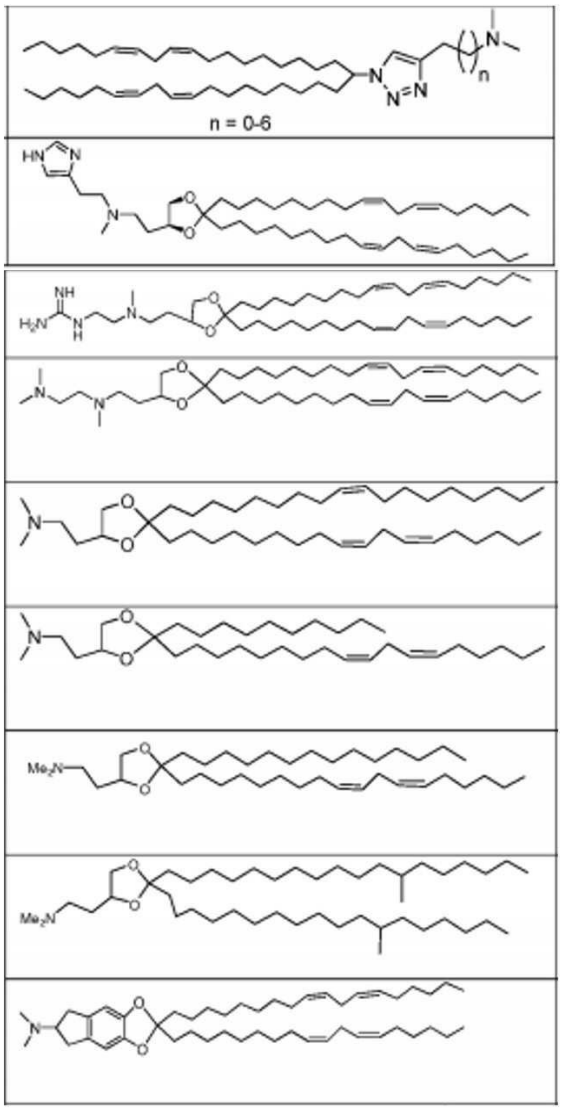
[1185]

[1186]



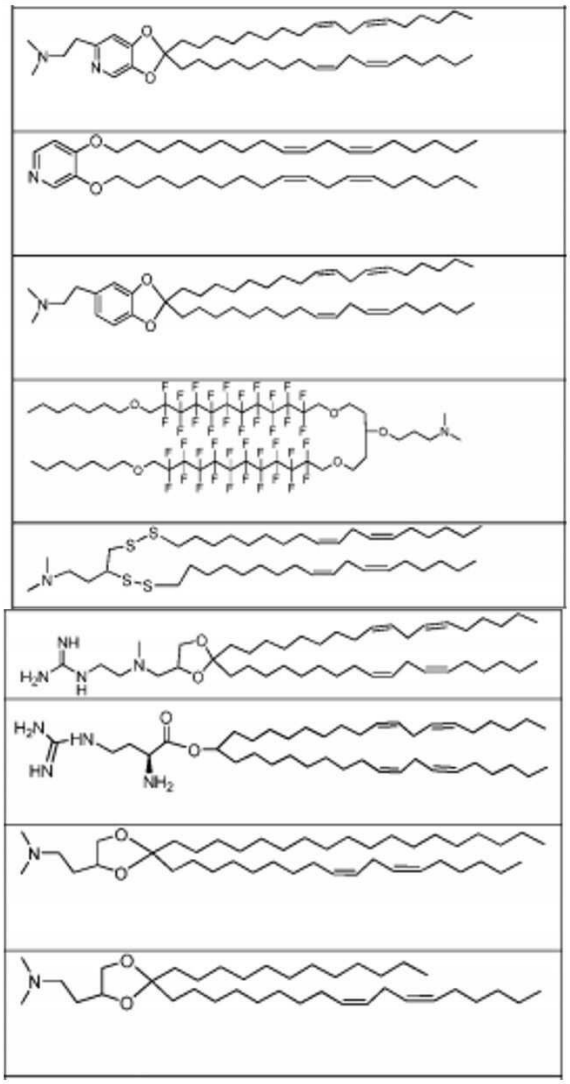
[1187]



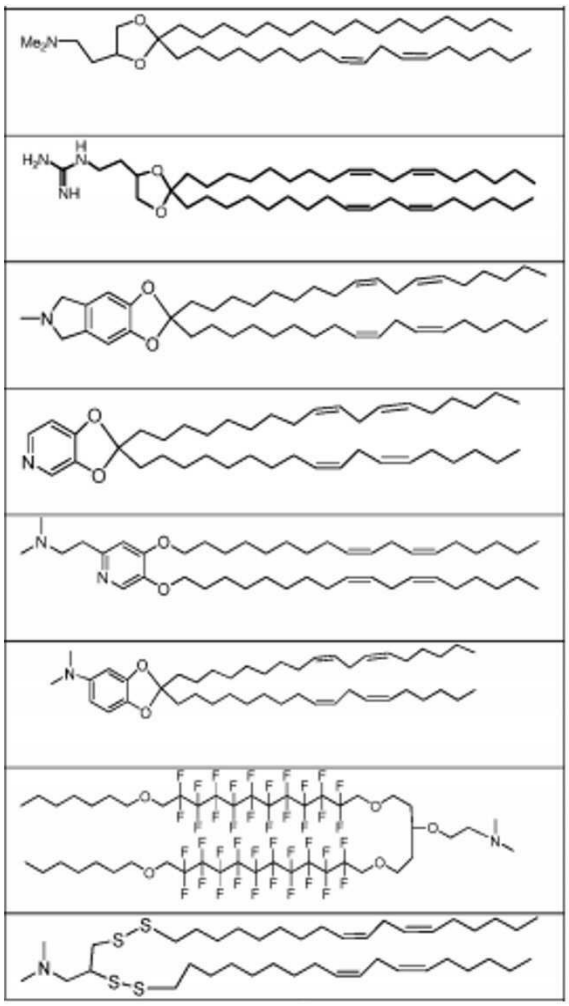


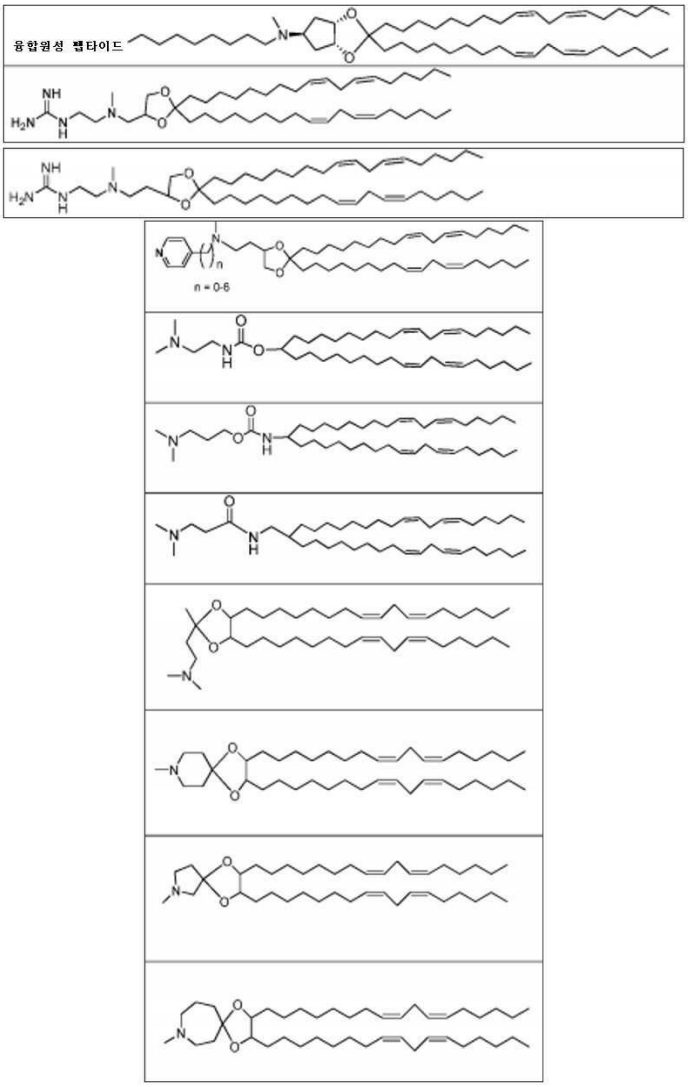
[1188]

[1189]

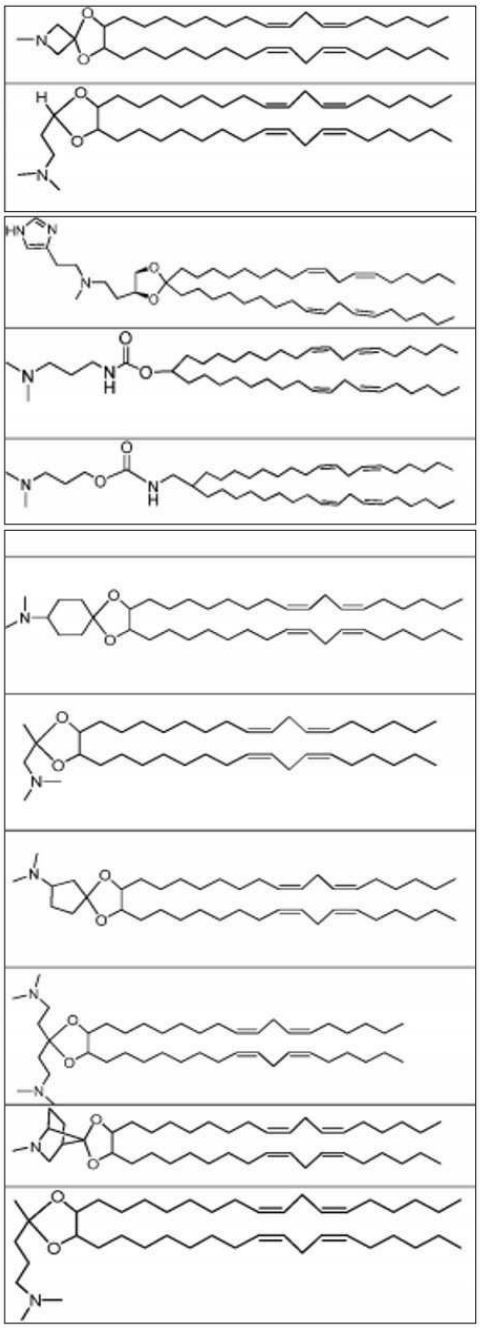


[1190]

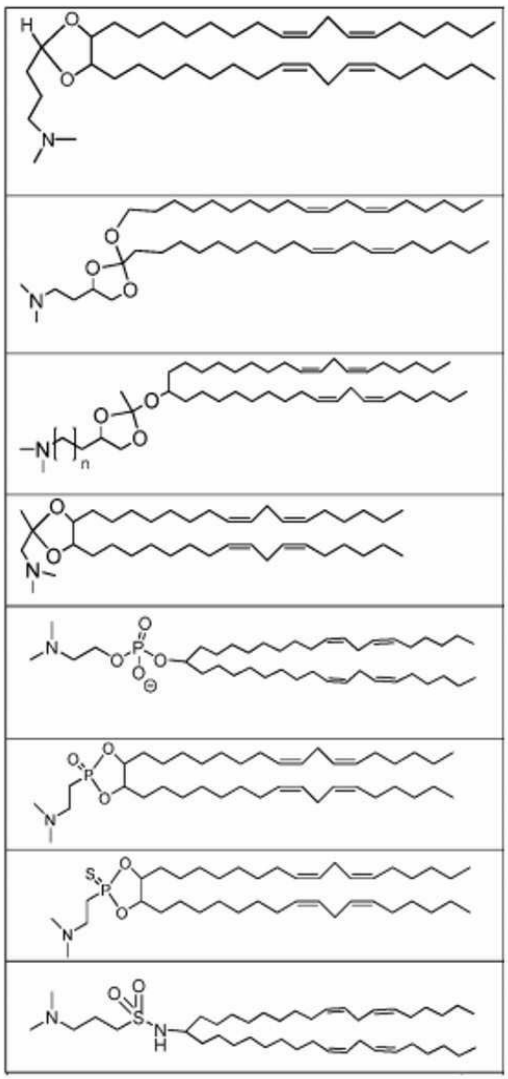




[1191]

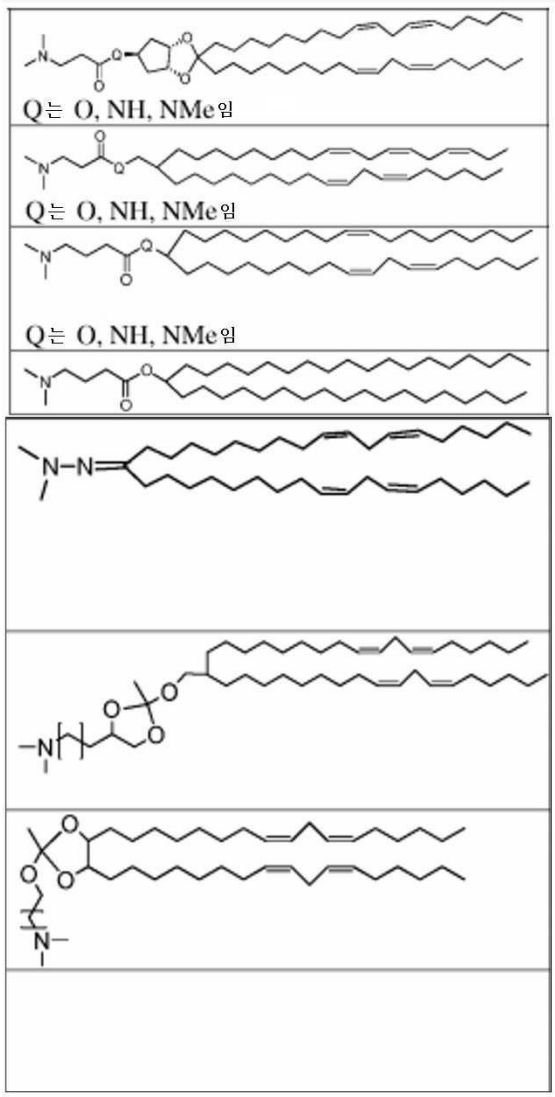


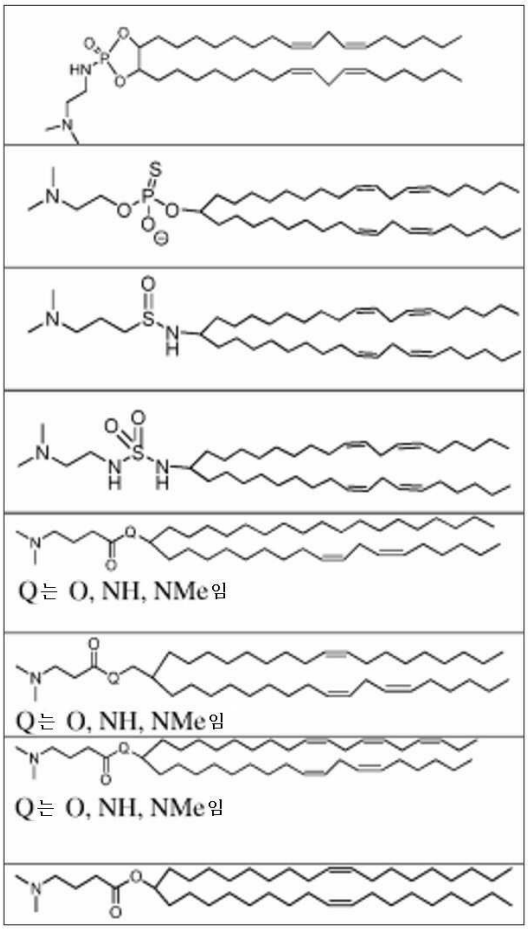
[1192]



[1193]

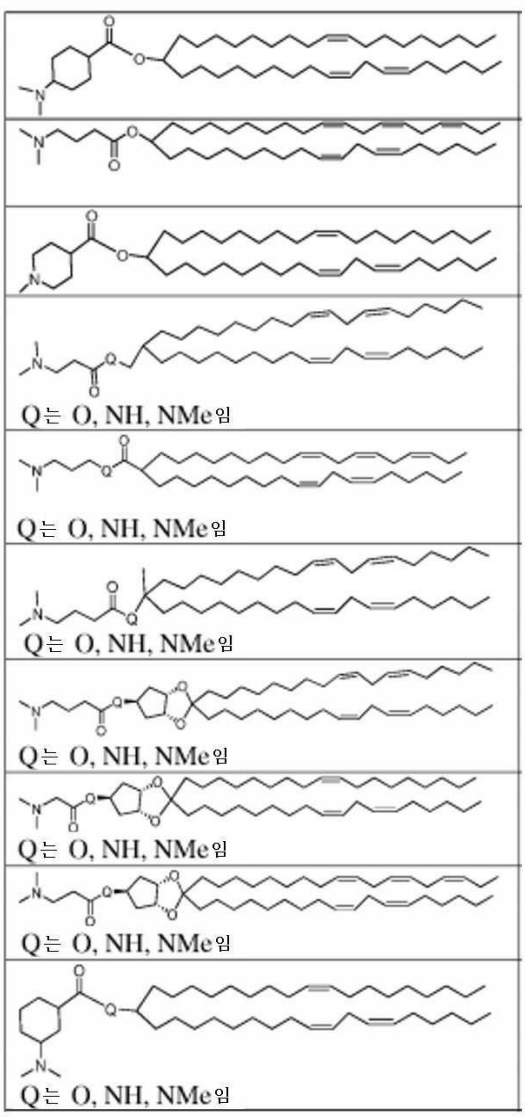
[1194]



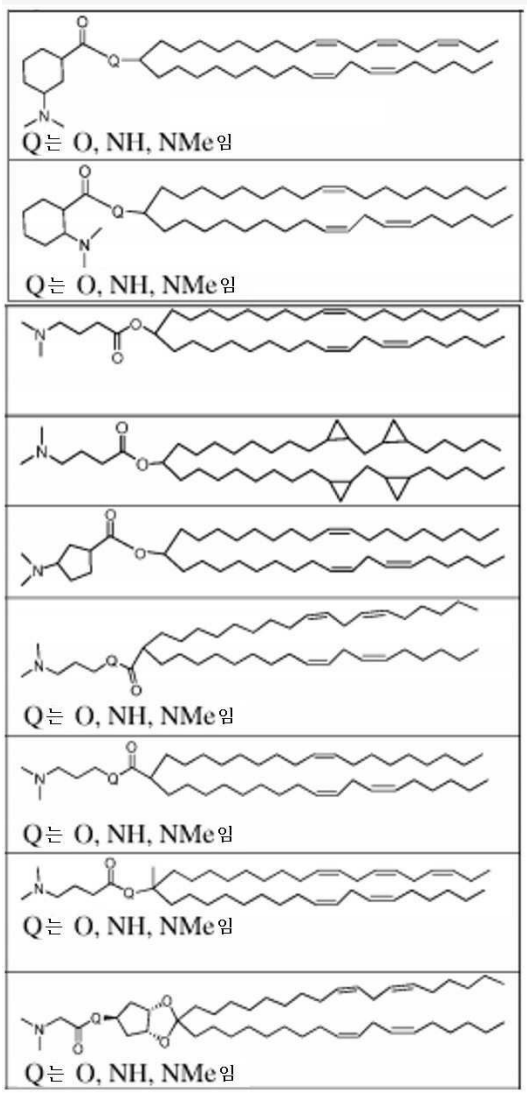


[1195]

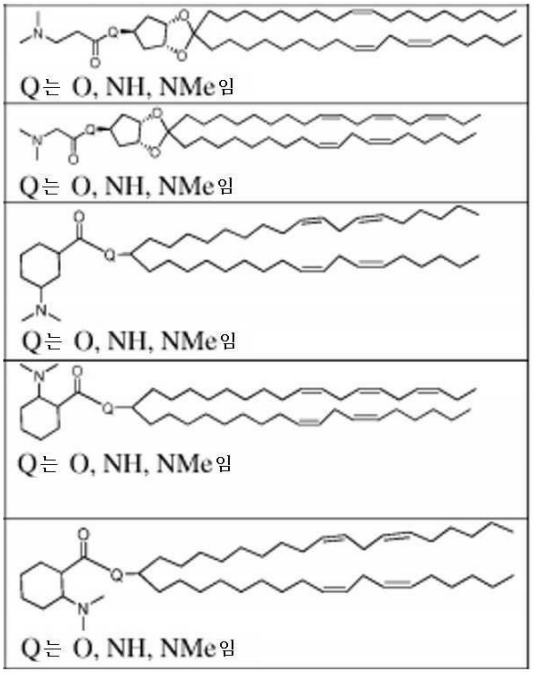
[1196]


Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임

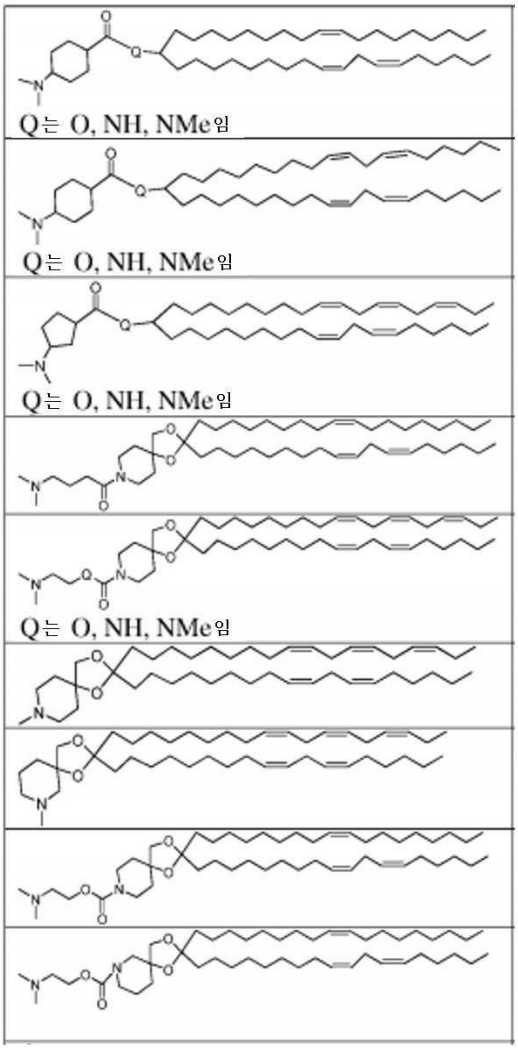
[1197]


Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임

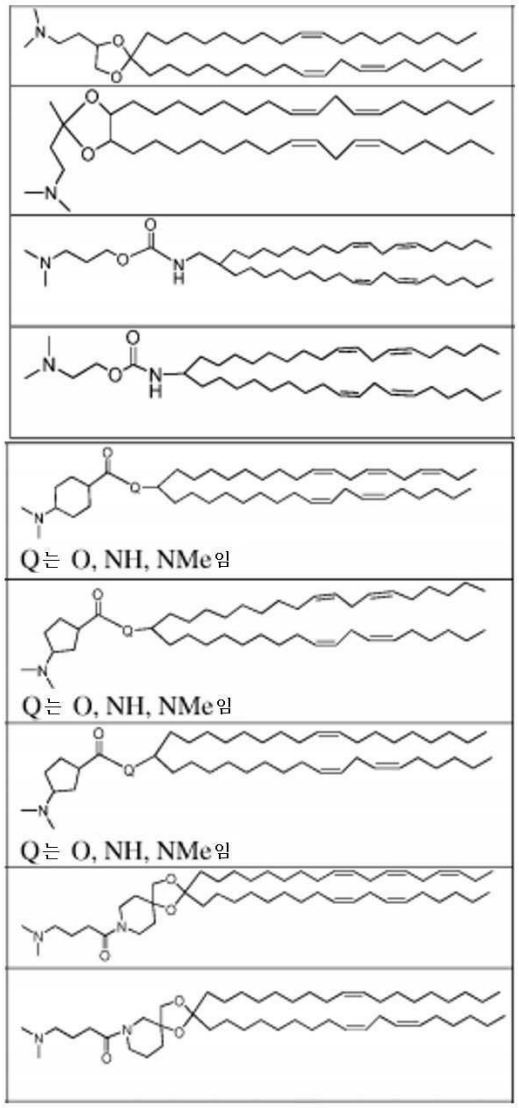
[1198]

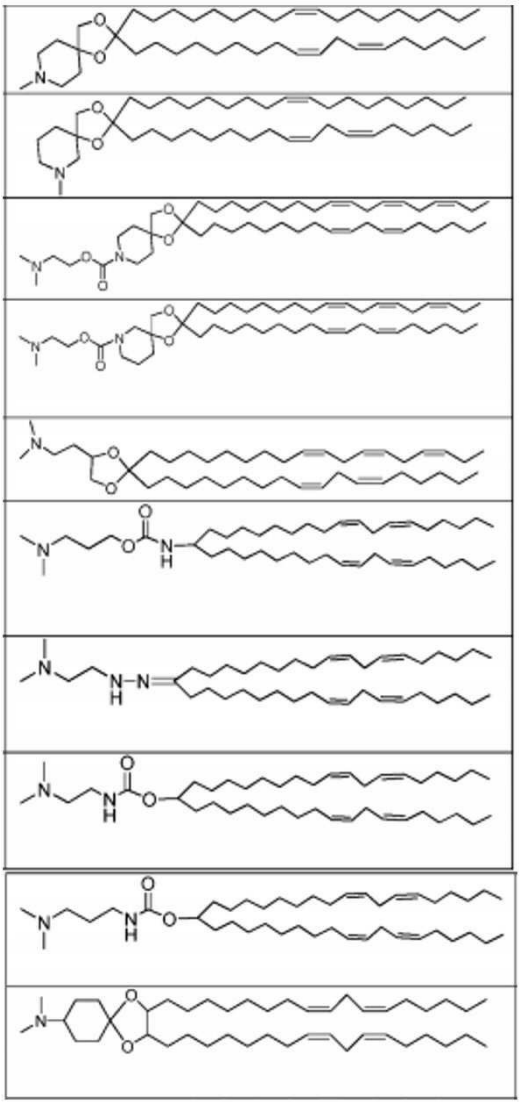

Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임
Q는 O, NH, NMe임

[1199]



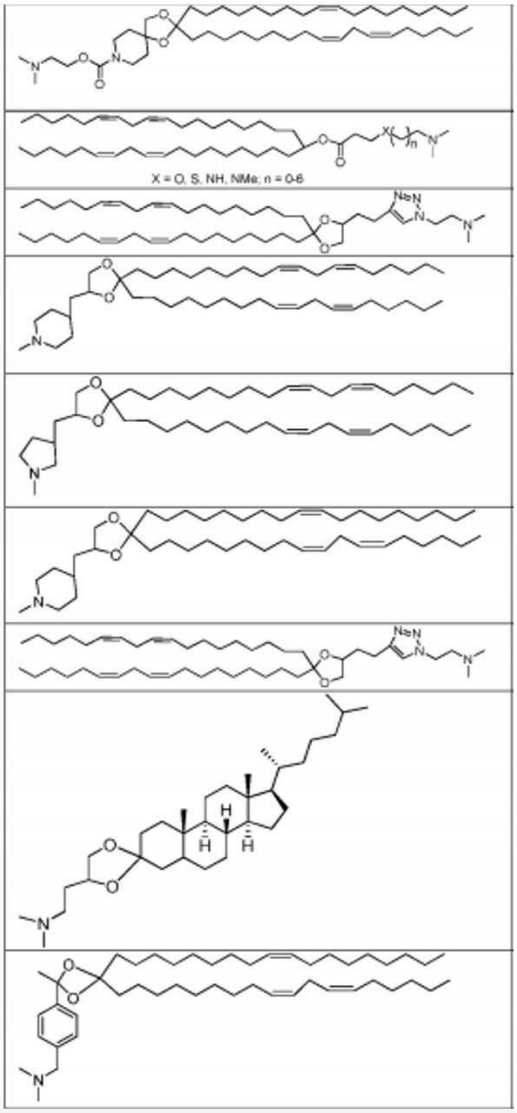
[1200]



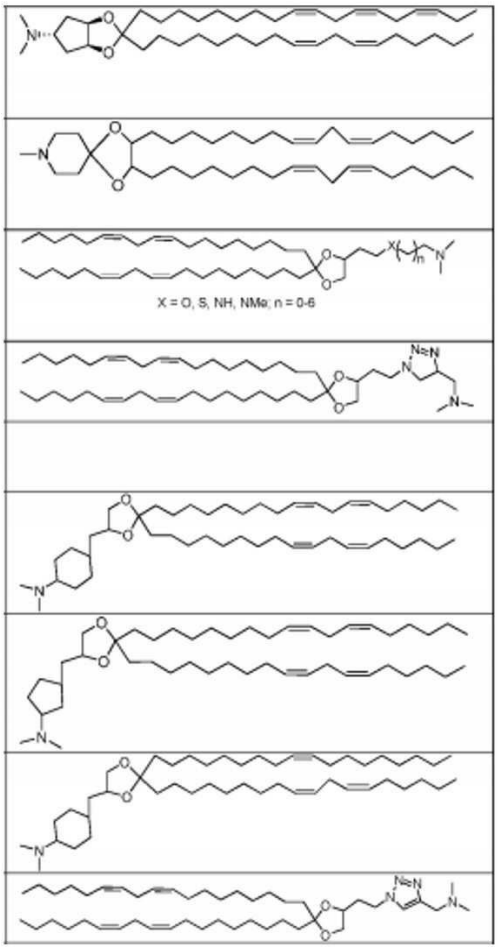


[1201]

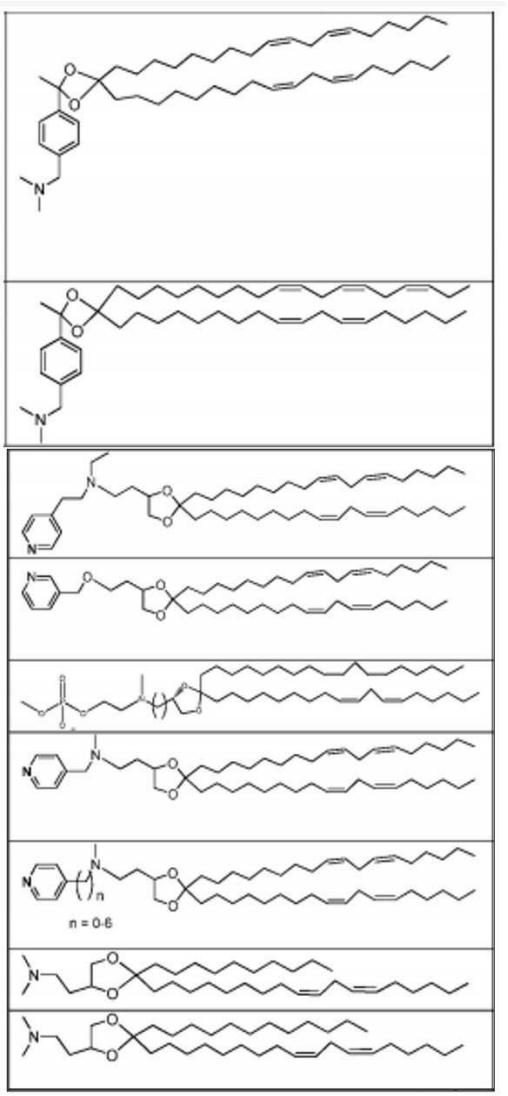
[1202]



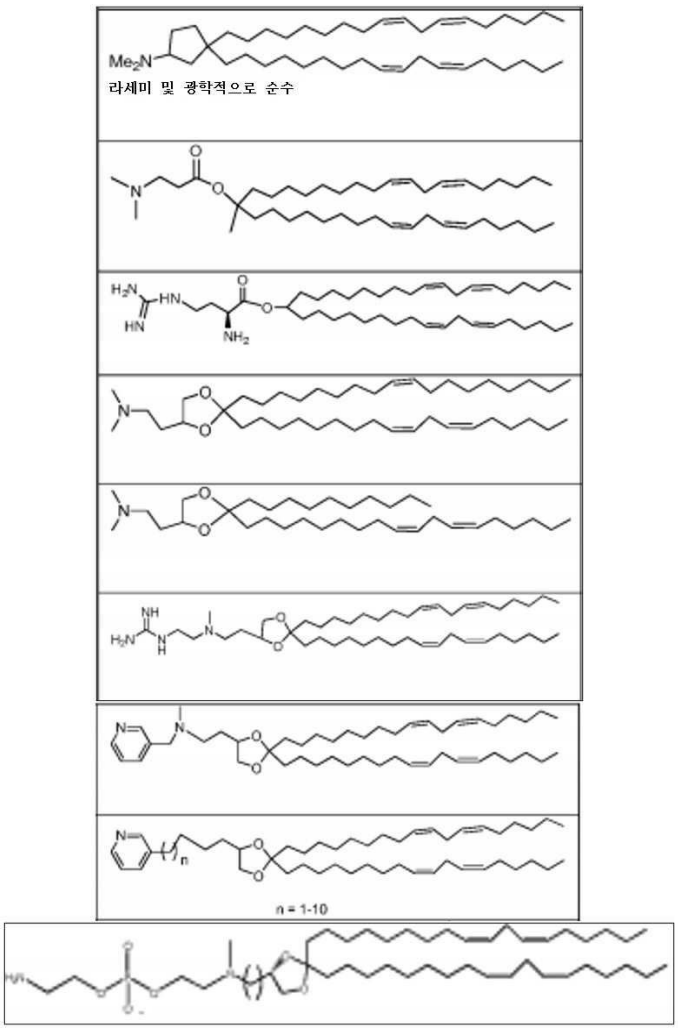
[1203]

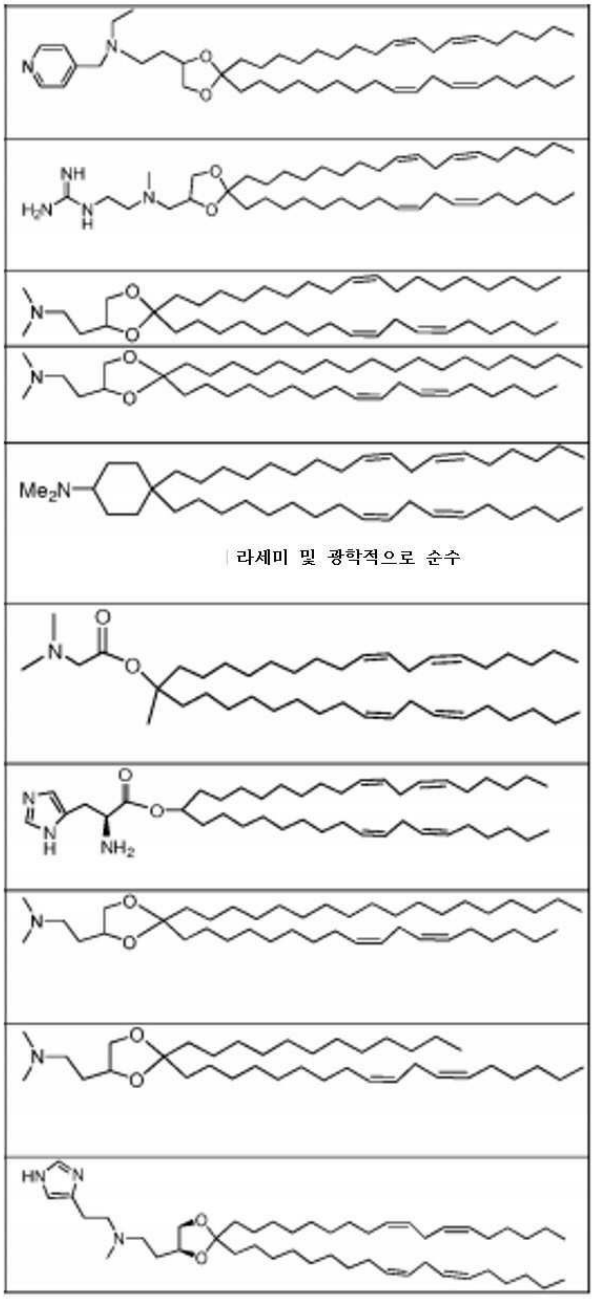


[1204]



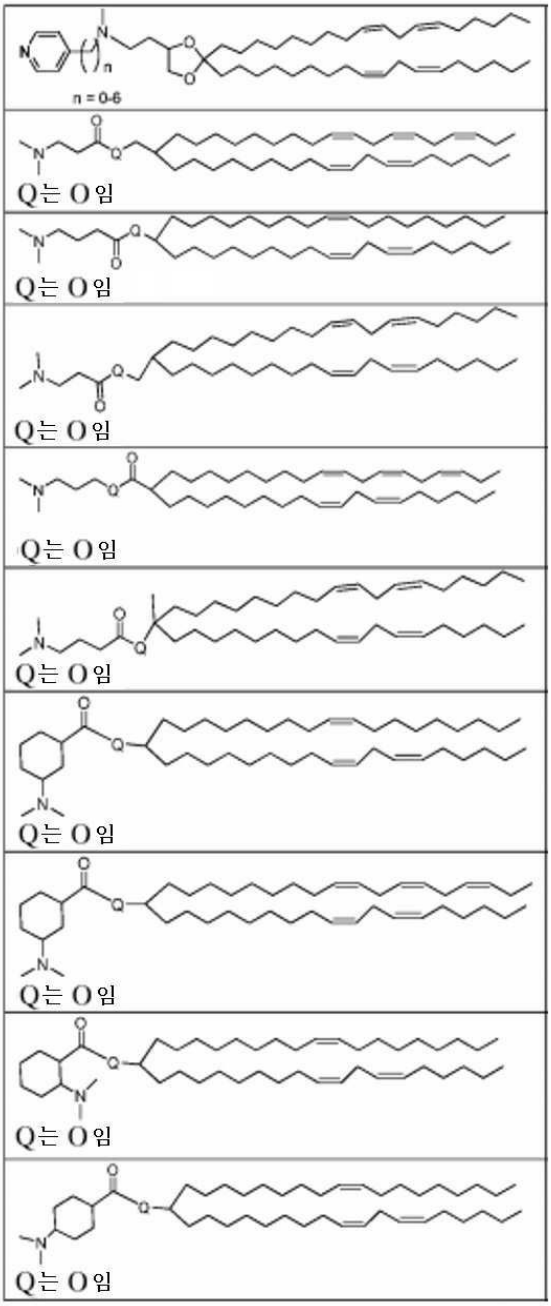
[1205]

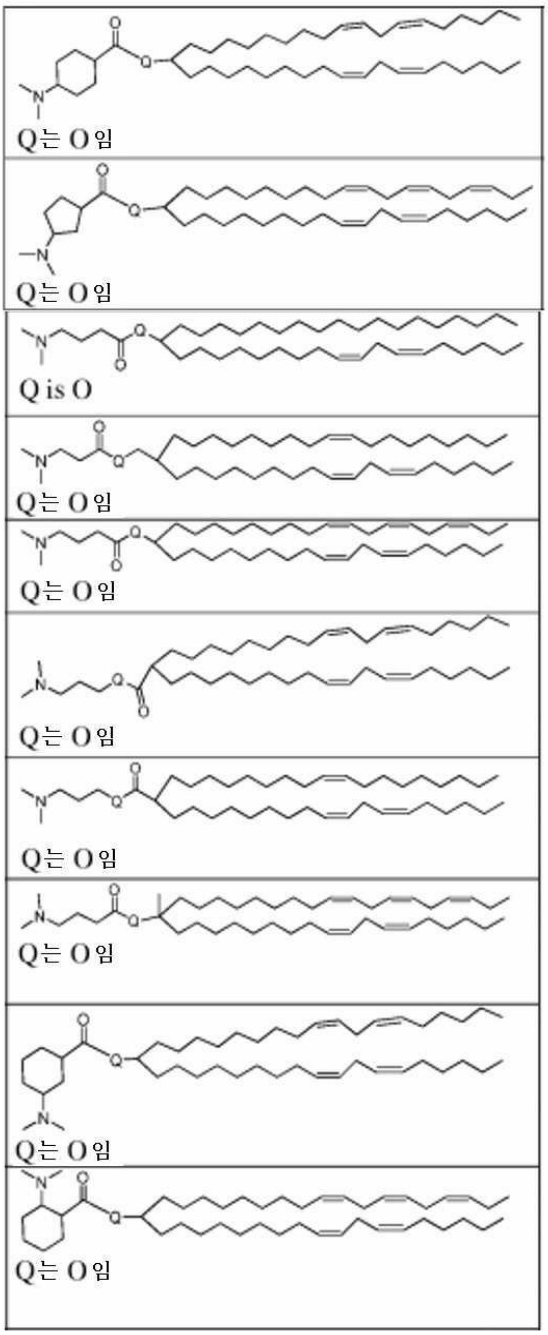




[1206]

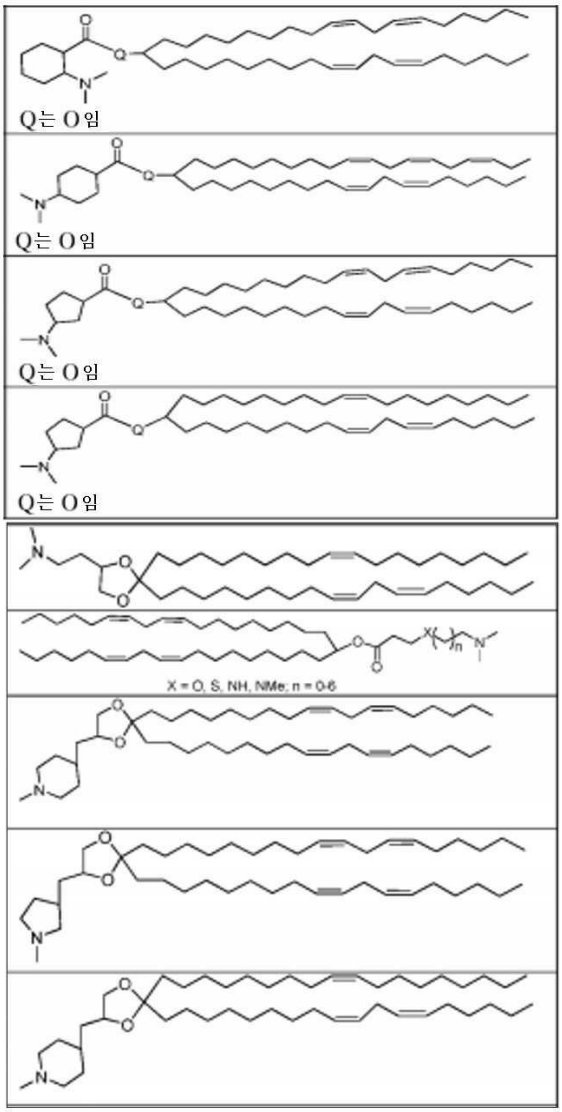
[1207]



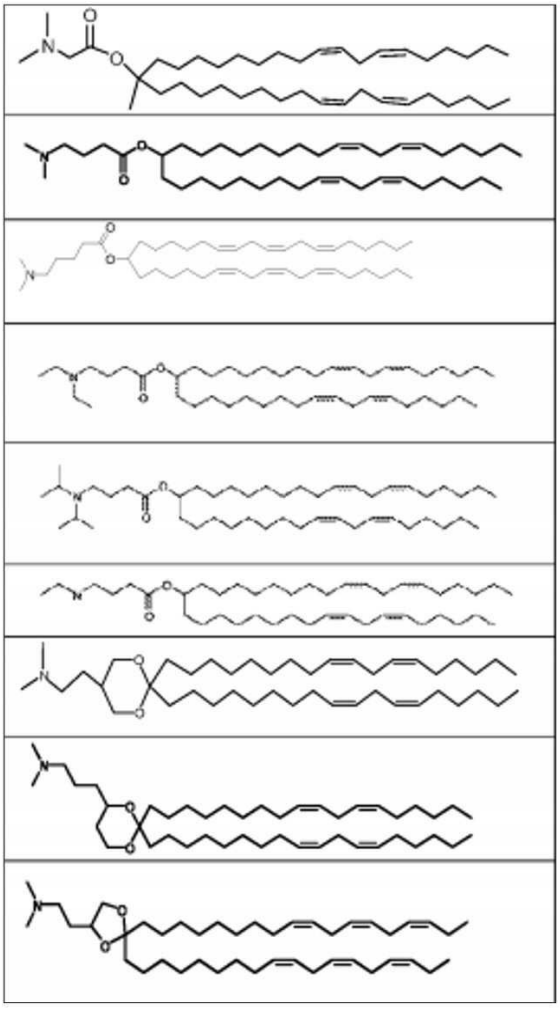


[1208]

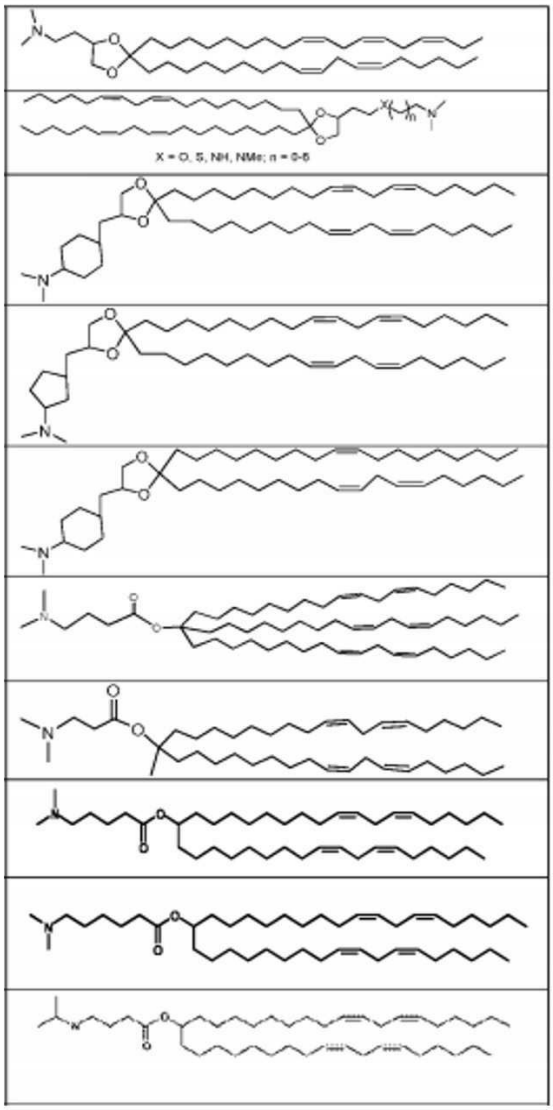
[1209]



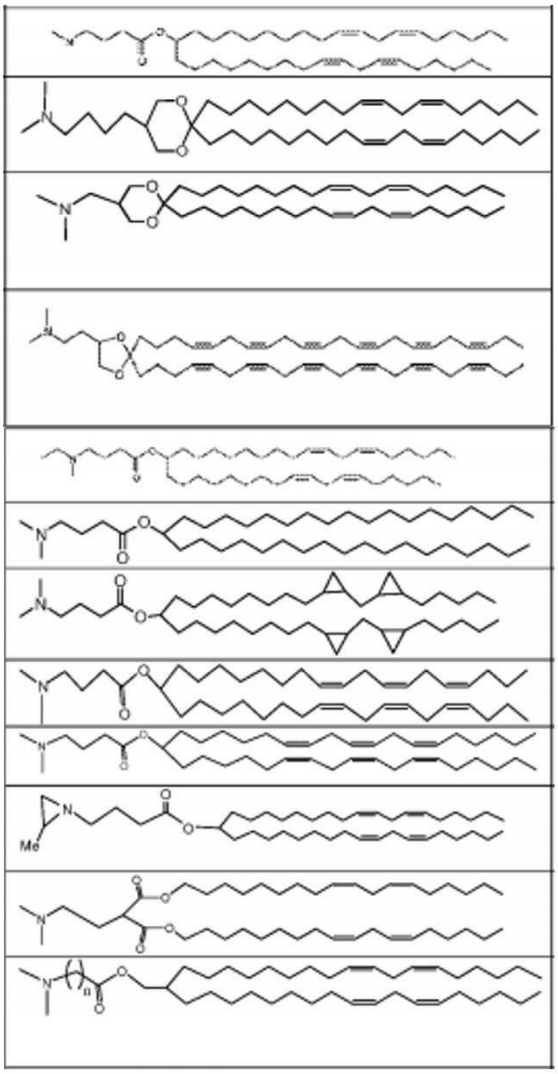
[1210]

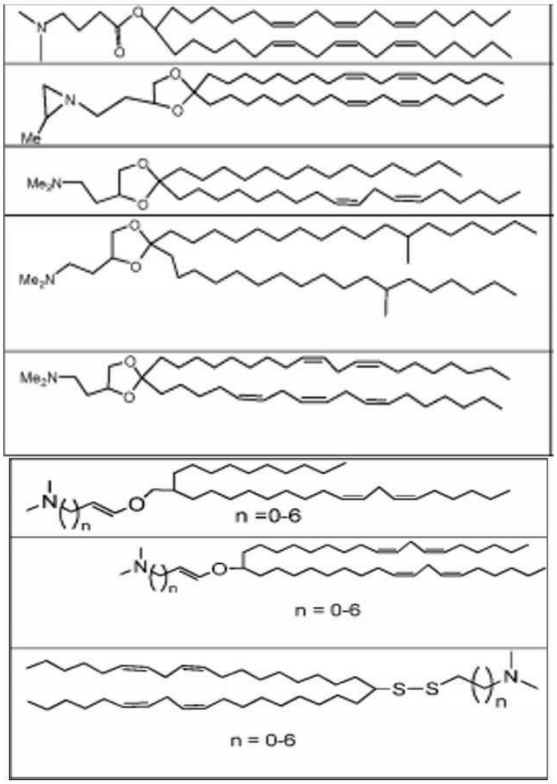


[1211]



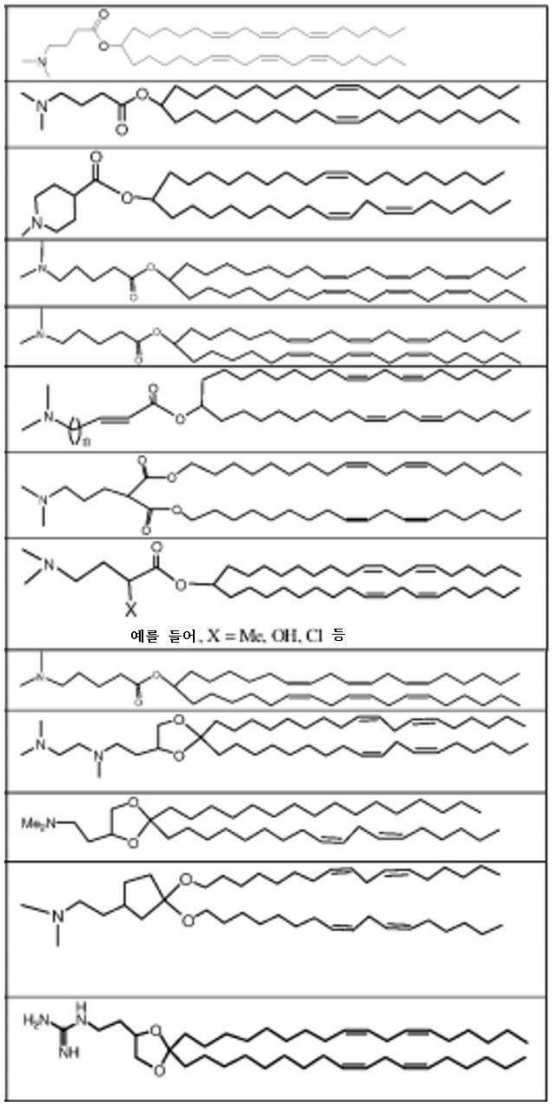
[1212]



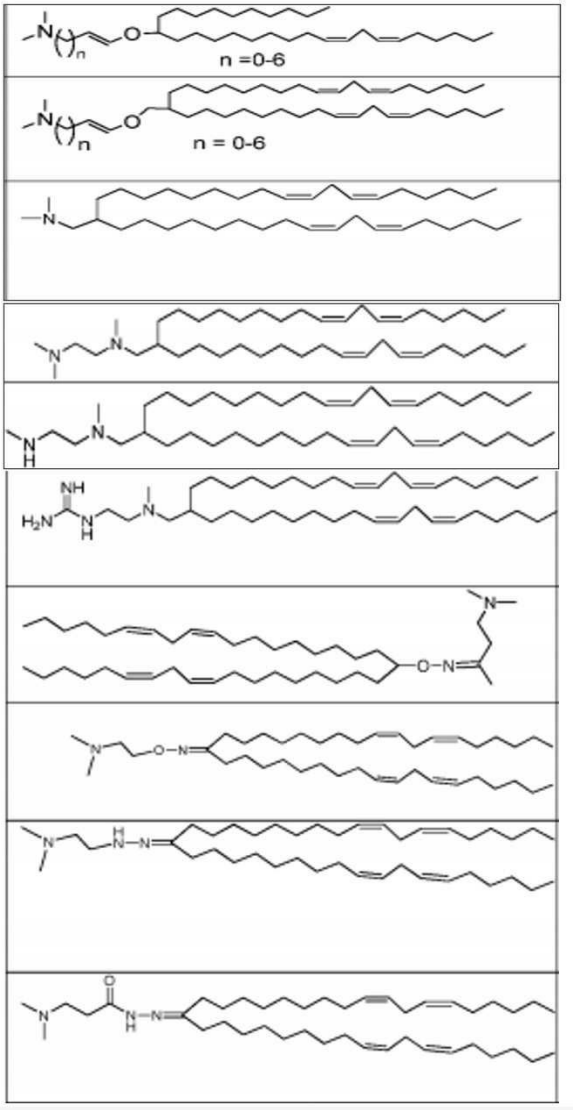


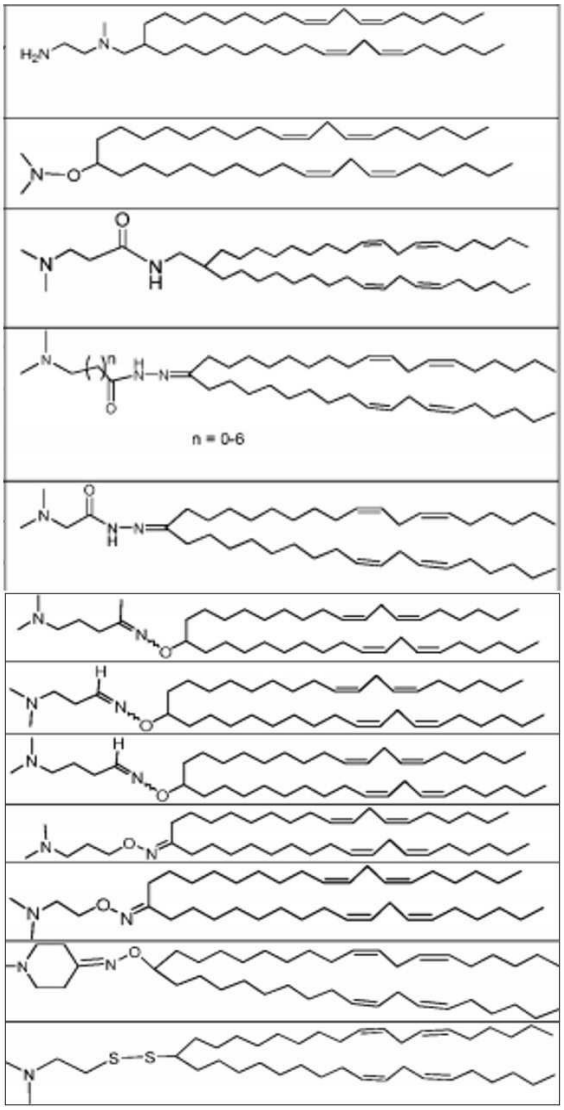
[1213]

[1214]

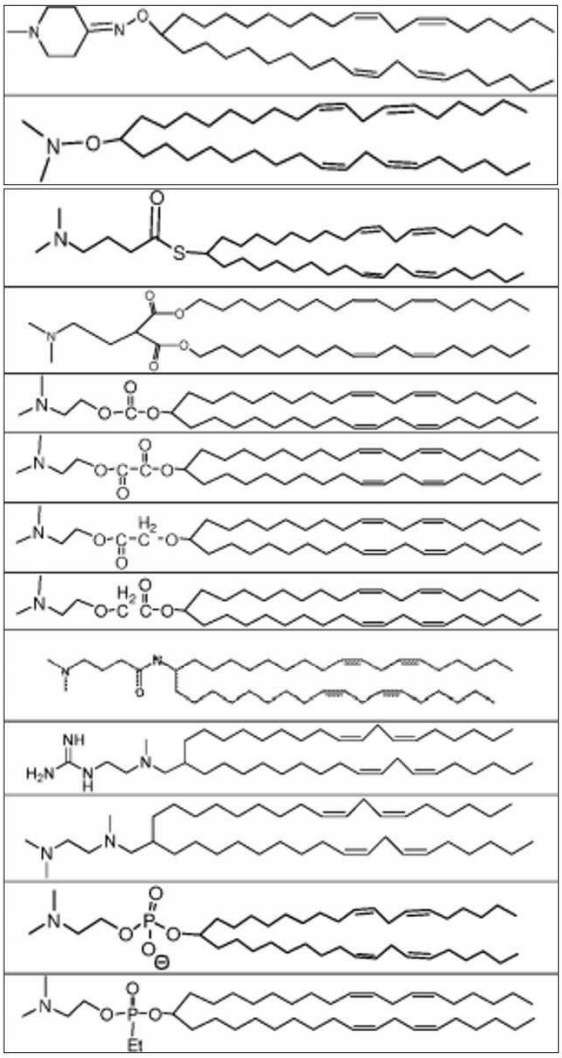


[1215]

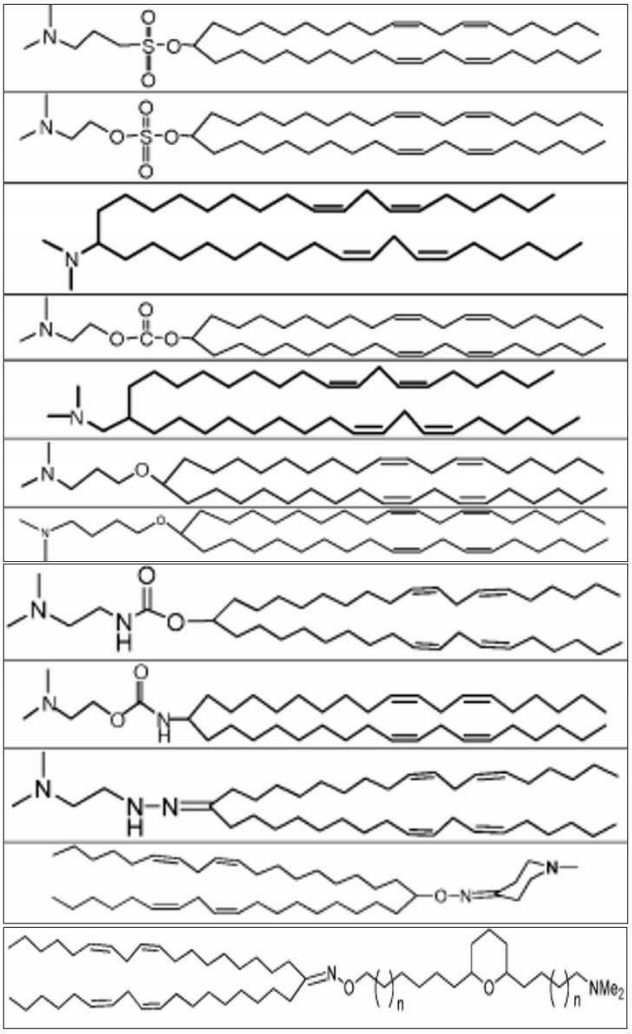




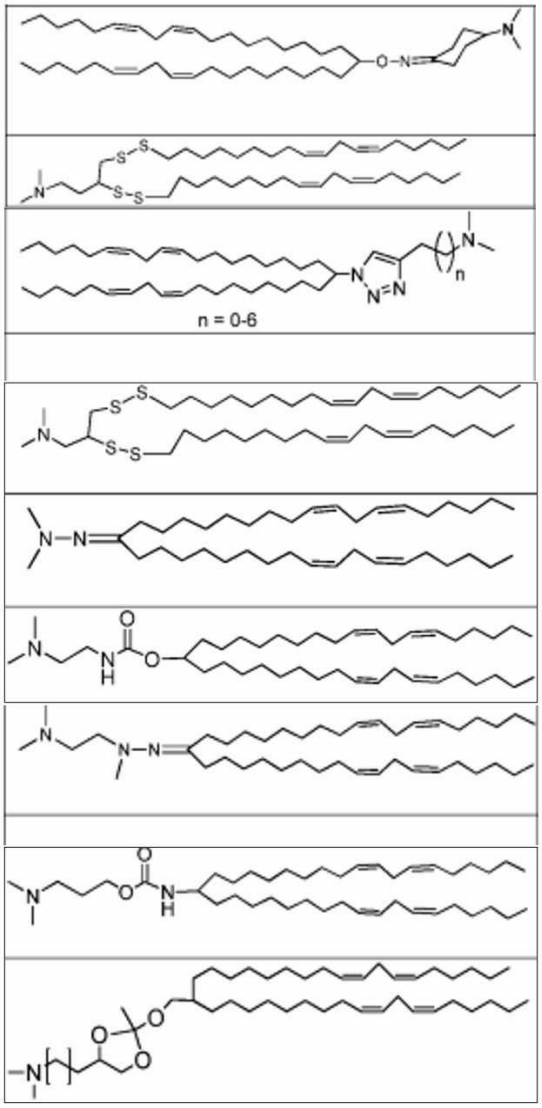
[1216]



[1217]

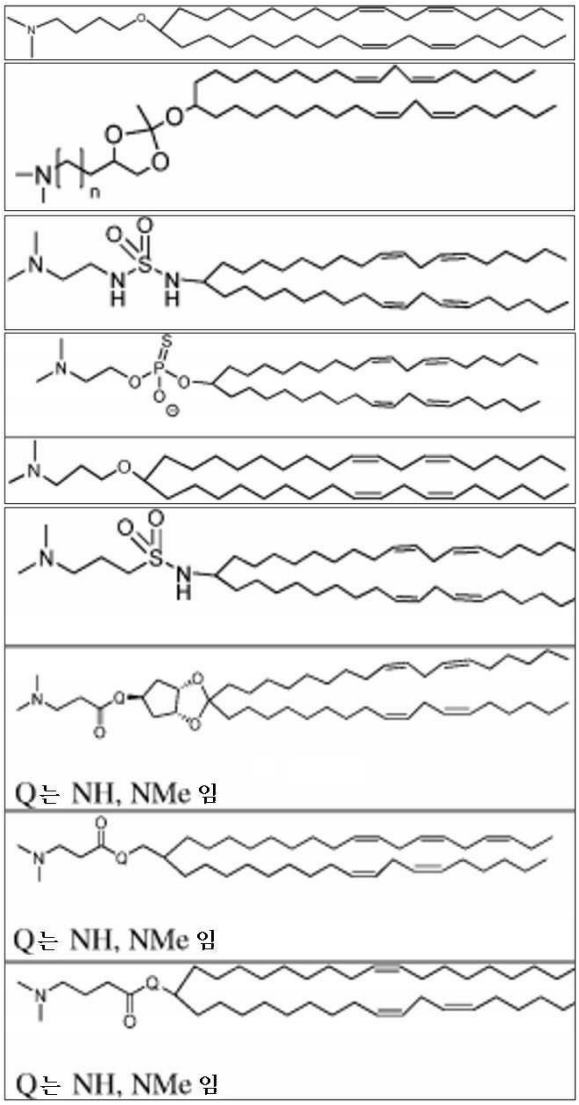


[1218]

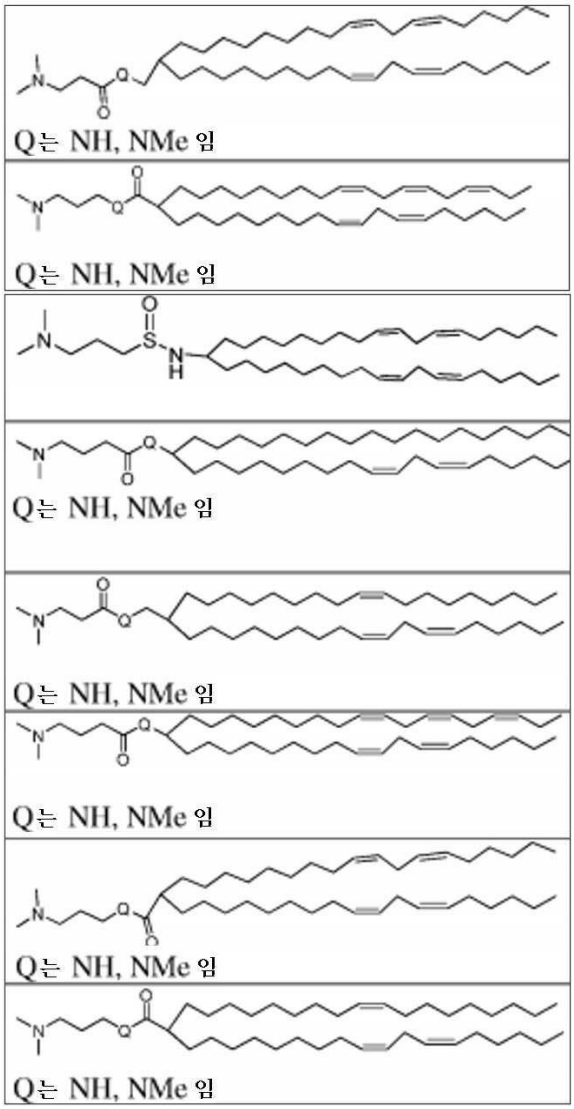


[1219]

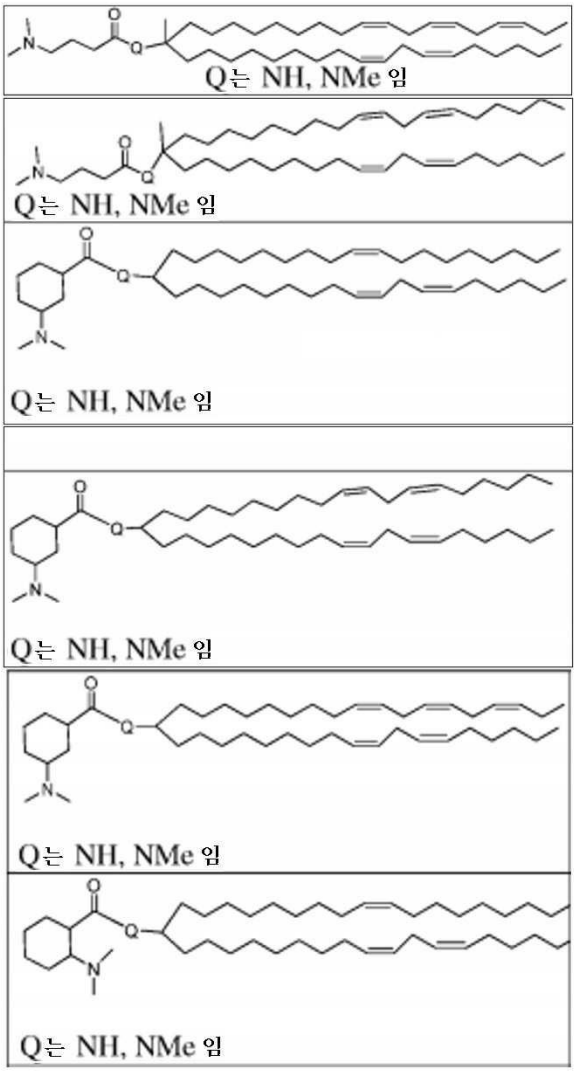
[1220]

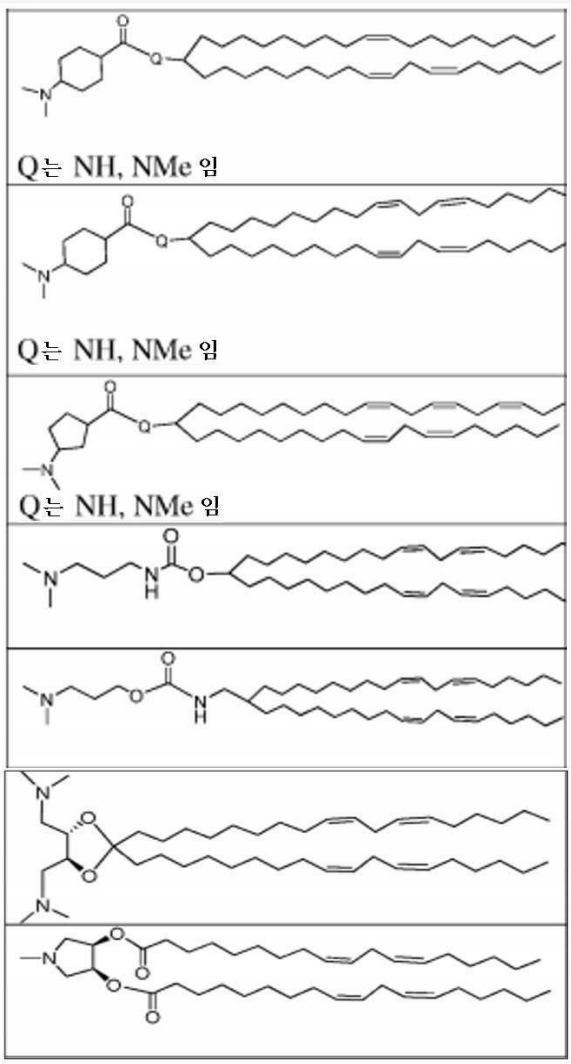


[1221]


Q는 NH, NMe 임
Q는 NH, NMe 임
Q는 NH, NMe 임
Q는 NH, NMe 임
Q는 NH, NMe 임
Q는 NH, NMe 임
Q는 NH, NMe 임

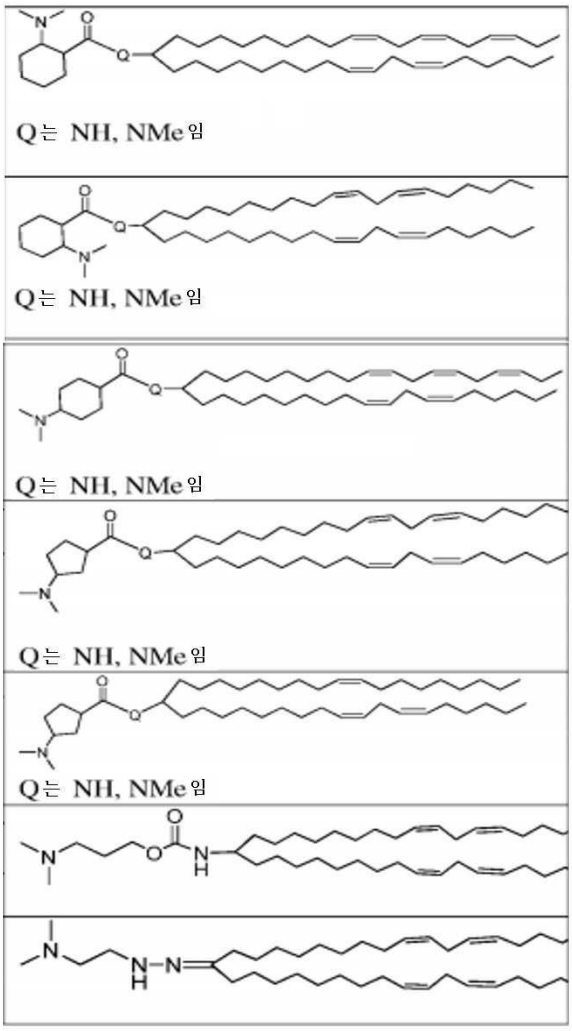
[1222]



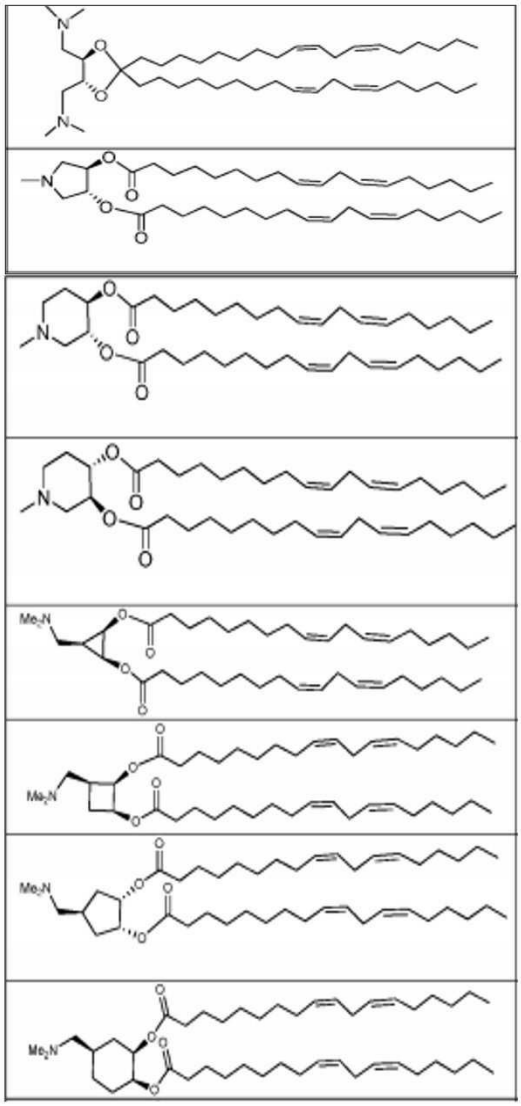


[1223]

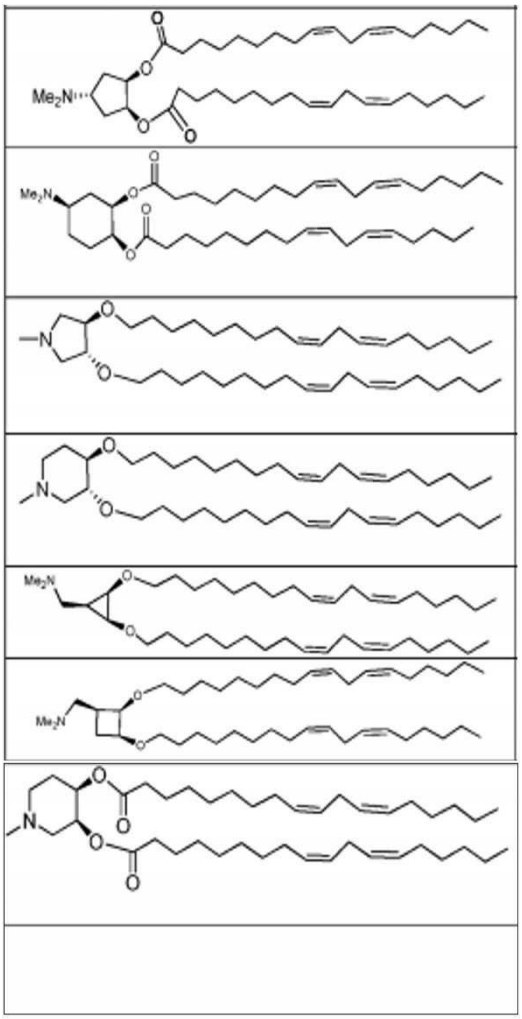
[1224]



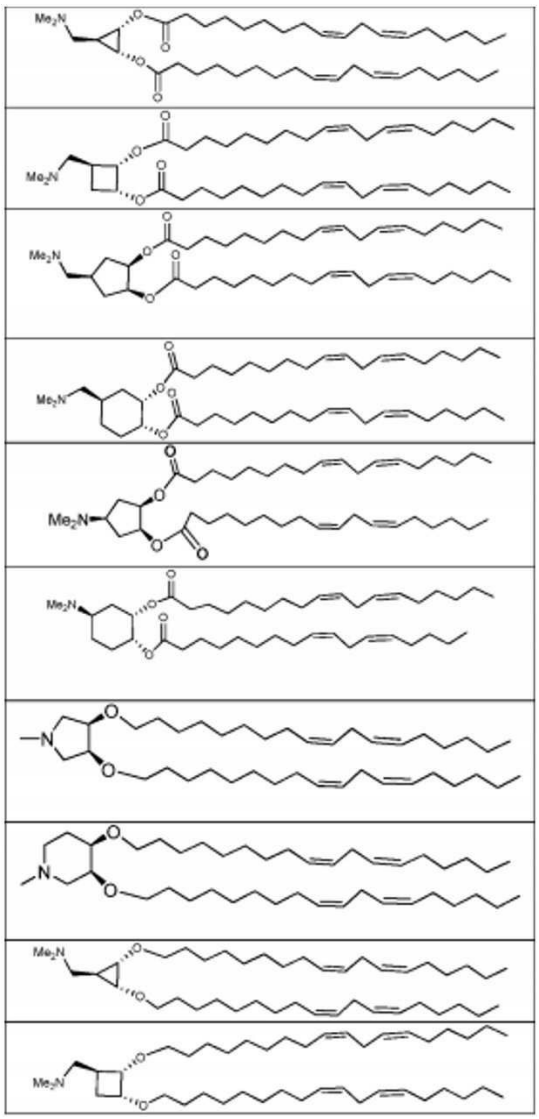
[1225]

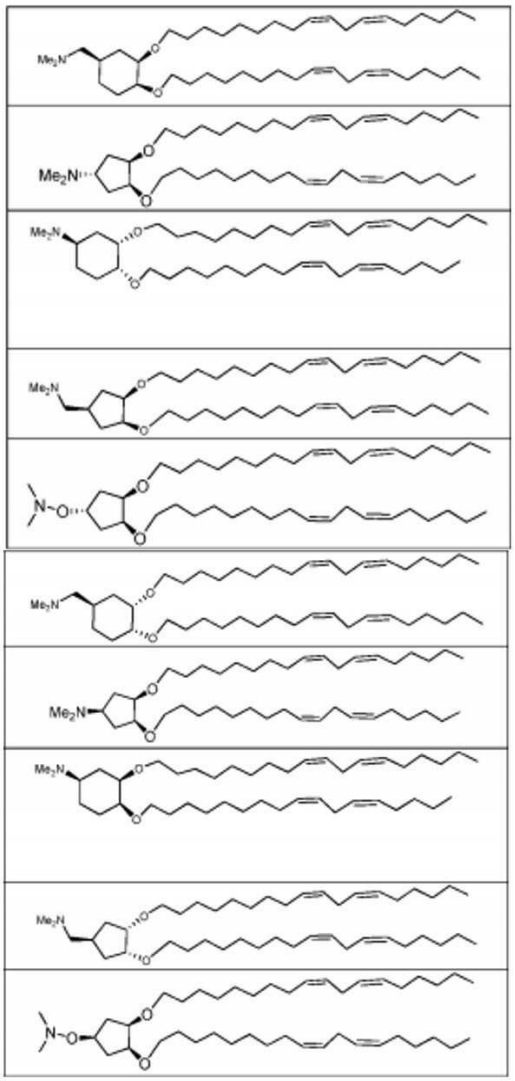


[1226]



[1227]

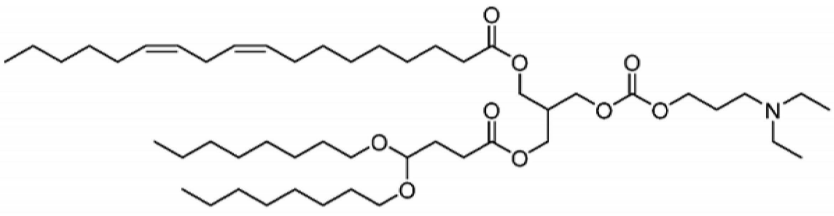




[1228]

[1229] 일부 실시형태에서, 전달 비히클은 지질 A, 지질 B, 지질 C 및/또는 지질 D를 포함한다. 일부 실시형태에서, 지질 A, 지질 B, 지질 C 및/또는 지질 D의 포함은 캡슐화 및/또는 엔도솜 탈출을 개선시킨다. 일부 실시형태에서, 지질 A, 지질 B, 지질 C 및/또는 지질 D는 국제 특허 출원 제PCT/US2017/028981호에 기재되어 있다.

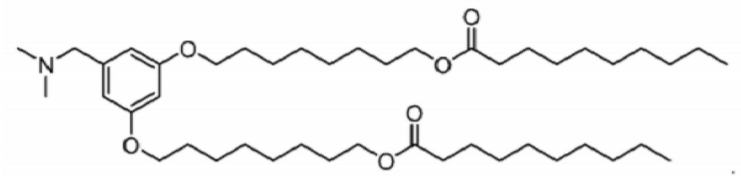
[1230] 일부 실시형태에서, 이온화 가능한 지질은 3-((4,44비스(옥틸옥시)부탄오일)옥시)-2-((((3-(다이에틸아미노)프로폭시)카보닐)옥시)메틸)프로필 (9Z,12Z)-옥타데카-9,12-다이엔오에이트라고도 불리는 (9Z,12Z)-3-((4,4- 비스(옥틸옥시)부탄오일)옥시)-2-((((3-(다이에틸아미노)프로폭시)카보닐)옥시)메틸)프로필 옥타데카9,12-다이엔오에이트인 지질 A이다. 지질 A는 하기로서 도시될 수 있다:



[1231]

[1232] 지질 A는 전문이 참조에 의해 포함된 국제 공개 제WO2015/095340호(예를 들어, pp. 84 내지 86)에 따라서 합성될 수 있다.

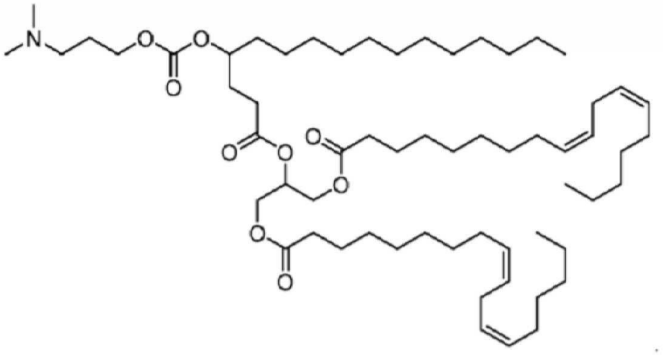
[1233] 일부 실시형태에서, 이온화 가능한 지질은 ((5-((다이메틸아미노)메틸)-1,3-페닐렌)비스(옥시))비스(옥탄-8,1-다이일)비스(데칸오에이트)인 지질 B이다. 지질 B는 하기로서 도시될 수 있다:



[1234]

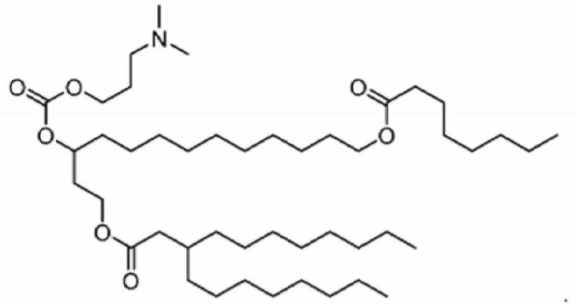
[1235] 지질 B는 전문이 참조에 의해 포함된 국제 공개 제WO2014/136086호(예를 들어, pp. 107 내지 09)에 따라서 합성될 수 있다.

[1236] 일부 실시형태에서, 이온화 가능한 지질은 2-((4-(((3- (다이메틸아미노)프로폭시)카보닐)옥시)헥사데칸오일)옥시)프로판-1,3-다이일(9Z,9'Z,12Z,12'Z)- 비스(옥타데카-9,12-다이엔오에이트)인 지질 C이다. 지질 C는 하기로서 도시될 수 있다:



[1237]

[1238] 일부 실시형태에서, 이온화 가능한 지질은 3-(((3-(다이메틸아미노)프로폭시)카보닐)옥시)- 13-(옥탄오일옥시) 트라이데실 3-옥틸운데칸오에이트인 지질 D이다. 지질 D는 하기로서 도시될 수 있다:



[1239]

[1240] 지질 C 및 지질 D는 전문이 참조에 의해 포함된 국제 공개 제WO2015/095340호에 따라서 합성될 수 있다.

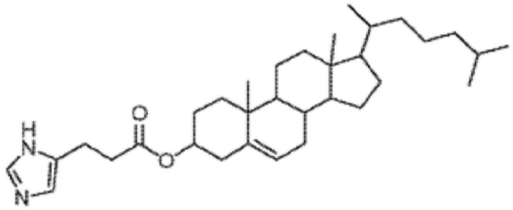
[1241] 일부 실시형태에서, 이온화 가능한 지질은 미국 출원 공보 제20190321489호에 기재되어 있다. 일부 실시형태에서, 이온화 가능한 지질은 참조에 의해 본 명세서에 포함된 국제 특허 공개 제WO 2010/053572호에 기재되어 있다. 일부 실시형태에서, 이온화 가능한 지질은 WO 2010/053572의 단락 [00225]에 기재된 C12-200이다.

[1242] 여러 이온화 가능한 지질이 문헌에 기술되어 있으며, 그 중 다수는 상업적으로 이용 가능하다. 특정 실시형태에서, 이러한 이온화 가능한 지질은 본 명세서에 기재된 전달 비히클에 포함된다. 일부 실시형태에서, 이온화 가능한 지질 N-[1-(2,3-다이올레일옥시)프로필]-N,N,N-트라이메틸암모늄 클로라이드 또는 "DOTMA"가 사용된다(문헌[Felgner *et al.* (Proc. Nat'l Acad. Sci. 84, 7413 (1987)]; 미국 특허 제4,897,355호). DOTMA는 단독으로 제형화되거나 중성 지질, 다이올레오일포스파티딜에탄올아민 또는 "DOPE" 또는 다른 양이온성 또는 비-양이온성 지질과 지질 나노입자에 조합될 수 있다. 다른 적합한 양이온성 지질은 예를 들어, 미국 가특허 61/617,468호(출원일 2012년 5월 29일)(참조에 의해 본 명세서에 포함됨)에 기재된 바와 같은 이온화 가능한 양이온성 지질, 예를 들어, (15Z,18Z)-N,N-다이메틸-6-(9Z,12Z)-옥타데카-9,12-다이엔-1-일)테트라코사-15,18-다이엔-1-아민(HGT5000), (15Z,18Z)-N,N-다이메틸-6-((9Z,12Z)-옥타데카-9,12-다이엔-1-일)테트라코사-4,15,18-트라이엔-1-아민(HGT5001) 및 (15Z,18Z)-N,N-다이메틸-6-((9Z,12Z)-옥타데카-9,12-다이엔-1-일)테트라코사-5,15,18-트라이엔-1-아민(HGT5002), C12-200(WO 2010/053572에 기재됨), 2-(2,2-다이((9Z,12Z)-옥타데카-9,12-다이엔-1-일)-1,3-다이옥솔란-4-일)-N,N-다이메틸에탄아민(DLinKC2-DMA))(WO 2010/042877; 문헌[Semple *et al.*, Nature

Biotech. 28:172-176 (2010)] 참조), 2-(2,2-다이((9Z,2Z)-옥타데카-9,12-다이엔-1-일)-1,3-다이옥솔란-4-일)-N,N-다이메틸에탄아민(DLin-KC2-DMA), (3S,10R,13R,17R)-10,13-다이메틸-17-((R)-6-메틸헵탄-2-일)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-테트라데카하이드로-1H-사이클로펜타[a]페난트렌-3-일 3-(1H-이미다졸-4-일)프로판오에이트(ICE), (15Z,18Z)-N,N-다이메틸-6-(9Z,12Z)-옥타데카-9,12-다이엔-1-일)테트라코사-15,18-다이엔-1-아민(HGT5000), (15Z,18Z)-N,N-다이메틸-6-((9Z,12Z)-옥타데카-9,12-다이엔-1-일)테트라코사-4,15,18-트라이엔-1-아민(HGT5001), (15Z,18 Z)-N,N-다이메틸-6-((9Z,12Z)-옥타데카-9,12-다이엔-1-일)테트라코사-5,15,18-트라이엔-1-아민(HGT5002), 5-카복시스퍼밀글리신-다이옥타데실아마이드(DOGS), 2,3-다이올레일옥시-N-[2(스퍼민-카복스아미도)에틸]-N,N-다이메틸-1-프로판아미늄(DOSPA)(문헌[Behr *et al.* Proc. Nat.'l Acad. Sci. 86, 6982 (1989)]; 미국 특허 제5,171,678호; 제5,334,761호), 1,2-다이올레오일-3-다이메틸암모늄-프로판(DODAP), 1,2-다이올레일다이올레오일-3-트라이메틸암모늄-프로판 또는 (DOTAP)을 포함한다. 고려되는 이온화 가능한 지질은 또한 1,2-다이스트카릴옥시-N,N-다이메틸-3-아미노프로판(DSDMA), 1,2-다이올레일옥시-N,N-다이메틸-3-아미노프로판(DODMA), 1,2-다이리놀레일옥시-N,N-다이메틸-3-아미노프로판(DLinDMA), 1,2-다이리놀렌일옥시-N,N-다이메틸-3-아미노프로판(DLenDMA), N-다이올레일-N,N-다이메틸암모늄 클로라이드(DODAC), N,N-다이스테아릴-N,N-다이메틸암모늄 브로마이드(DDAB), N-(1,2-다이미리스틸옥시프로프-3-일)-N,N-다이메틸-N-하이드록시에틸 암모늄 브로마이드(DMRIE), 3-다이메틸아미노-2-(콜레스트-5-엔-3-베타-옥시부탄-4-옥시)-1-(시스,시스-9,12-옥타데카다이엔옥시)프로판(CLinDMA), 2-[5'-(콜레스트-5-엔-3-베타-옥시)-3'-옥사펜톡시)-3-다이메틸-1-(시스,시스-9',1-2'-옥타데카다이엔옥시)프로판(CpLinDMA), N,N-다이메틸-3,4-다이올레일옥시벤질아민(DMOBA), 1,2-N,N'-다이올레일카바밀-3-다이메틸아미노프로판(DOcarbDAP), 2,3-다이리놀레오일옥시-N,N-다이메틸프로필아민(DLinDAP), 1,2-N,N'-Di리놀레일카바밀-3-다이메틸아미노프로판(DLincarbDAP), 1,2-다이리놀레오일카바밀-3-다이메틸아미노프로판(DLinCDAP), 2,2-다이리놀레일-4-다이메틸아미노메틸-[1,3]-다이옥솔란(DLin-K-DMA), 2,2-다이리놀레일-4-다이메틸아미노에틸-[1,3]-다이옥솔란(DLin-K-XTC2-DMA) 또는 GL67 또는 이들의 혼합물을 포함한다. (문헌[Heyes, J., *et al.*, J Controlled Release 107: 276-287 (2005); Morrissey, D V., *et al.*, Nat. Biotechnol. 23(8): 1003-1007 (2005)]; PCT 공개 제WO2005/121348A1호). 전달 비히클(예를 들어, 지질 나노입자)을 제형화하기 위해서 콜레스테롤계 이온화 가능한 지질을 사용하는 것이 또한 본 발명에서 고려된다. 이러한 콜레스테롤계 이온화 가능한 지질은 단독으로 또는 다른 지질과 조합하여 사용될 수 있다. 적합한 콜레스테롤계 이온화 가능한 지질은 예를 들어, DC-콜레스테롤(N,N-다이메틸-N-에틸카복사미도콜레스테롤) 및 1,4-비스(3-N-올레일아미노-프로필)피페라진(문헌[Gao, *et al.* Biochem. Biophys. Res. Comm. 179, 280 (1991); Wolf *et al.* BioTechniques 23, 139 (1997)]; 미국 특허 제5,744,335호)를 포함한다.

[1243] 또한 양이온성 지질 예컨대, 다이알킬아미노계, 이미다졸계 및 구아니디늄계 지질이 고려된다. 예를 들어, 참조에 의해 본 명세서에 포함된 국제 출원 제PCT/US2010/058457호에 개시된 바와 같은 이온화 가능한 지질 (3S,10R, 13R, 17R)-10,13-다이메틸-17-((R)-6-메틸헵타n-2-일)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-테트라데카하이드로-1H-사이클로펜타[a]페난트렌-3-일 3-(1H-이미다졸-4-일)프로판오에이트(ICE)의 사용이 또한 고려된다.

[1244] 또한 이온화 가능한 지질, 예컨대, 다이알킬아미노계, 이미다졸계 및 구아니디늄계 지질이 고려된다. 예를 들어, 특정 실시형태는 1종 이상의 이미다졸계 이온화 가능한 지질, 예를 들어, 이미다졸 콜레스테롤 에스터 또는 "ICE" 지질, 하기 구조식 (XIII)에 나타낸 바와 같은 (3S, 10R, 13R, 17R)-10, 13-다이메틸-17-((R)-6-메틸헵탄-2-일)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-테트라데카하이드로-1H-사이클로펜타[a]페난트렌-3-일 3-(1H-이미다졸-4-일)프로판오에이트를 포함하는 조성물에 관한 것이다. 실시형태에서, circRNA의 전달을 위한 전달 비히클은 1종 이상의 이미다졸계 이온화 가능한 지질, 예를 들어, 이미다졸 콜레스테롤 에스터 또는 "ICE" 지질, 하기 구조식 (XIII)에 나타낸 바와 같은 (3S, 10R, 13R, 17R)-10, 13-다이메틸-17-((R)-6-메틸헵탄-2-일)-2, 3, 4, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17-테트라데카하이드로-1H-사이클로펜타[a]페난트렌-3-일 3-(1H-이미다졸-4-일)프로판오에이트를 포함할 수 있다.



[1245] (XIII)

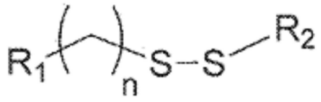
[1246] 특정 이론에 얽매이는 것은 아니지만, 이미다졸계 양이온성 지질 ICE의 융합원성(fusogenicity)은 전통적인 이

온화 가능한 지질에 비해 더 낮은 pKa를 갖는 이미다졸기에 의해 촉진되는 엔도솜 파괴와 관련이 있는 것으로 여겨진다. 엔도솜 파괴는 결국 삼투성 팽창 및 리포솜 막의 파괴를 촉진하고, 그 다음 그 안에 로딩된 핵산(들) 내용물의 표적 세포 내로의 형질주입 또는 세포내 방출이 뒤따른다.

[1247] 이미다졸계 이온화 가능한 지질은 또한 다른 이온화 가능한 지질에 비해 감소된 독성을 특징으로 한다.

[1248] 일부 실시형태에서, 이온화 가능한 지질은 미국 특허 공개 제20190314284호에 의해 기재되어 있다. 특정 실시형태에서, 이온화 가능한 지질은 구조 3, 4, 5, 6, 7, 8, 9, 또는 10(예를 들어, HGT4001, HGT4002, HGT4003, HGT4004 및/또는 HGT4005)에 기재되어 있다. 특정 실시형태에서, 하나 이상의 절단 가능한 작용기(예를 들어, 다이설파이드)는 (예를 들어, 산화성, 환원성 또는 산성 조건 하에서) 예를 들어, 친수성 작용성 헤드기를 화합물의 친유성 작용성 테일기로부터 해리시킴으로써, 하나 이상의 표적 세포의 지질 이중층에서 상 전이를 용이하게 한다. 예를 들어, 전달 비히클(예를 들어, 지질 나노입자)이 구조 3 내지 10의 지질 중 하나 이상을 포함하는 경우, 하나 이상의 표적 세포의 지질 이중층에서의 상 전이는 circRNA를 하나 이상의 표적 세포로 전달하는 것을 촉진한다.

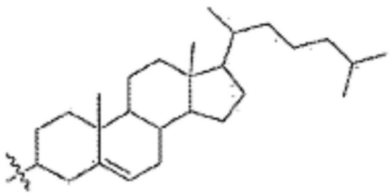
[1249] 특정 실시형태에서, 이온화 가능한 지질은 하기 구조식 (XIV)에 의해 기재된다:

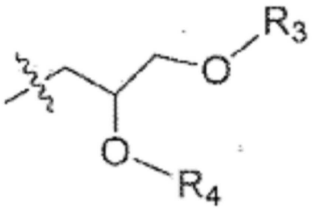
[1250]  (XIV)

[1251] 식 중,

[1252] R_1은 이미다졸, 구아니디늄, 아미노, 이민, 엔아민, 선택적으로 치환된 알킬 아미노(예를 들어, 알킬 아미노, 예컨대, 다이메틸아미노) 및 피리딜을 포함하거나 이들로 이루어진 군으로부터 선택되고;

[1253] R_2는 구조식 XV 및 구조식 XVI으로 이루어진 군으로부터 선택되고;

[1254]  XV

[1255]  XVI

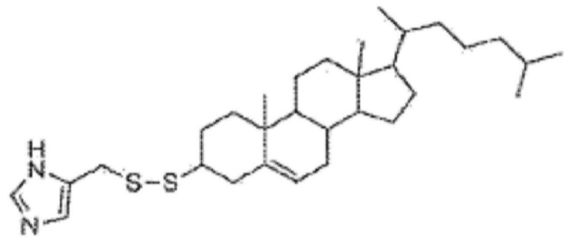
[1256] R_3 및 R_4는 각각 독립적으로 선택적으로 치환된, 가변적인 포화 또는 불포화 C_6-C_{20} 알킬 및 선택적으로 치환된, 가변적인 포화 또는 불포화 C_6-C_{20} 아실로 이루어진 군으로부터 선택되고; n은 0 또는 임의의 양의 정수(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 또는 그 초과)이다. 특정 실시형태에서, R_3 및 R_4는 각각 선택적으로 치환된, 다중불포화 C_{18} 알킬이지만, 다른 실시형태에서 R_3 및 R_4는 각각 비치환된, 다중불포화 C_{18} 알킬이다. 특정 실시형태에서, R_3 및 R_4 중 하나 이상은 (9Z,12Z)-옥타데카-9,12-다이엔이다.

[1257] 본 명세서에는 또한 구조식 XIV의 화합물을 포함하는 약제학적 조성물이 개시되며, R_1은 이미다졸, 구아니디늄, 아미노, 이민, 엔아민, 선택적으로 치환된 알킬 아미노(예를 들어, 알킬 아미노, 예컨대, 다이메틸아미노) 및 피리딜을 포함하거나 이들로 이루어진 군으로부터 선택되고; R_2는 구조식 XV이고; n은 0 또는 임의의 양의 정수이다. 추가로 본 명세서에는 구조식 XIV의 화합물을 포함하는 약제학적 조성물이 개시되며, R_1은 이미다졸, 구아

니디늄, 아미노, 이민, 엔아민, 선택적으로 치환된 알킬 아미노(예를 들어, 알킬 아미노, 예컨대, 다이메틸아미노) 및 피리딜을 포함하거나 이들로 이루어진 군으로부터 선택되고; R_2는 구조식 XVI이고; R_3 및 R_4는 각각 독립적으로 선택적으로 치환된, 가변적인 포화 또는 불포화 C_6-C_{20} 알킬 및 선택적으로 치환된, 가변적인 포화 또는 불포화 C_6-C_{20} 아실로 이루어진 군으로부터 선택되고; n은 0 또는 임의의 양의 정수이다. 특정 실시형태에서, R_3 및 R_4는 각각 선택적으로 치환된, 다중불포화 C_{18} 알킬이지만, 다른 실시형태에서 R_3 및 R_4는 각각 비치환된, 다중불포화 C_{18} 알킬(예를 들어, 옥타데카-9,12-다이엔)이다.

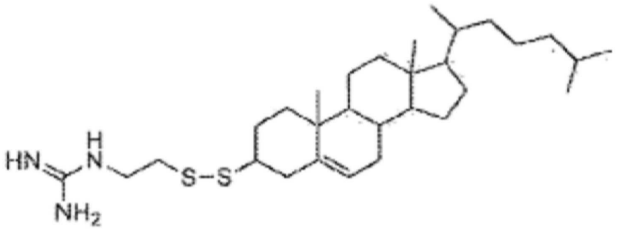
[1258] 특정 실시형태에서, R_1기 또는 헤드-기는 극성 또는 친수성 기(예를 들어, 이미다졸, 구아니디늄 및 아미노기 중 하나 이상)이고, 예를 들어, 구조식 XIV에 도시된 바와 같은, 이황화(S-S) 절단 가능한 링커 기에 의해서 R_2 지질기에 결합되어 있다. 다른 고려되는 절단 가능한 링커 기는 예를 들어, 알킬기(예를 들어, C_1 내지 C_{10} 알킬)에 결합된(예를 들어, 공유 결합된) 하나 이상의 이황화(S-S) 링커 기를 포함하는 조성물을 포함할 수 있다. 특정 실시형태에서, R1 기는 C_1-C_{20} 알킬기에 의해서 절단 가능한 링커 기에 공유 결합되거나(예를 들어, n은 1 내지 20임), 대안적으로 절단 가능한 링커 기에 직접 결합될 수 있다(예를 들어, n은 0임). 특정 실시형태에서, 이황화 링커 기는 시험관내 및/또는 생체내에서 절단 가능하다(예를 들어, 효소에 의해서 절단 가능하거나 산성 또는 환원 조건에 노출될 때 절단 가능함).

[1259] 특정 실시형태에서, 본 발명은 구조식 XVII을 갖는 화합물 5-(((10,13-다이메틸-17-(6-메틸헵타n-2-일)-2,3,4,7,8,9,10,11,12,13,14,15,16,17-테트라데카하이드로-1H-사이클로펜타[a]페난트렌-3-일)다이설파닐)메틸)-1H-이미다졸("HGT4001"이라고 지칭됨)에 관한 것이다.



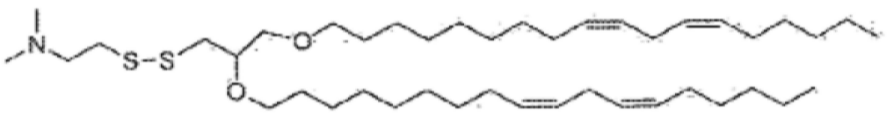
[1260] XVII

[1261] 특정 실시형태에서, 본 발명은 구조식 XVIII을 갖는 화합물 1-(2-(((3S,10R,13R)-10,13-다이메틸-17-((R)-6-메틸헵타n-2-일)-2,3,4,7,8,9,10,11,12,13,14,15,16,17-테트라데카하이드로-1H-사이클로펜타[a]페난트렌-3-일)다이설파닐)에틸)구아니딘("HGT4002"라고 지칭됨)에 관한 것이다.



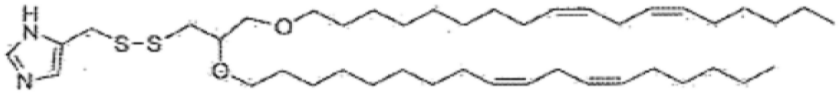
[1262] XVIII

[1263] 특정 실시형태에서, 본 발명은 구조식 XIX를 갖는 화합물 2-((2,3-비스((9Z,12Z)-옥타데카-9,12-다이엔-1-일옥시)프로필)다이설파닐)-N,N-다이메틸에탄아민("HGT4003"이라고 지칭됨)에 관한 것이다.

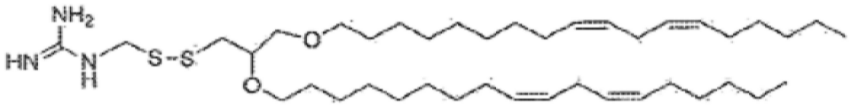


[1264] XIX

[1265] 다른 실시형태에서, 본 발명은 구조식 XX의 구조를 갖는 화합물 5-(((2,3-비스((9Z,12Z)-옥타데카-9,12-다이엔-1-일옥시)프로필)다이설파닐)메틸)-1H-이미다졸("HGT4004"라고 지칭됨)에 관한 것이다.

[1266]  XX

[1267] 추가의 다른 실시형태에서, 본 발명은 구조식 XXI을 갖는 화합물 1-(((2,3-비스((9Z,12Z)-옥타데카-9,12-다이엔-1-일옥시)프로필)다이설파닐)메틸)구아니딘("HGT4005"라고 지칭됨)에 관한 것이다.

[1268]  XXI

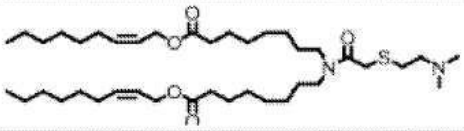
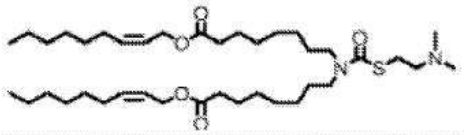
[1269] 특정 실시형태에서, 구조 3 내지 10으로 기재된 화합물은 이온화 가능한 지질이다.

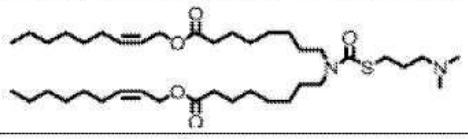
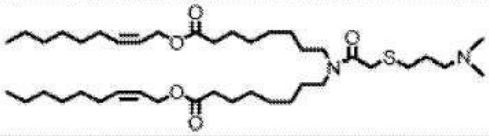
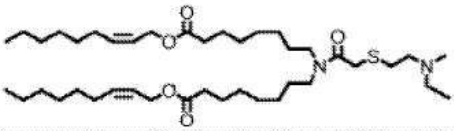
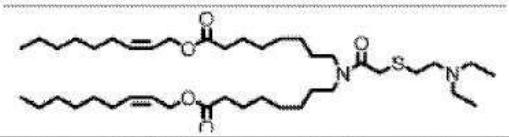
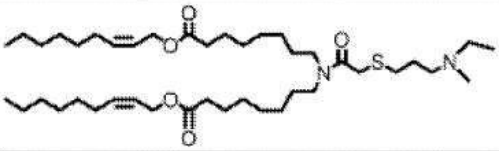
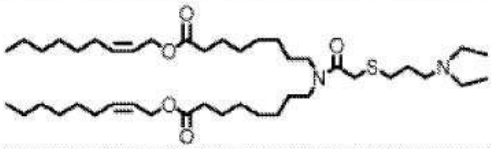
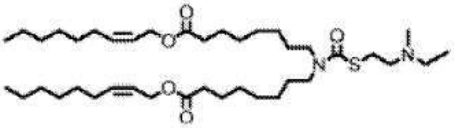
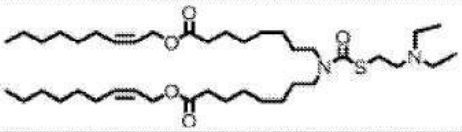
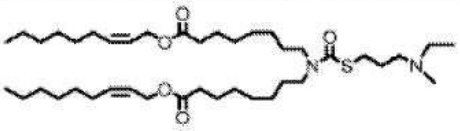
[1270] 화합물 및 특히 구조 3 내지 8로서 기재된 이미다졸계 화합물(예를 들어, HGT4001 및 HGT4004)은 특히 전통적인 이온화 가능한 지질과 비교하여 감소된 독성을 특징으로 한다. 일부 실시형태에서, 본 명세서에 기재된 전달 비히클은 이러한 약제학적 또는 리포솜 조성물에서 다른 더 독성인 이온화 가능한 지질이 감소되거나 달리 제거될 수 있도록 1종 이상의 이미다졸계 이온화 가능한 지질 화합물을 포함한다.

[1271] 이온화 가능한 지질은 전문이 참조에 의해 본 명세서에 포함된 국제 특허 출원 제PCT/US2019/025246호 및 미국 특허 공개 제2017/0190661호 및 제2017/0114010호에 개시된 것을 포함한다. 이온화 가능한 지질은 하기 표 12, 13, 14 또는 15로부터 선택된 지질을 포함할 수 있다.

표 12

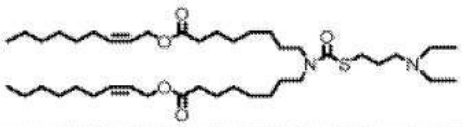
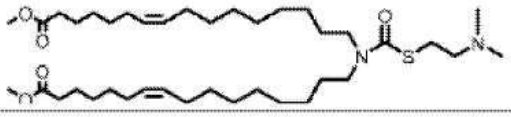
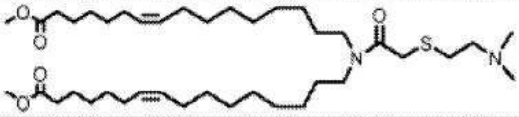
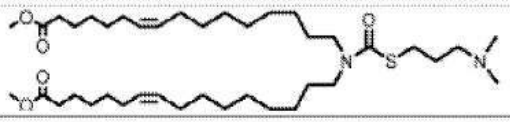
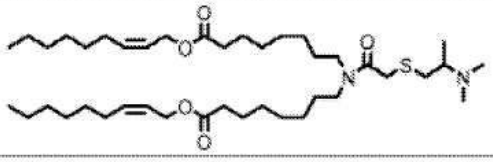
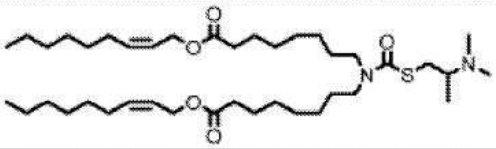
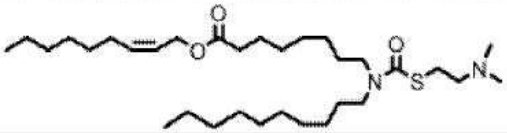
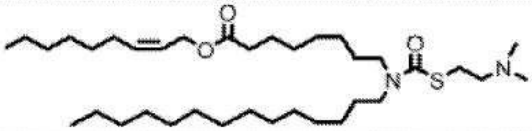
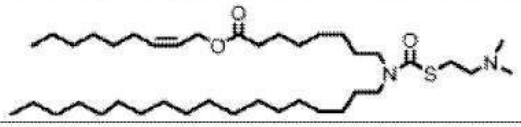
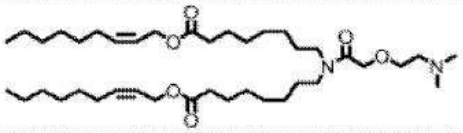
[1272]

ATX-001	
ATX-002	

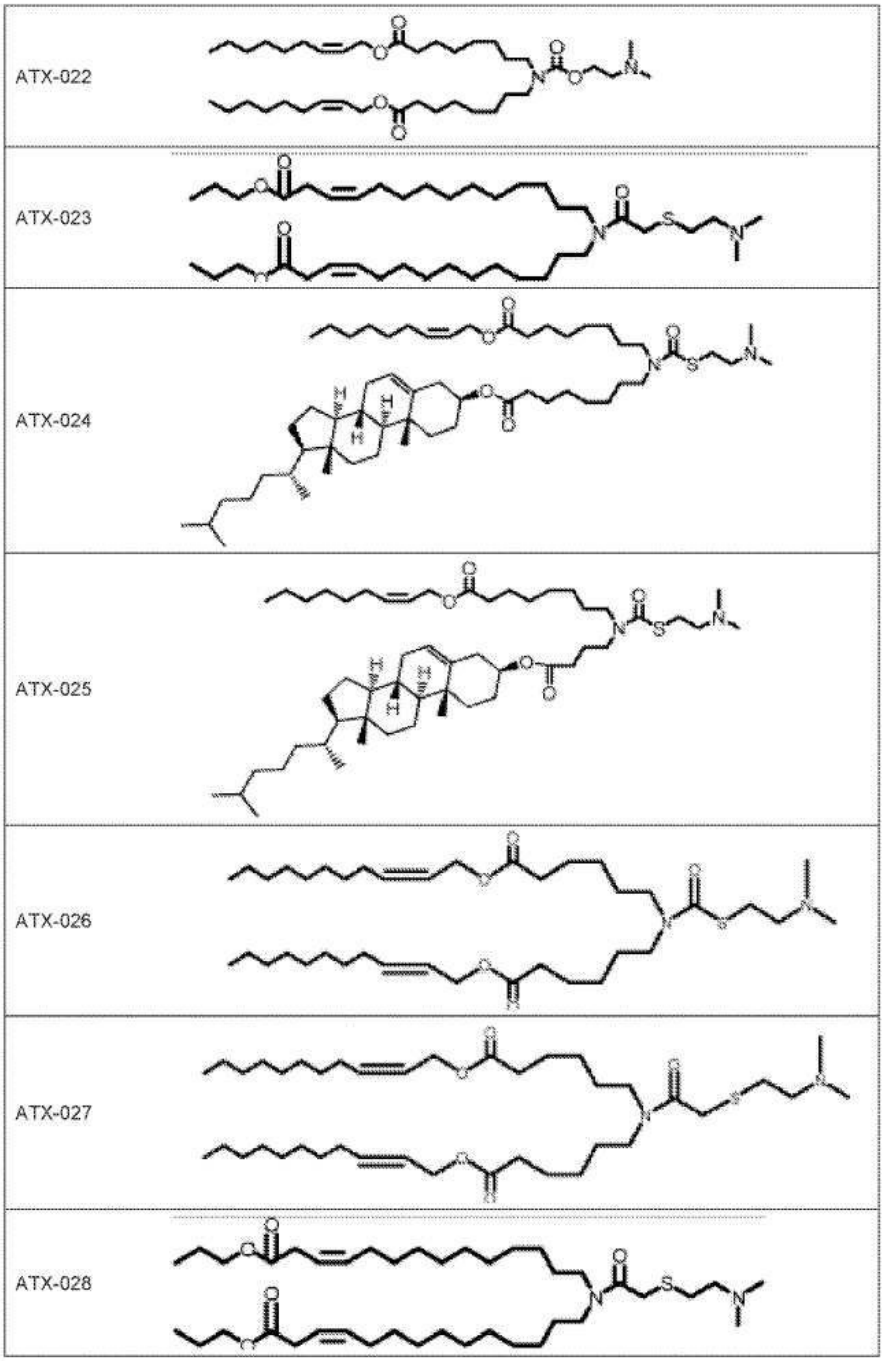
ATX-003	
ATX-004	
ATX-005	
ATX-006	
ATX-007	
ATX-008	
ATX-009	
ATX-010	
ATX-011	

[1273]

[1274]

ATX-012	
ATX-013	
ATX-014	
ATX-015	
ATX-016	
ATX-017	
ATX-018	
ATX-019	
ATX-020	
ATX-021	

[1275]



[1276]

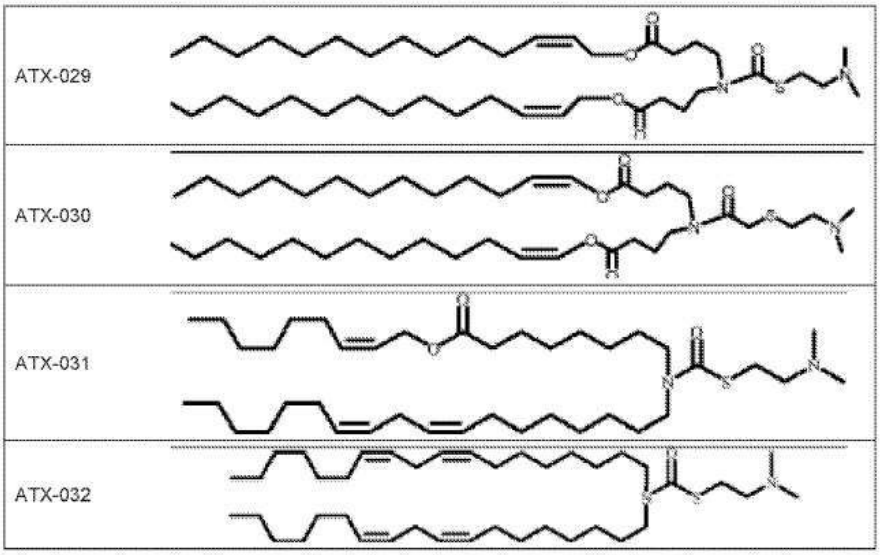
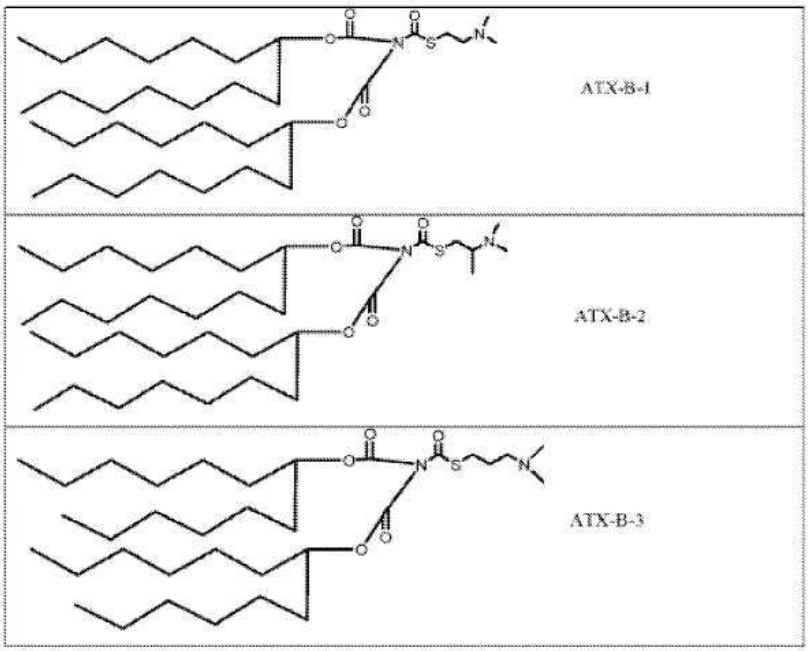
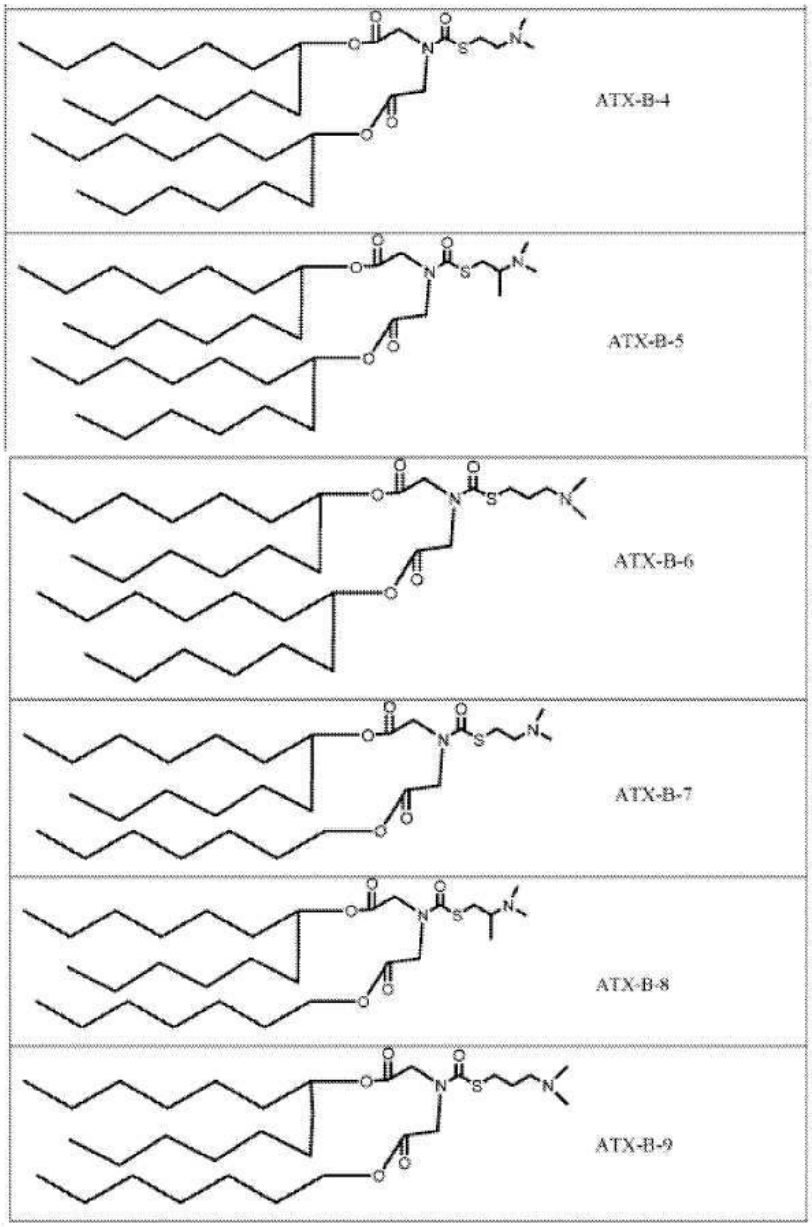


표 13

[1277]



[1278]



[1279]

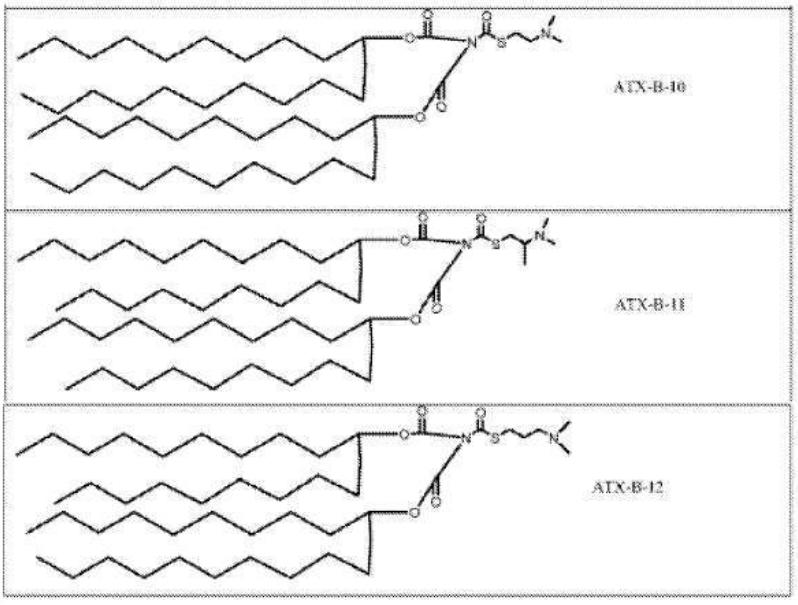
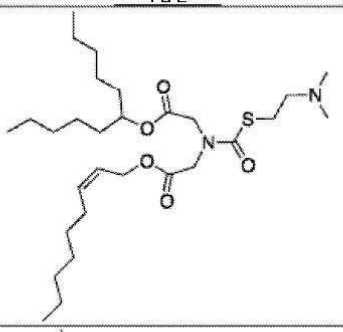
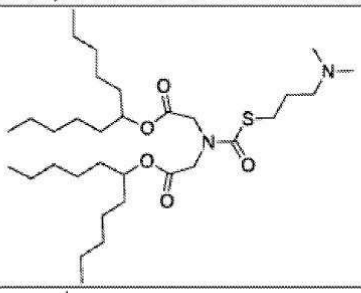
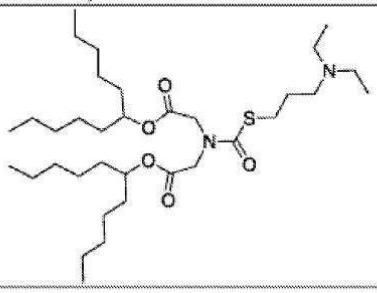
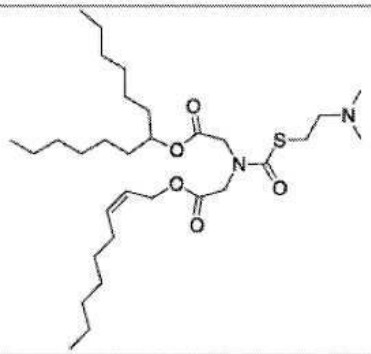
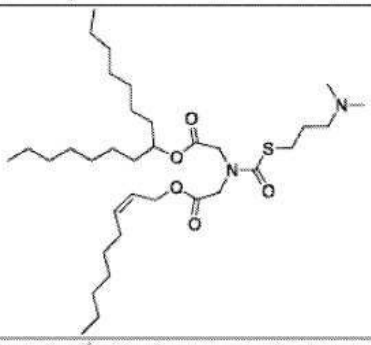
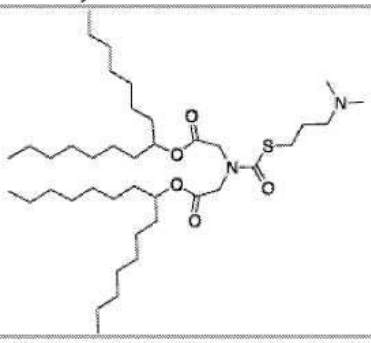


표 14

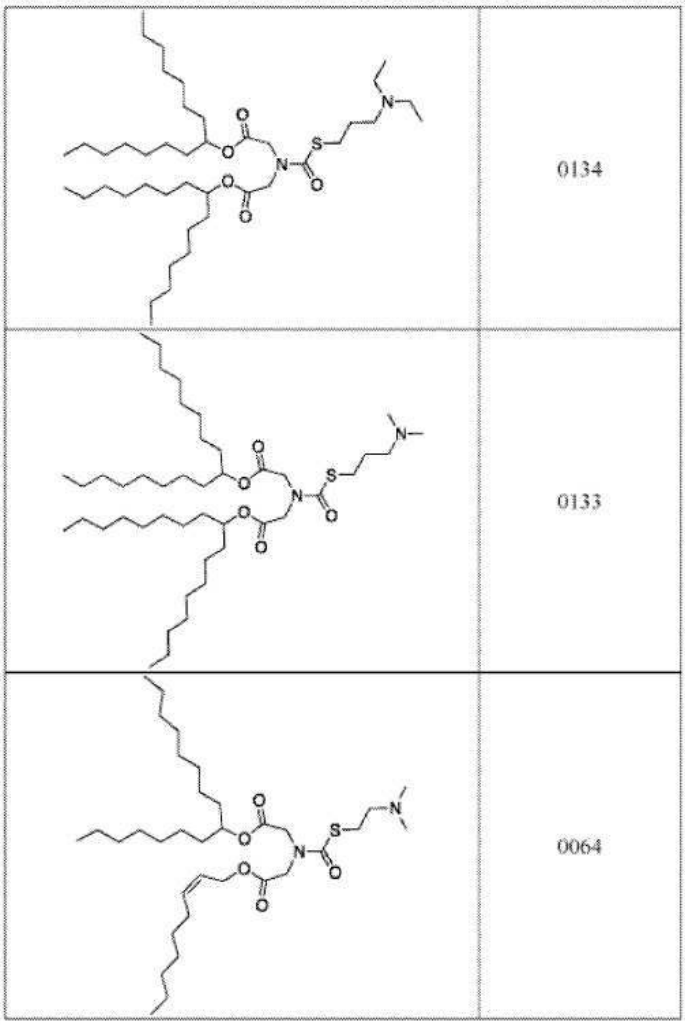
화합물	ATX-#
	0063
	0130
	0131

[1280]

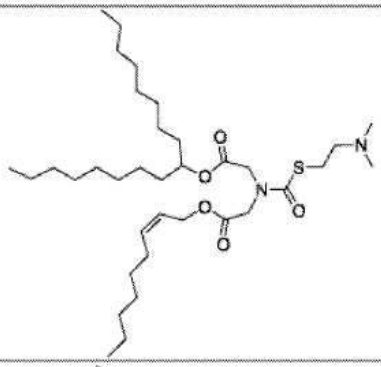
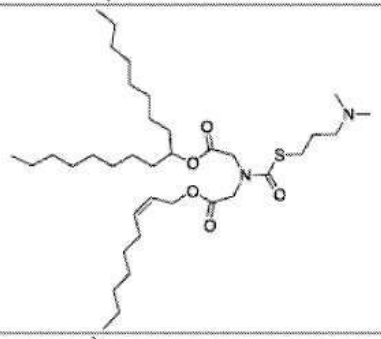
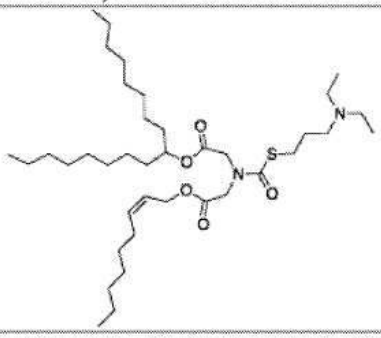
[1281]

	0044
	0111
	0132

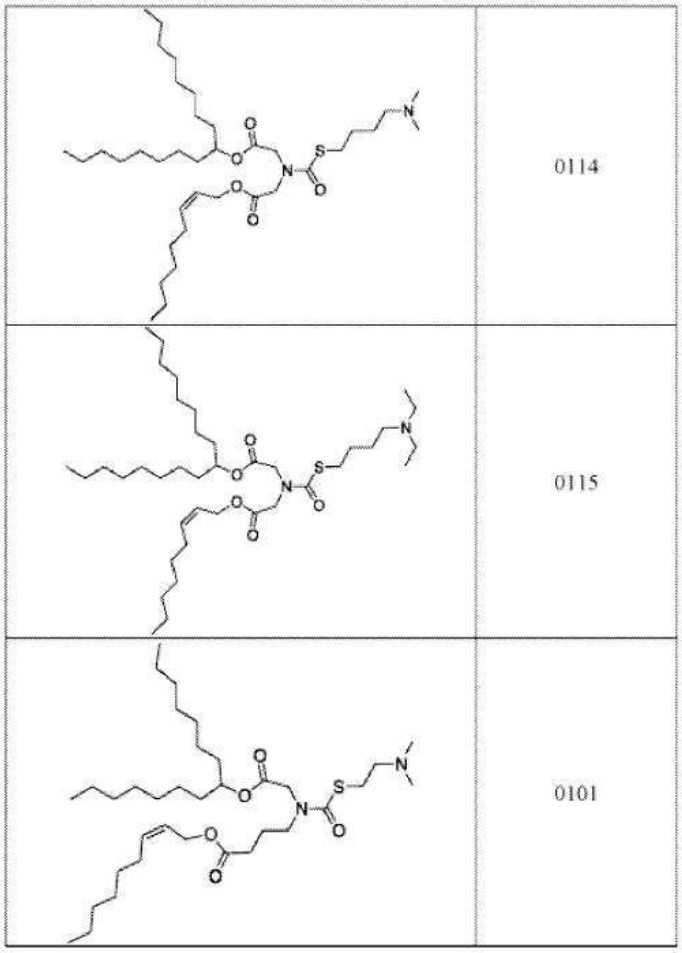
[1282]



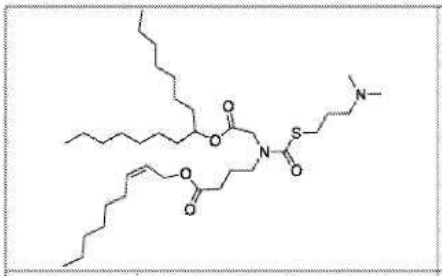
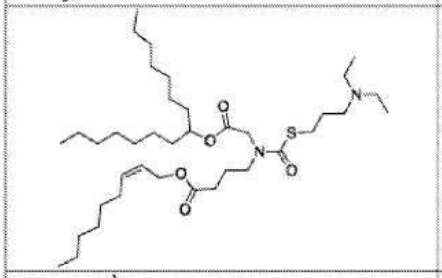
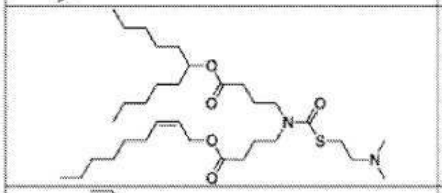
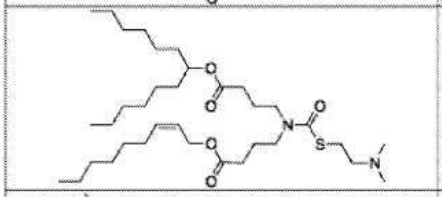
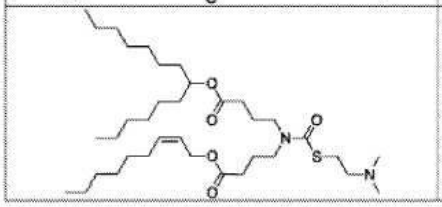
[1283]

	0061
	0100
	0117

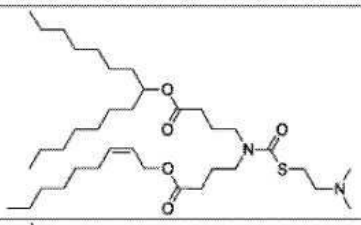
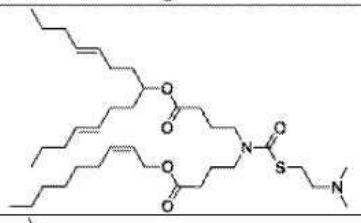
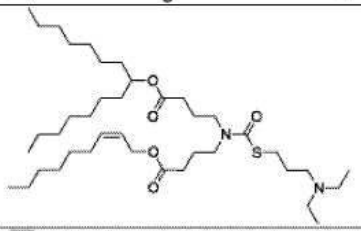
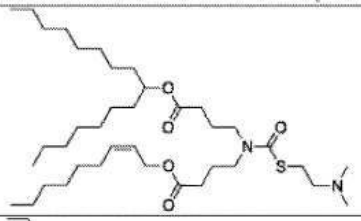
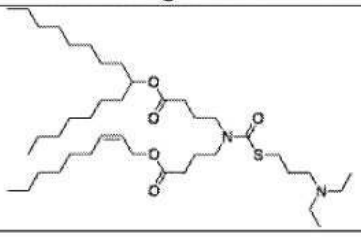
[1284]



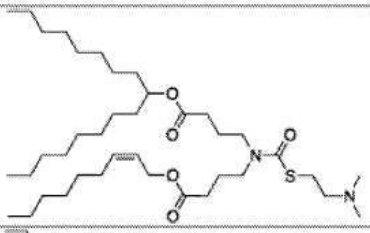
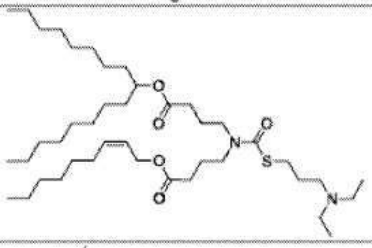
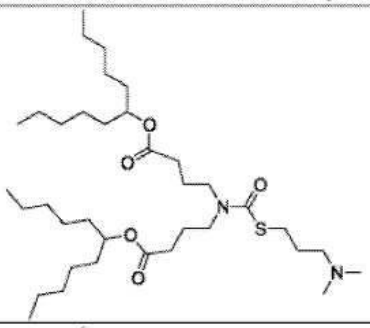
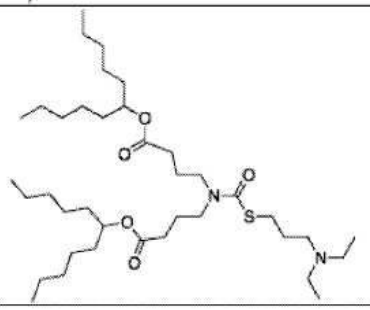
[1285]

	0106
	0116
	0043
	0086
	0058

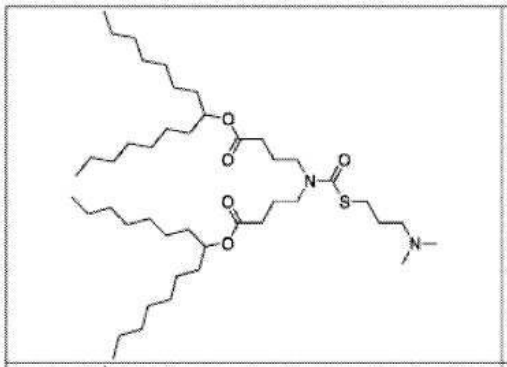
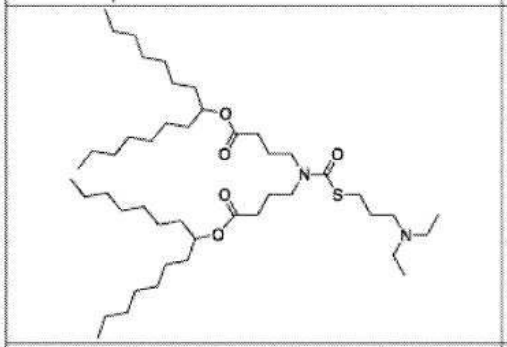
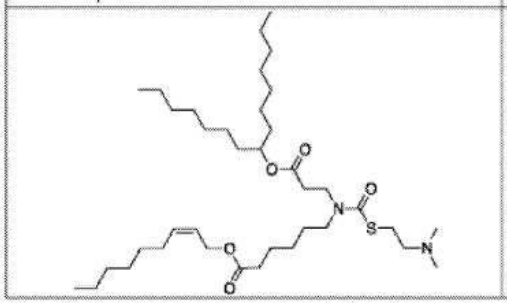
[1286]

	0081
	0123
	0122
	0057
	0088

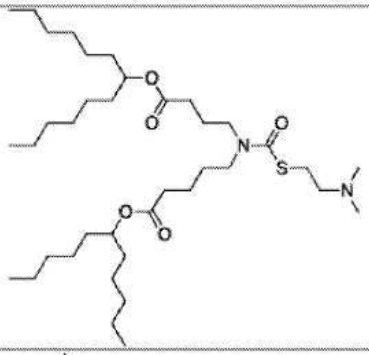
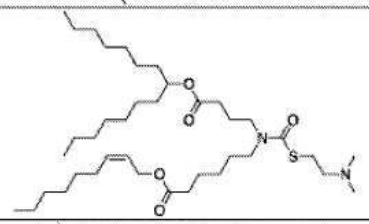
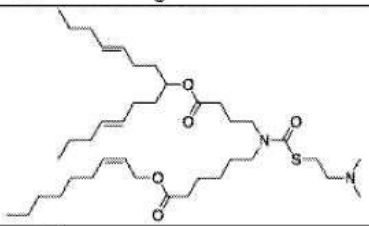
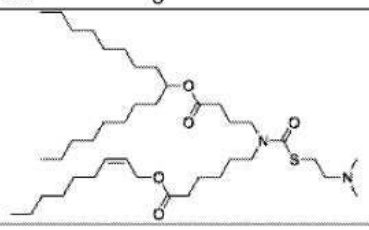
[1287]

	0087
	0124
	0128
	0127

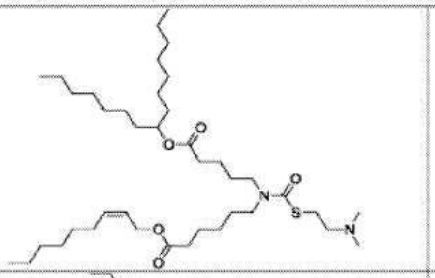
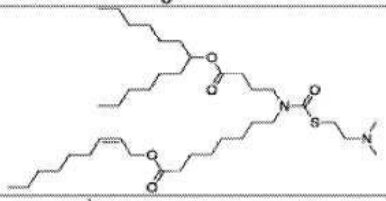
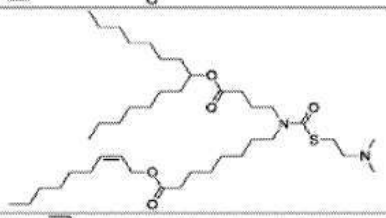
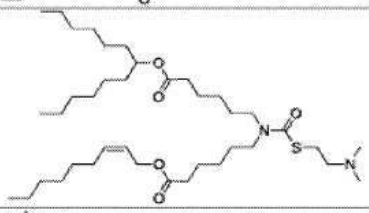
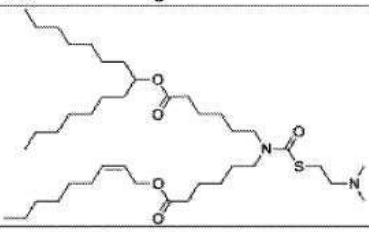
[1288]

	0126
	0129
	0082

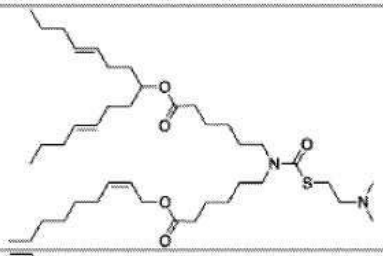
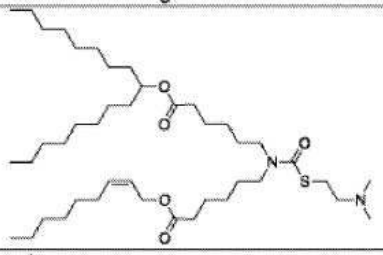
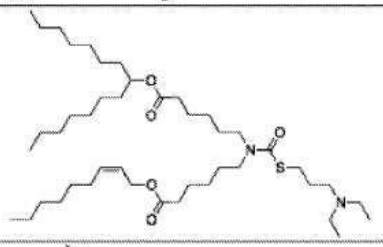
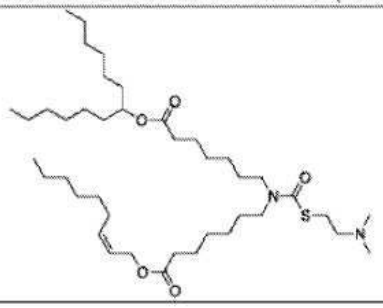
[1289]

	0085
	0083
	0121
	0091

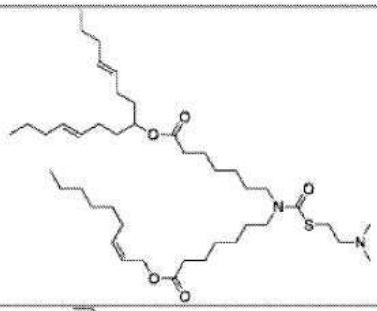
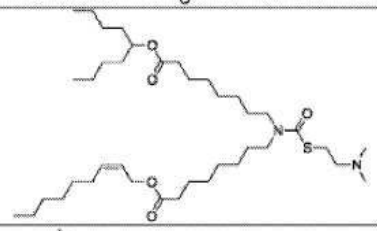
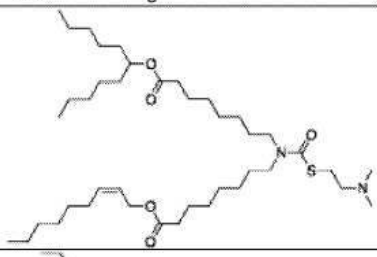
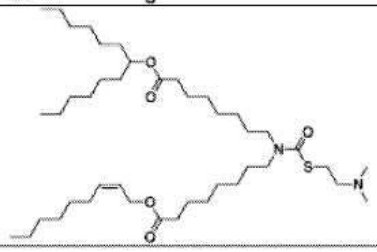
[1290]

	0102
	0098
	0092
	0084
	0095

[1291]

	0125
	0094
	0109
	0110

[1292]

	0118
	0108
	0107
	0093

[1293]

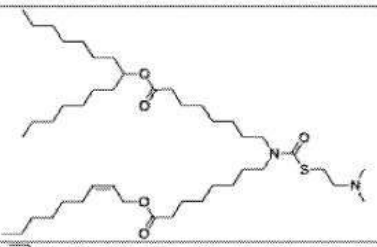
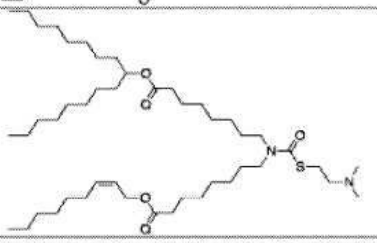
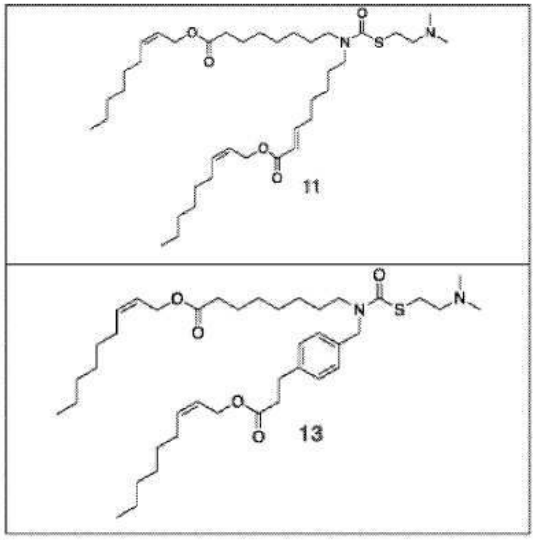
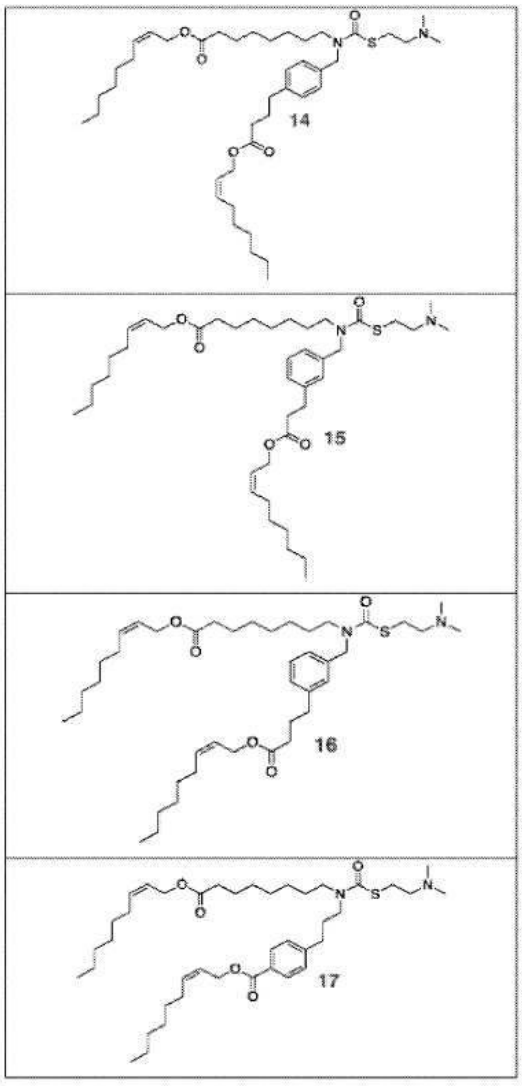
	0097
	0096

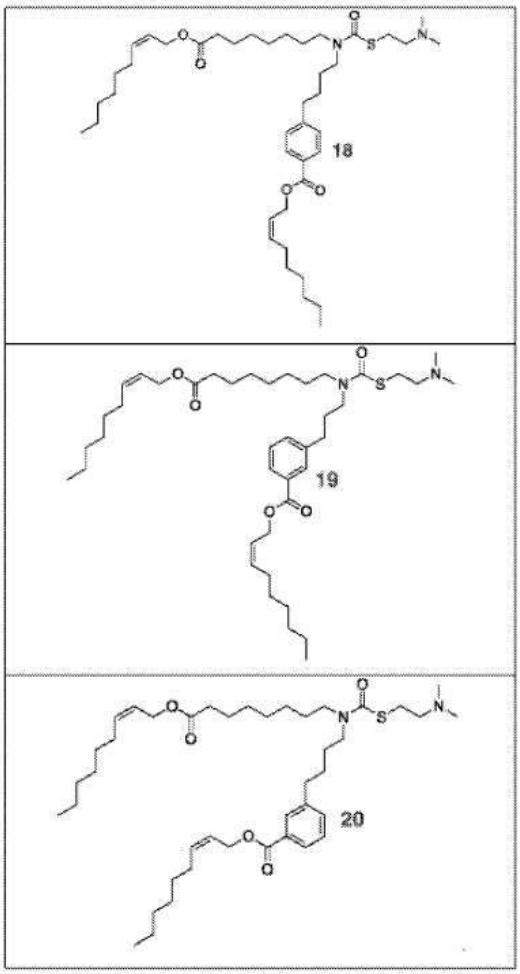
표 15

[1294]



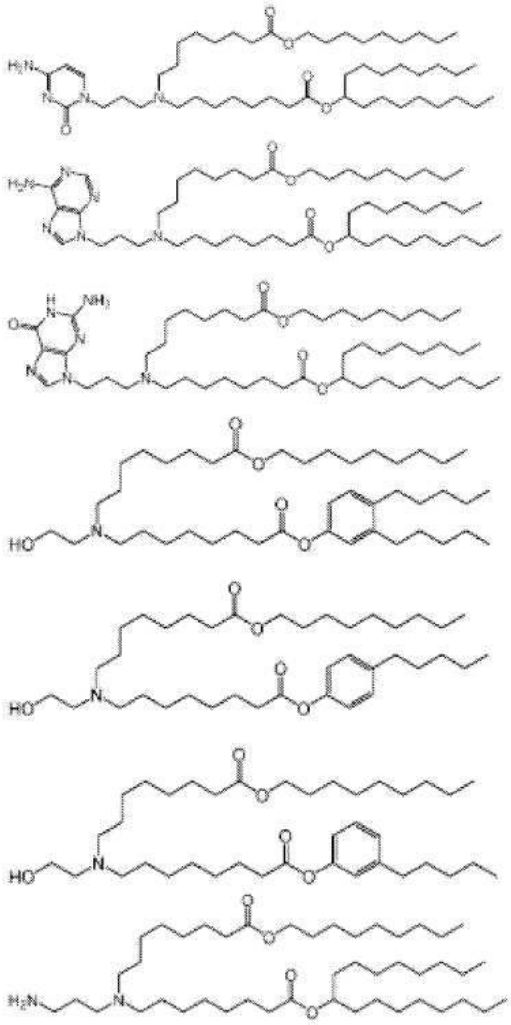
[1295]



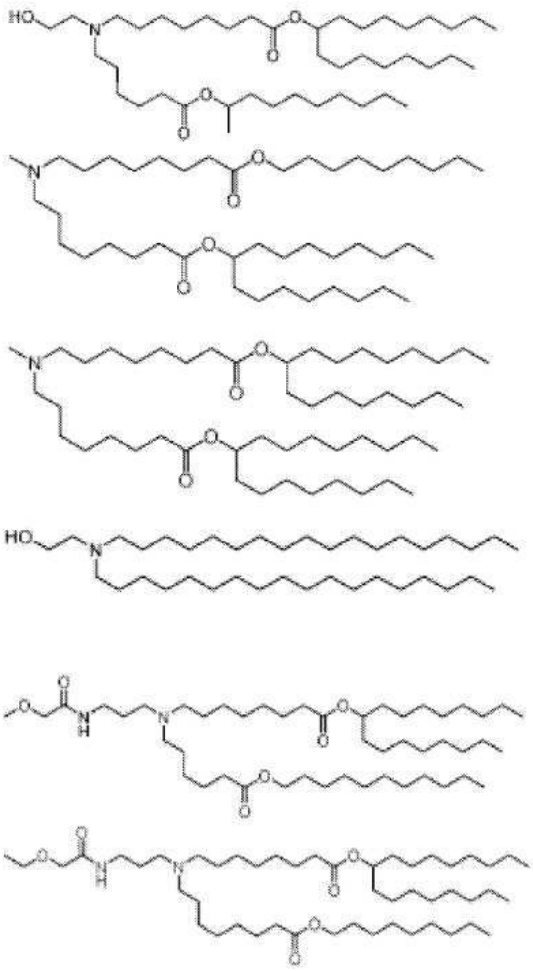


[1296]

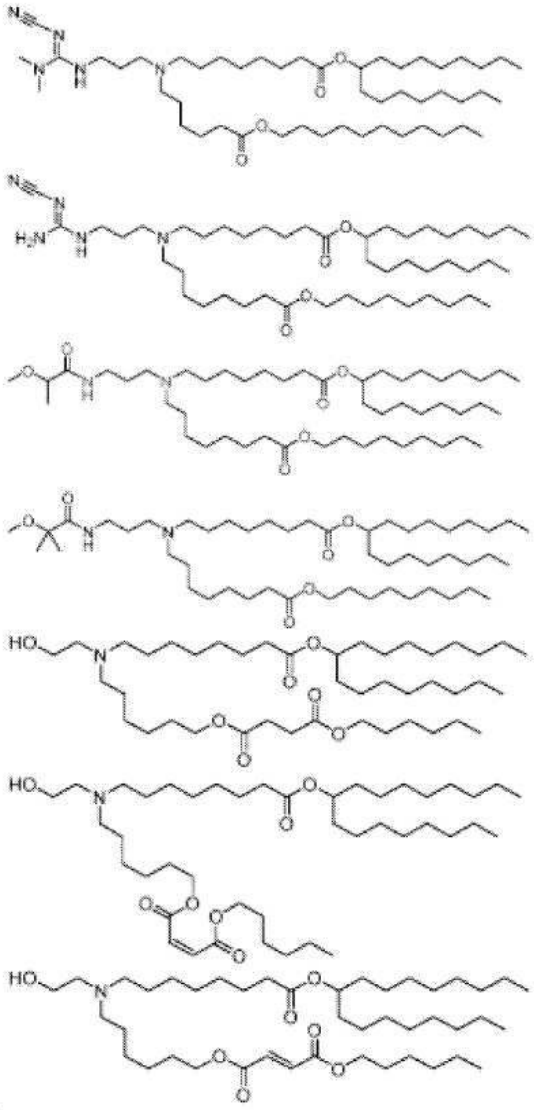
[1297] 일부 실시형태에서, 이온화 가능한 지질은 국제 특허 출원 제PCT/US2019/015913호에 기재된 바와 같다. 일부 실시형태에서, 이온화 가능한 지질은 하기로부터 선택된다:



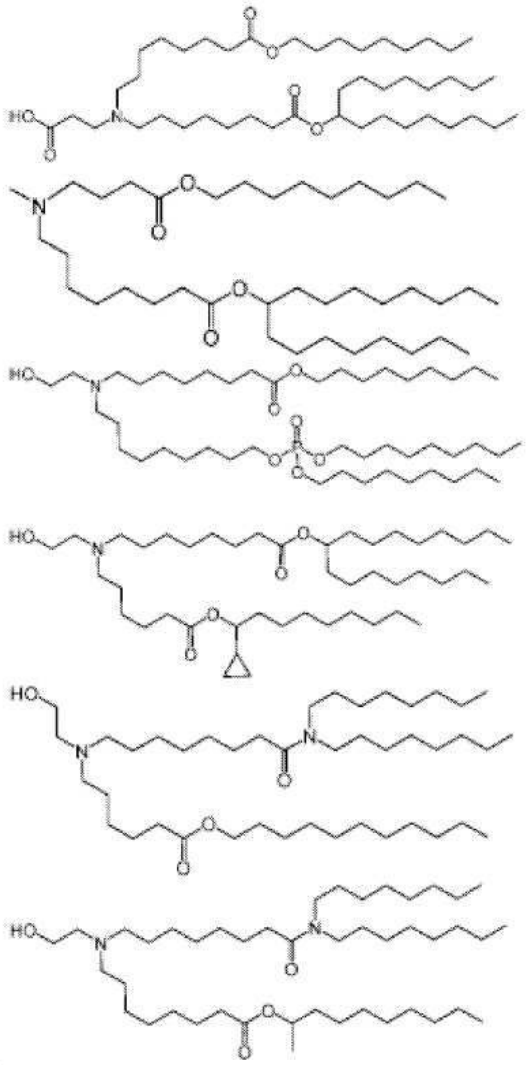
[1298]



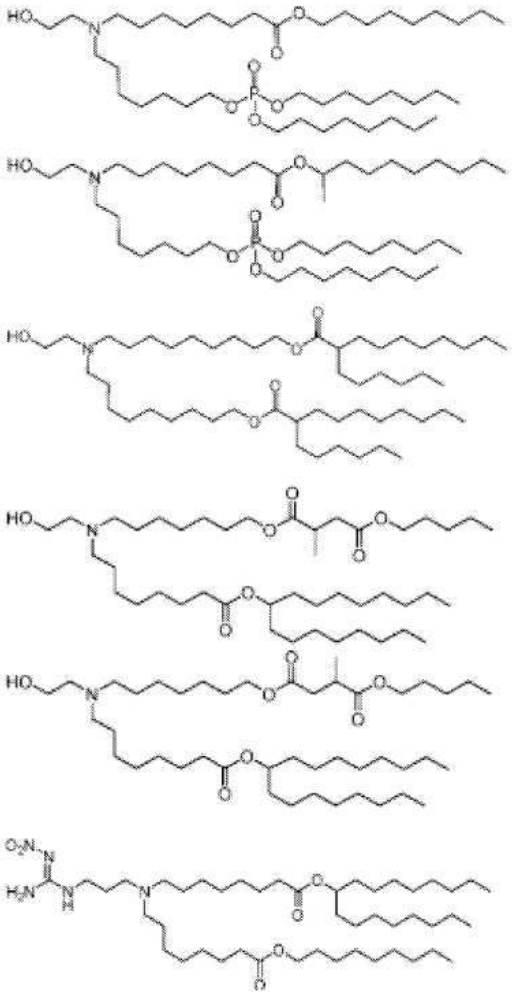
[1299]



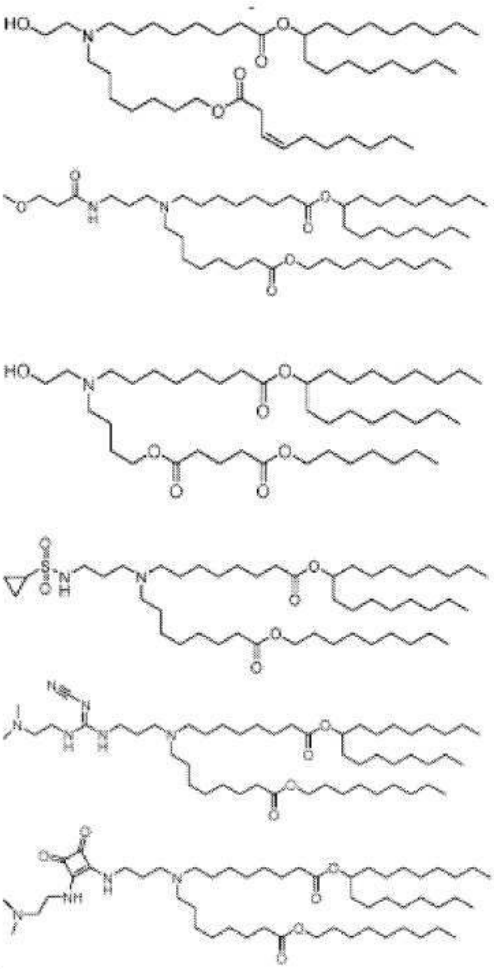
[1300]



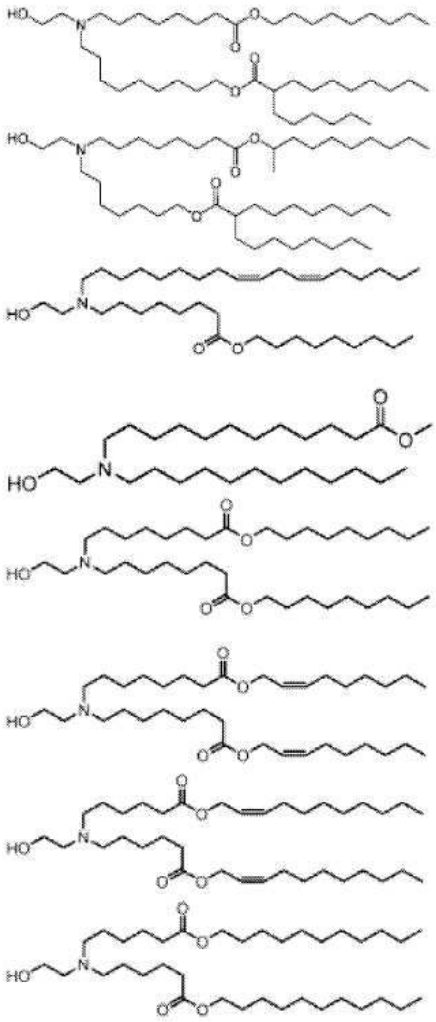
[1301]



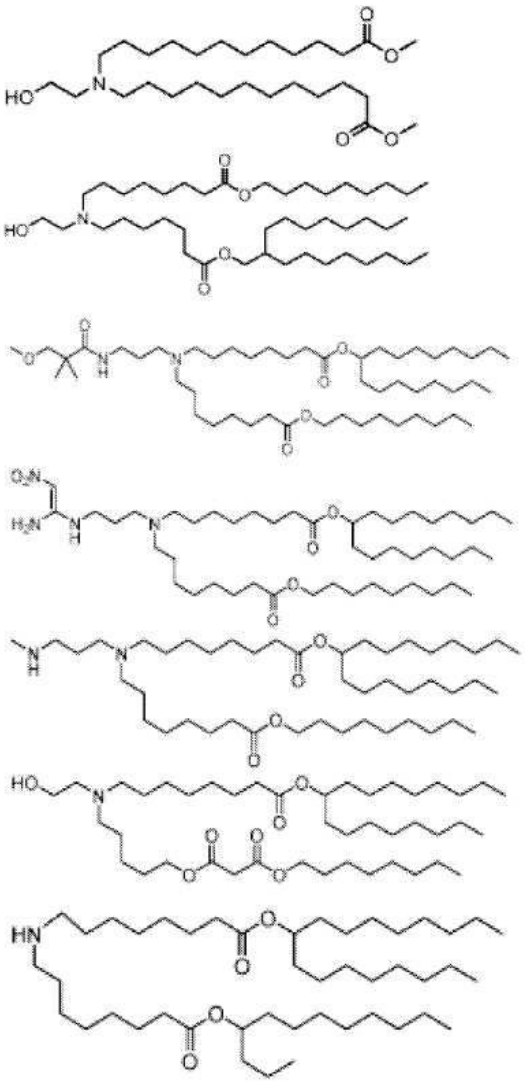
[1302]



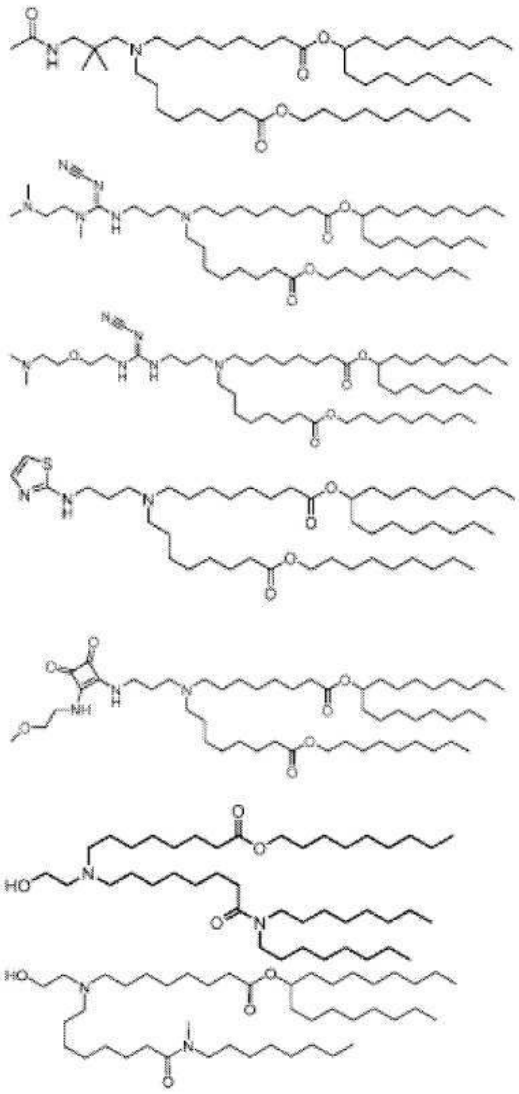
[1303]



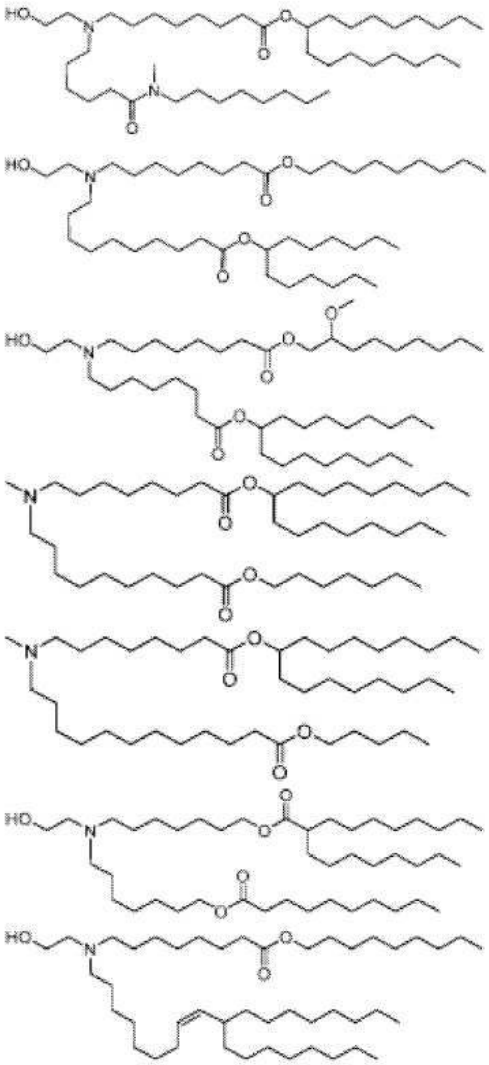
[1304]



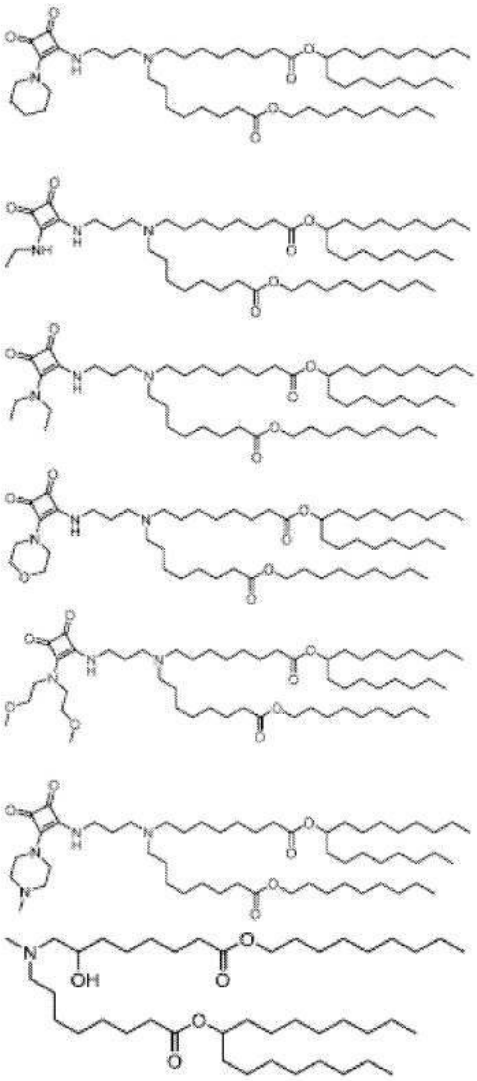
[1305]



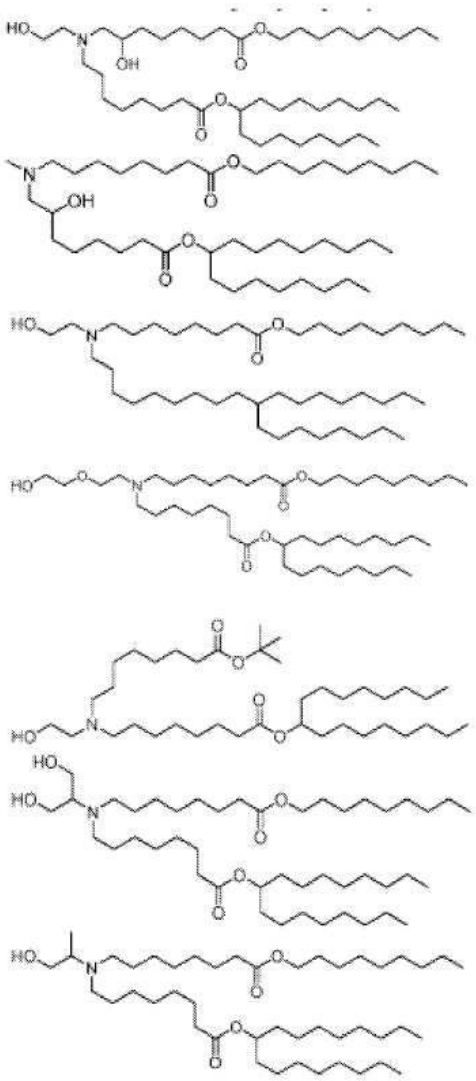
[1306]



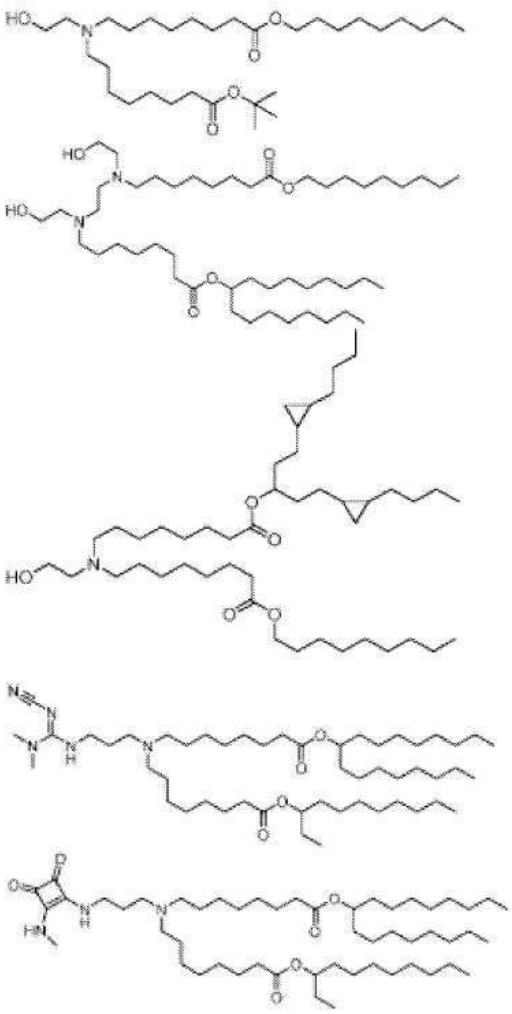
[1307]



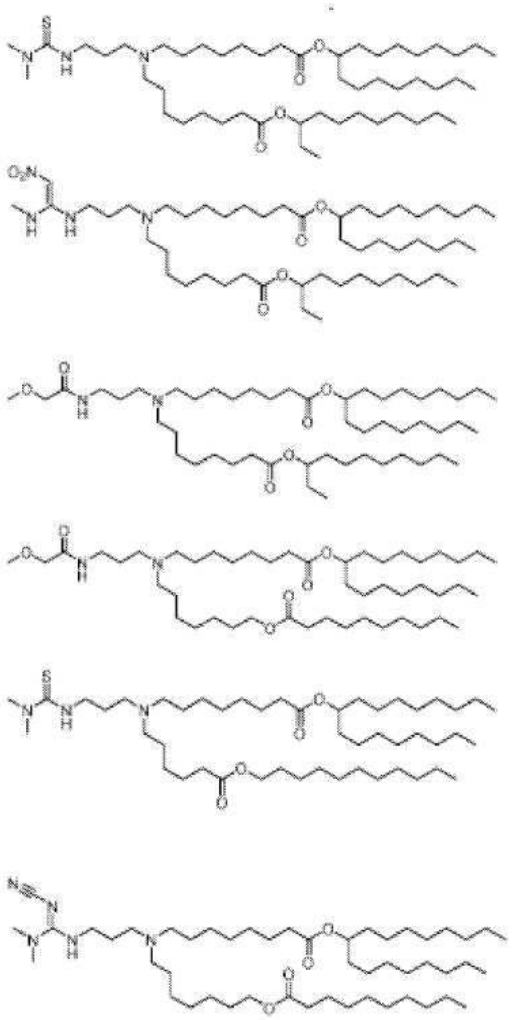
[1308]



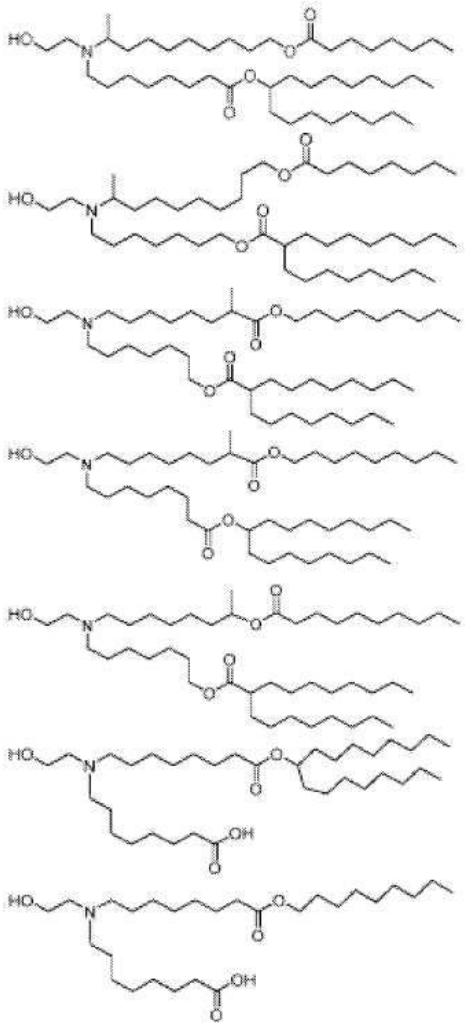
[1309]



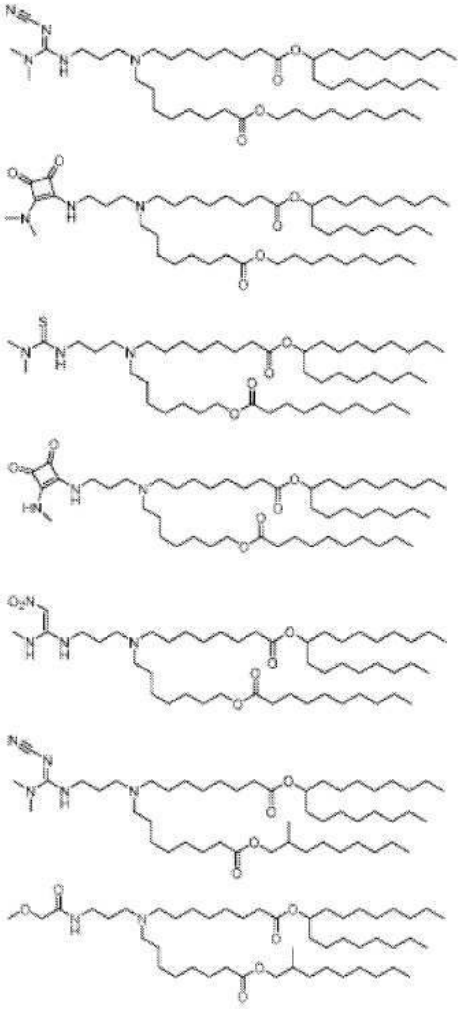
[1310]



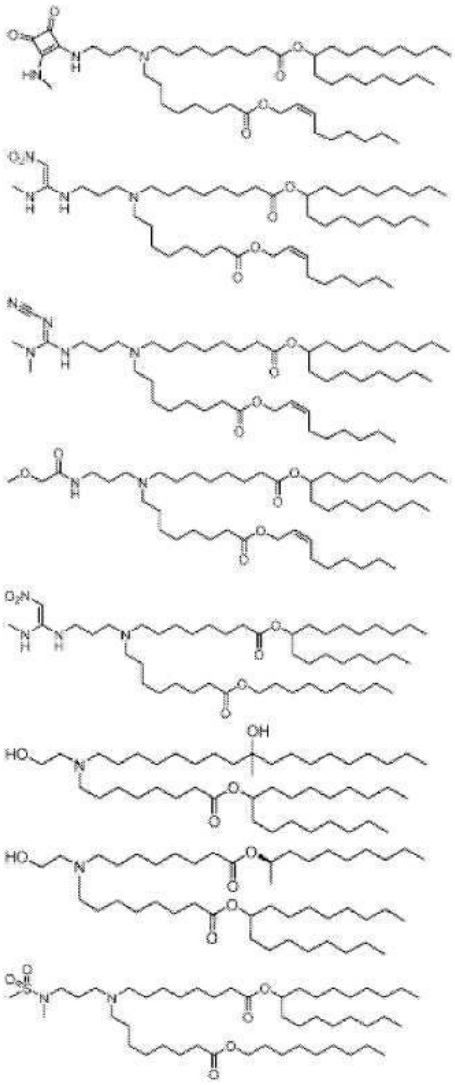
[1311]



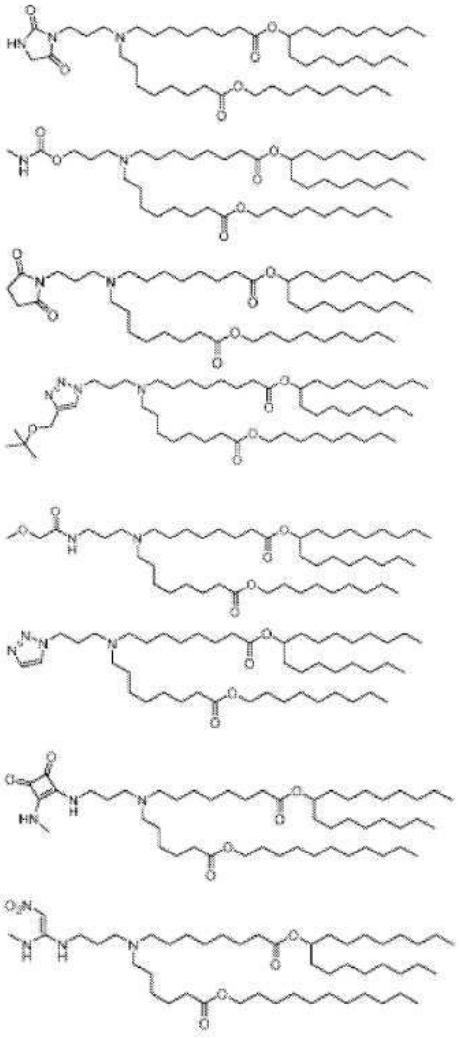
[1312]



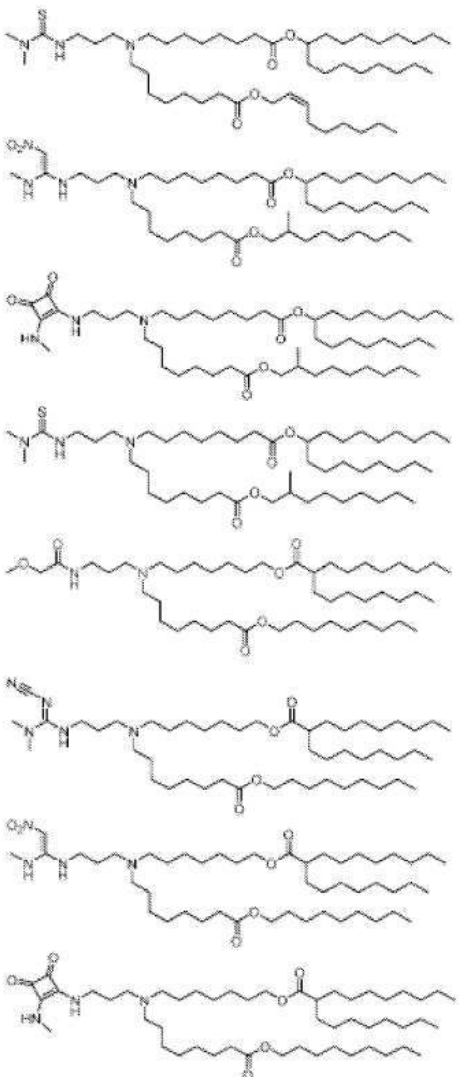
[1313]



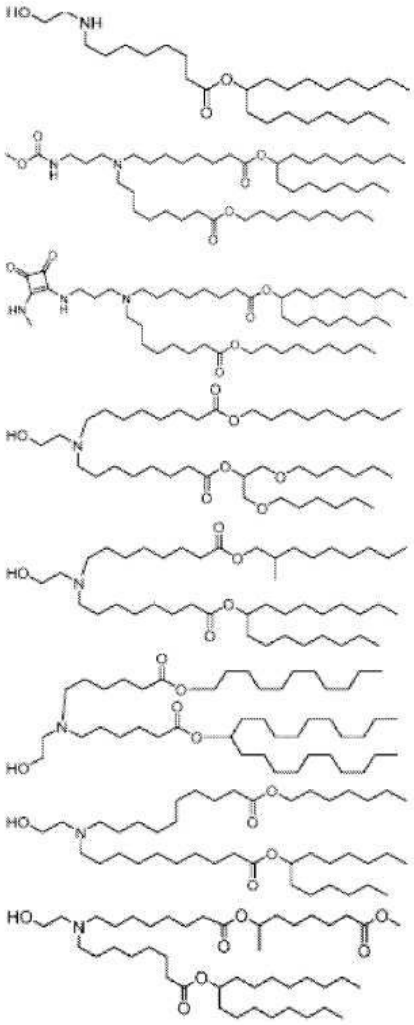
[1314]



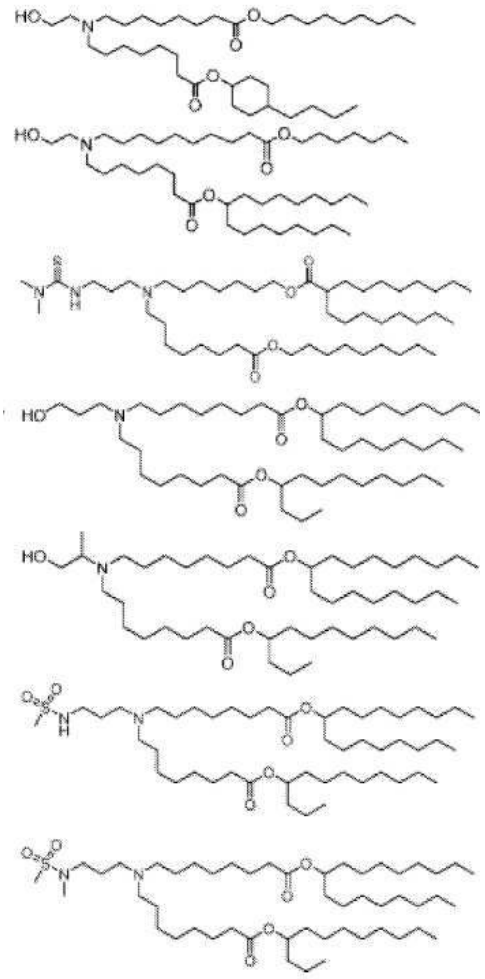
[1315]



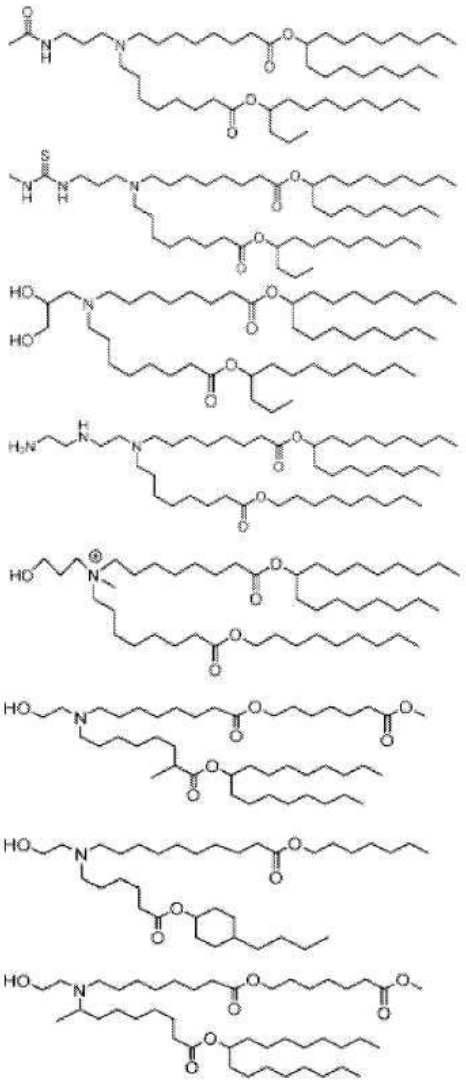
[1316]



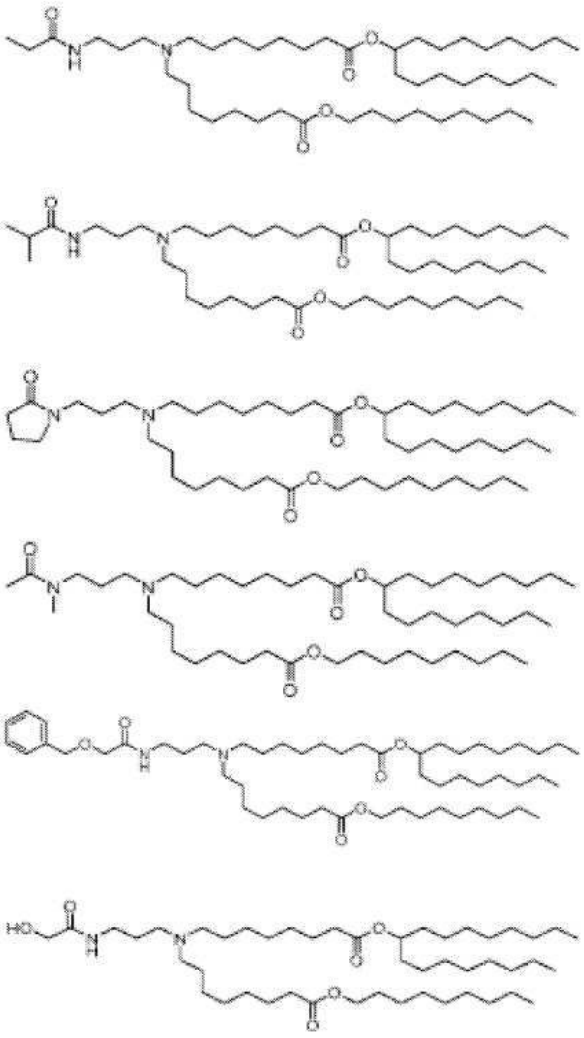
[1317]



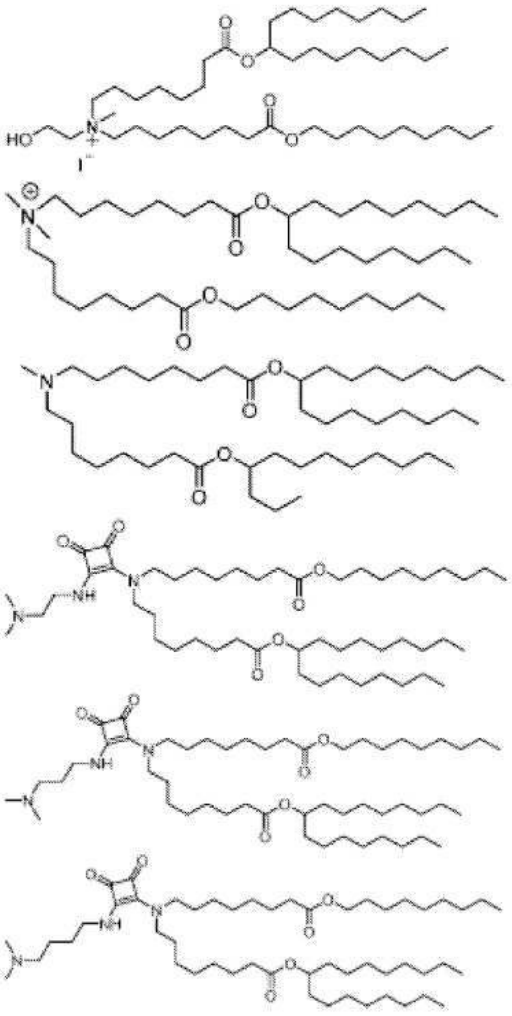
[1318]



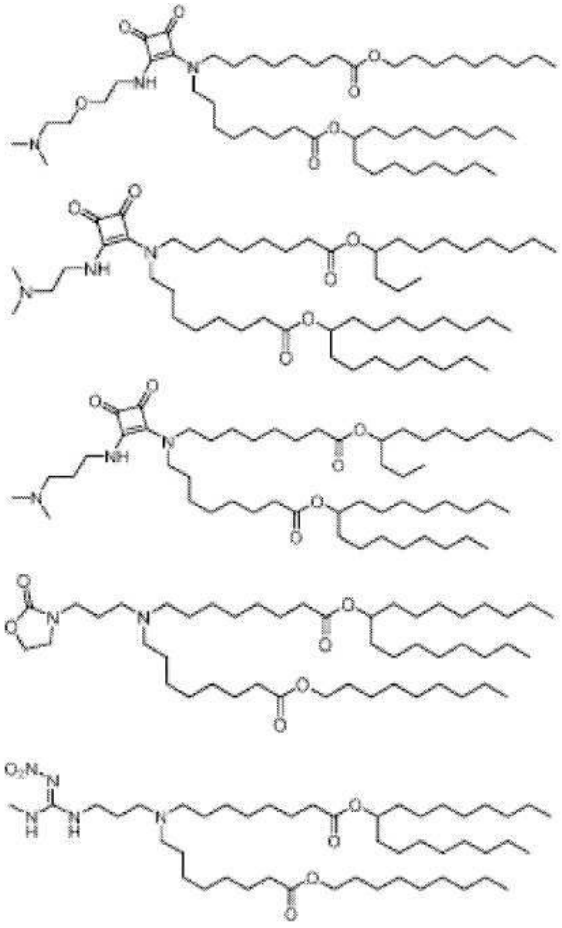
[1319]



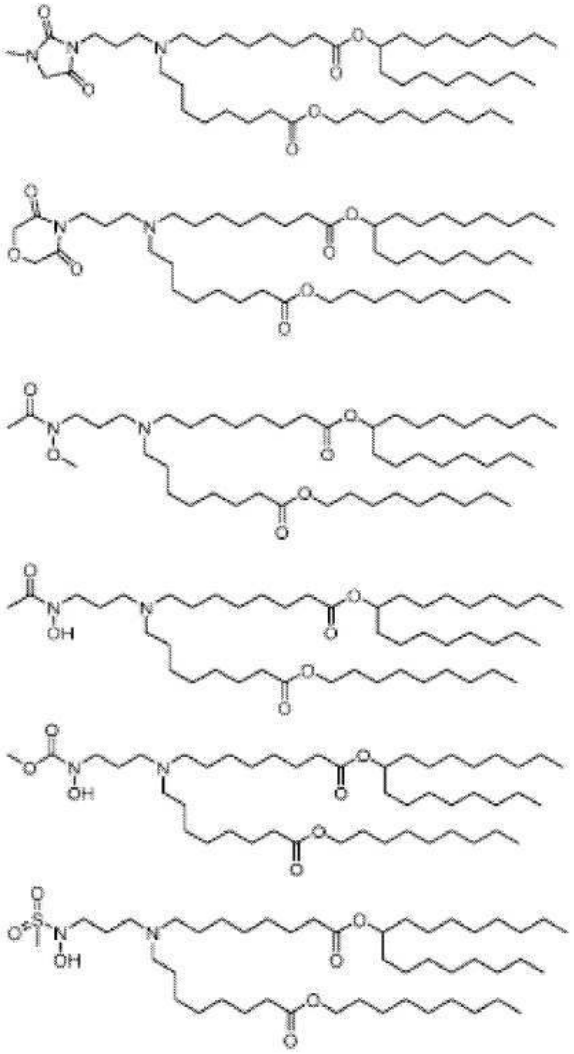
[1320]



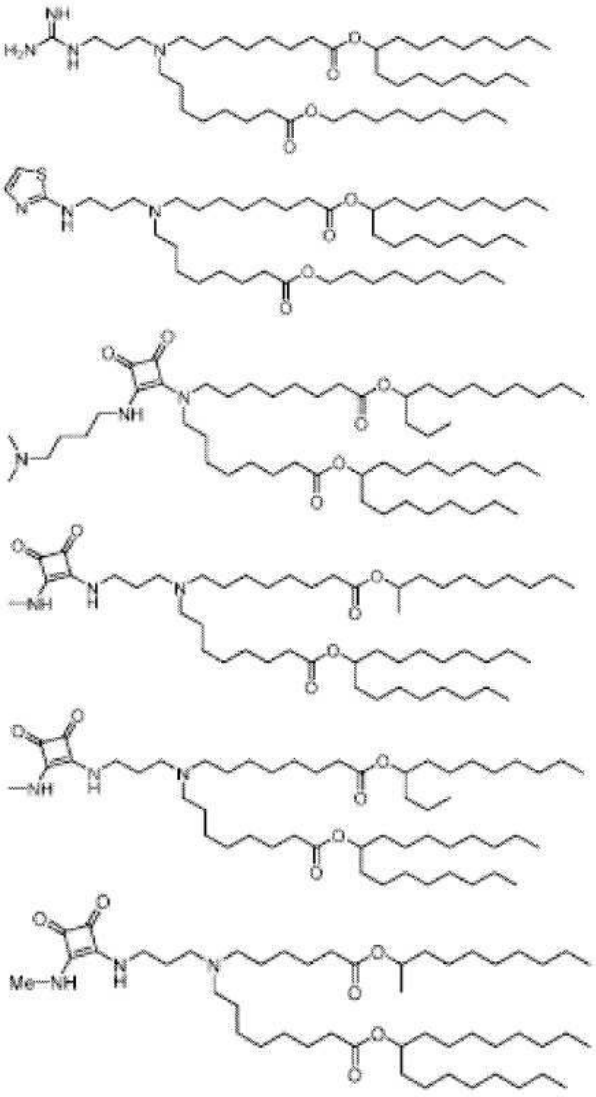
[1321]



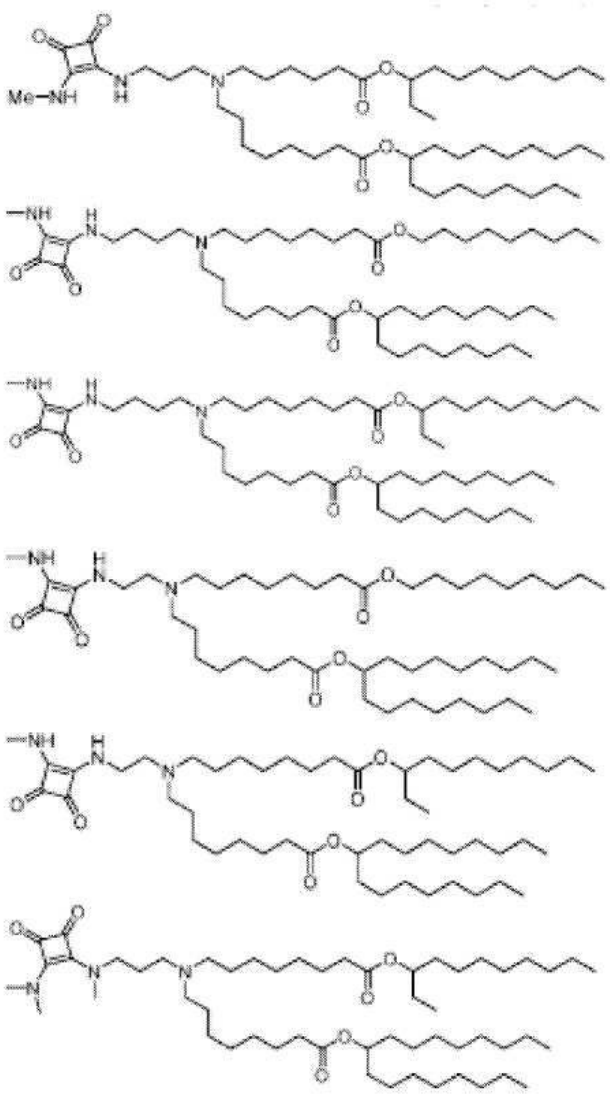
[1322]



[1323]

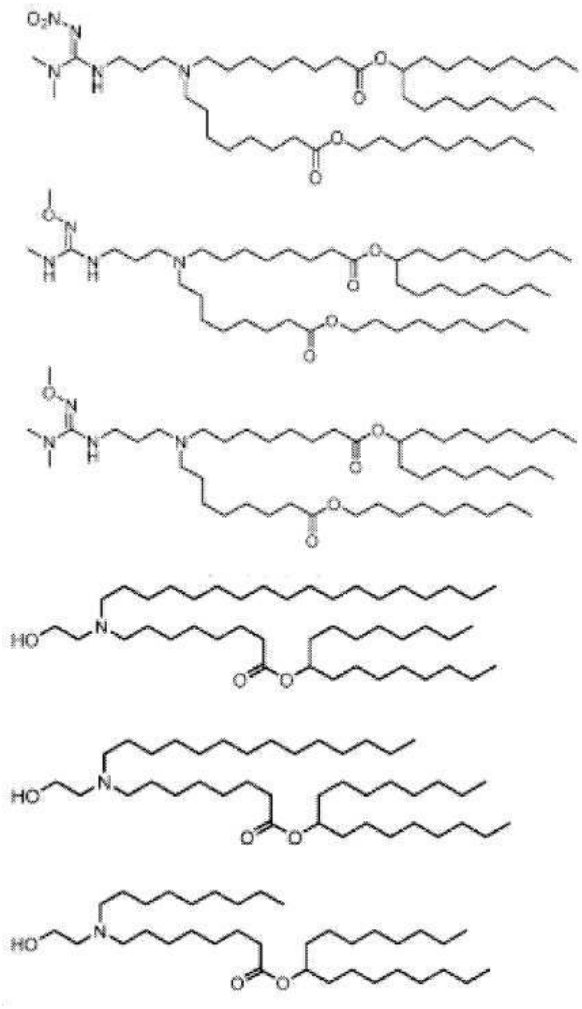


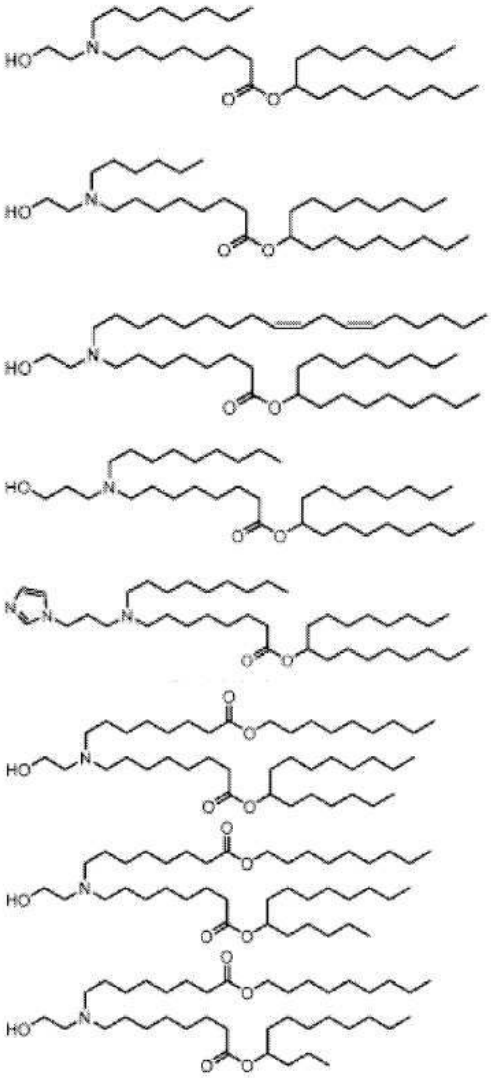
[1324]



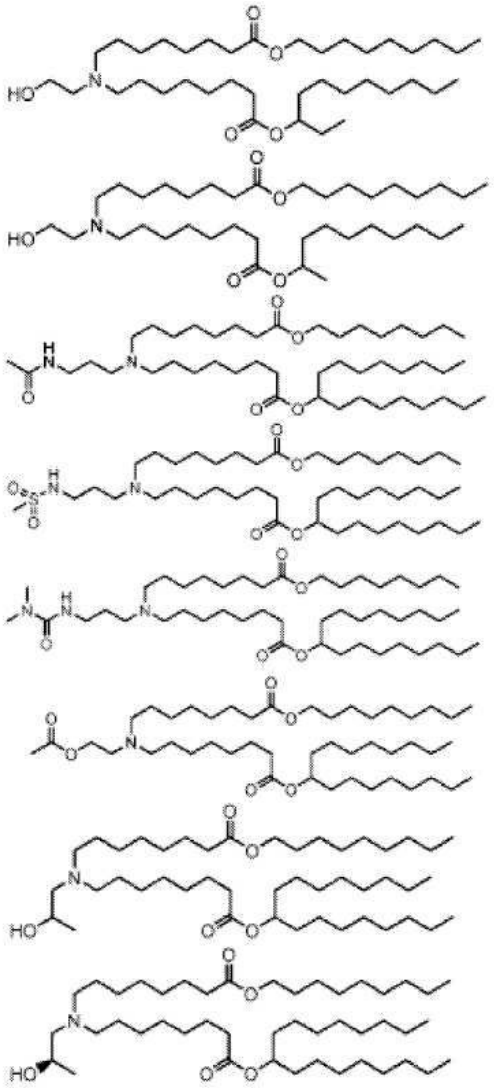
[1325]

[1326]

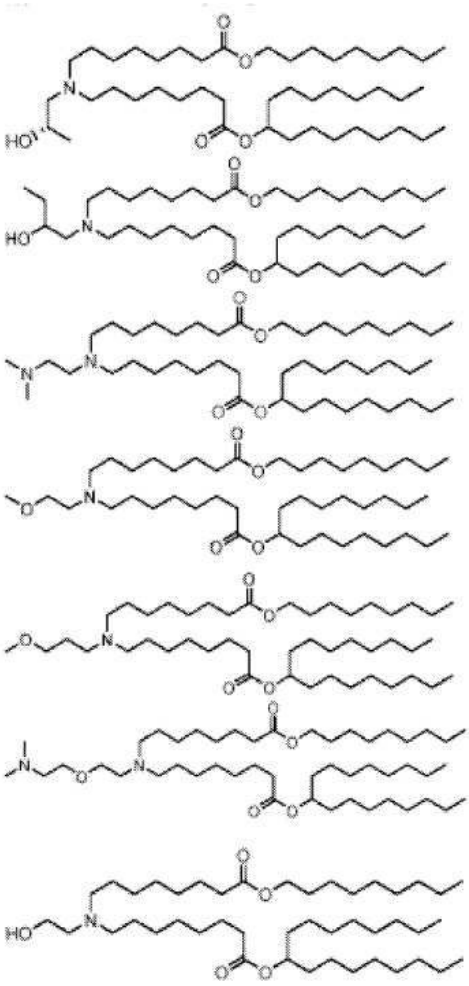




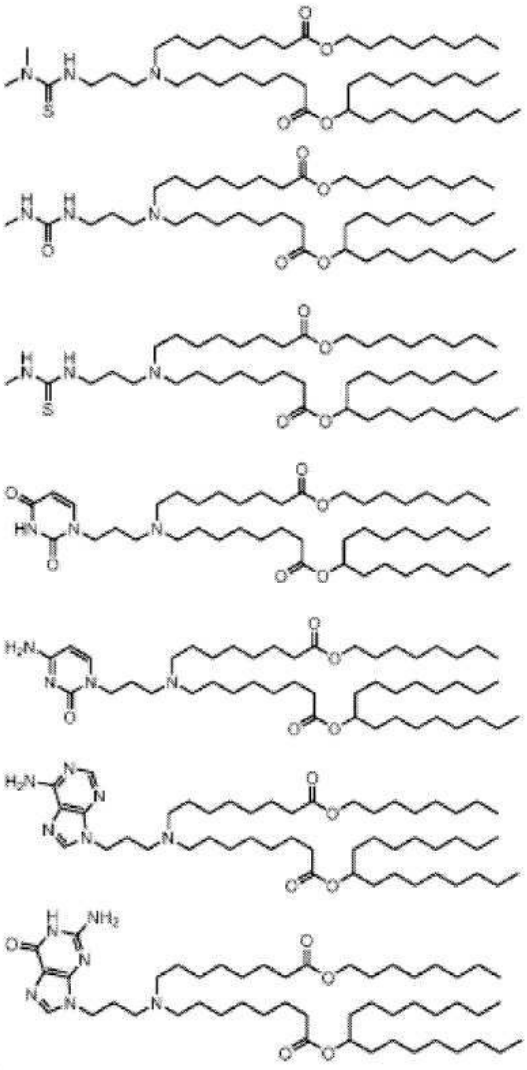
[1327]



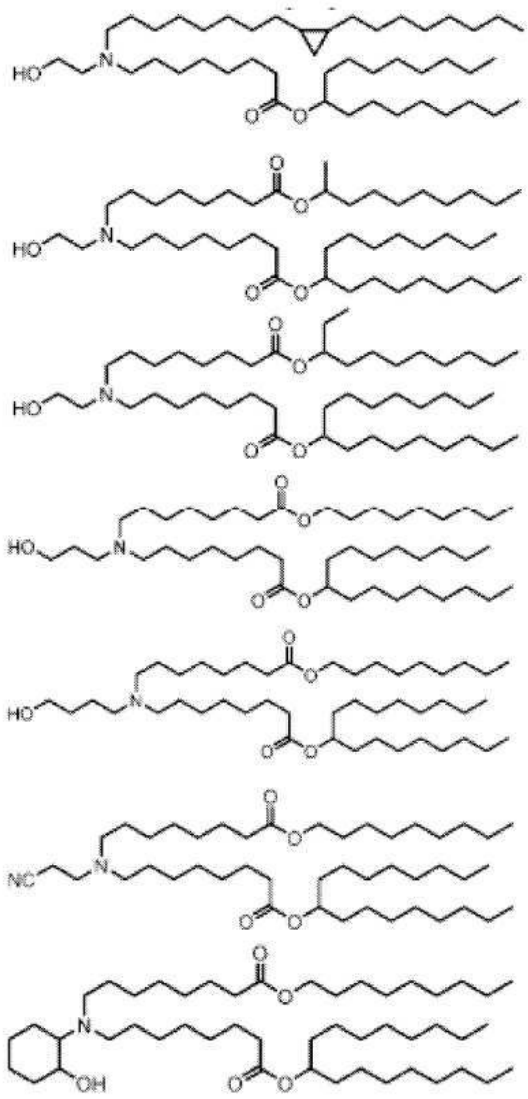
[1328]

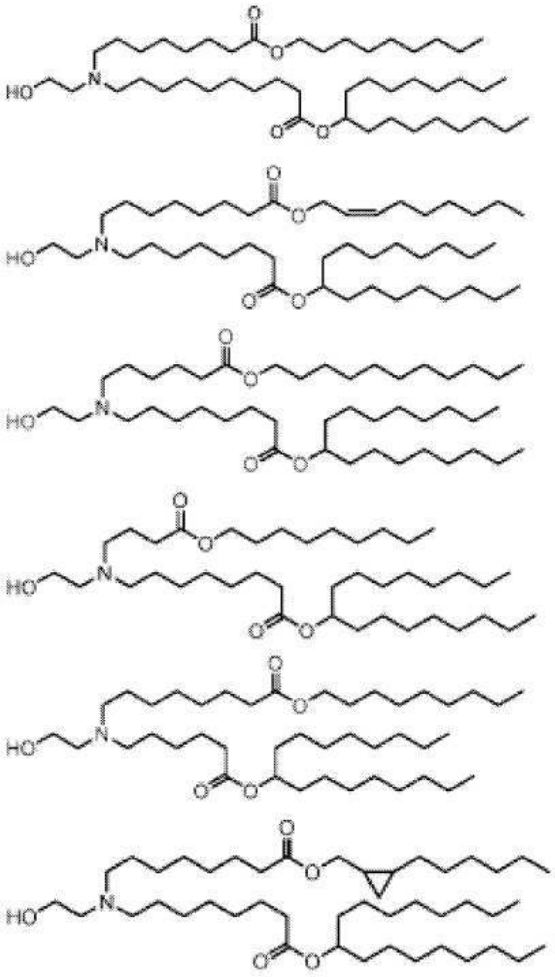


[1329]

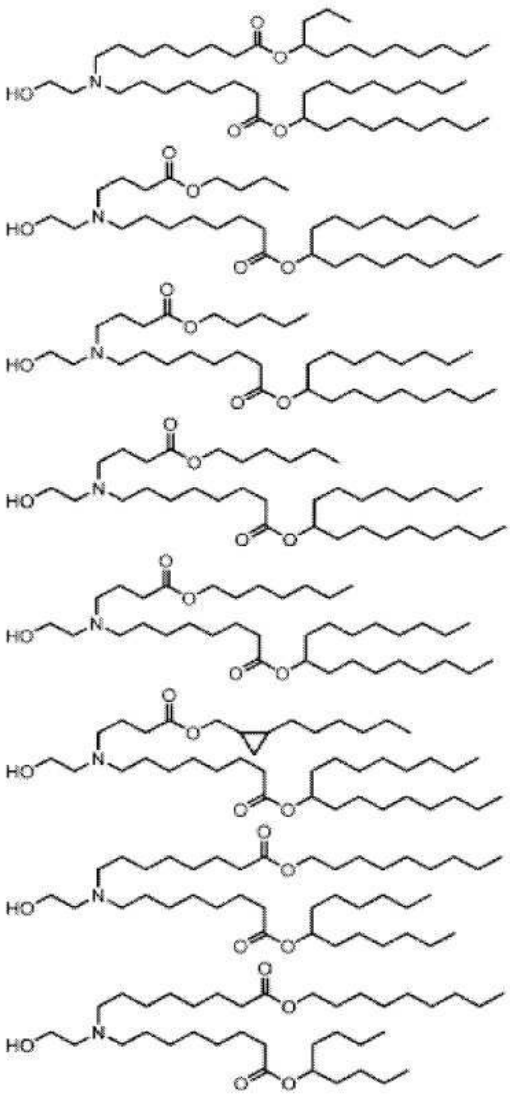


[1330]

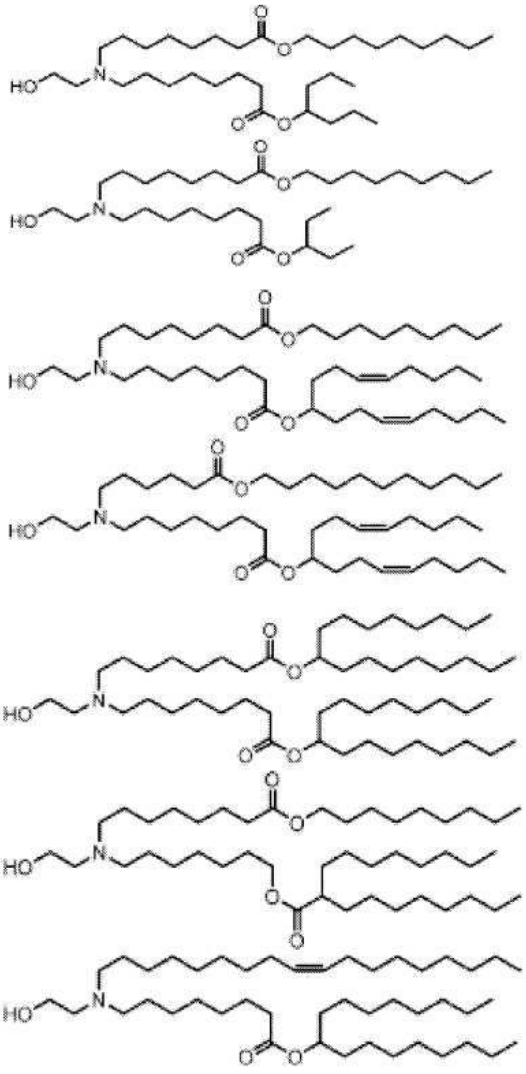
[1331] 

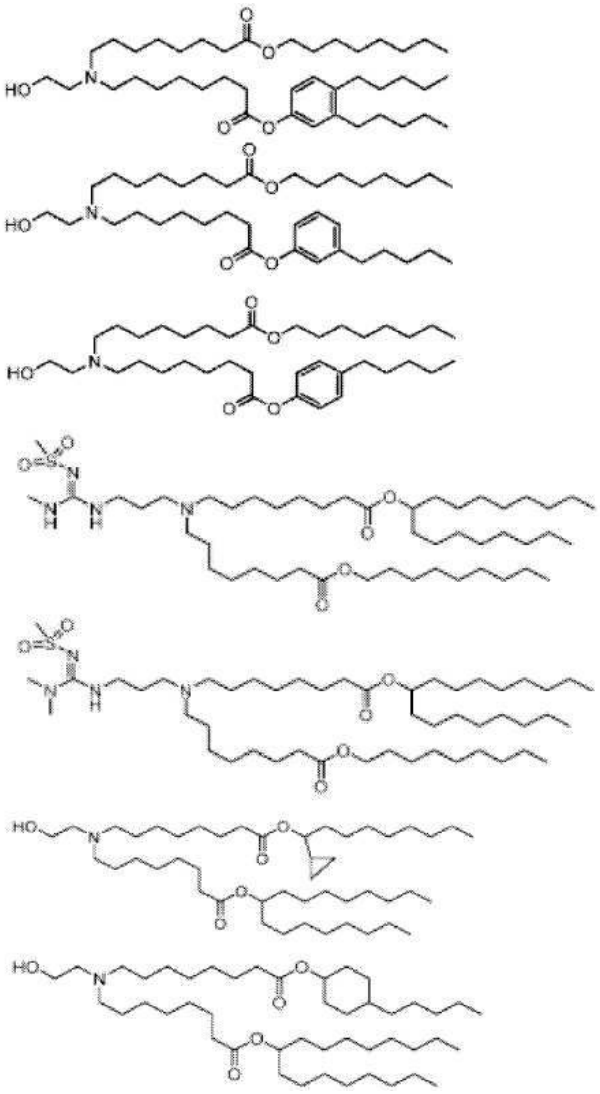


[1332]

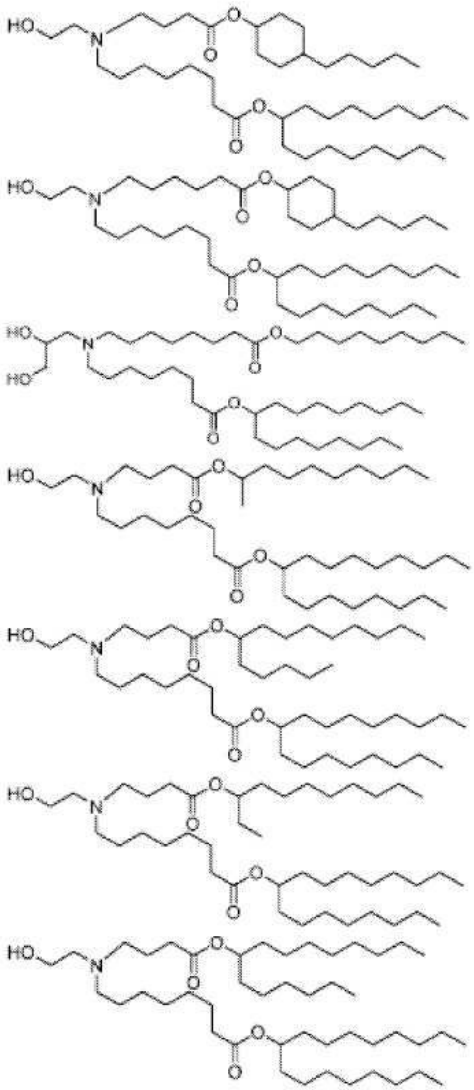


[1333]

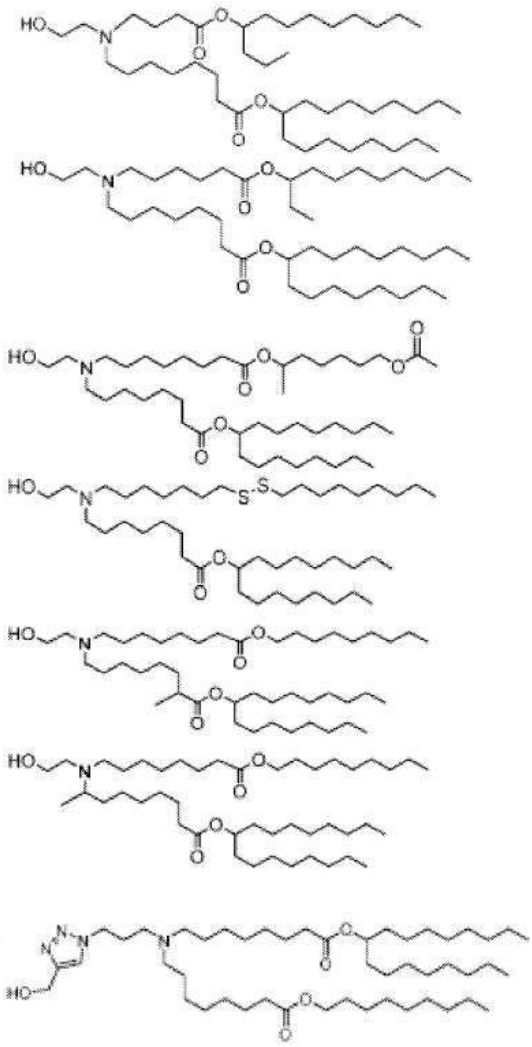
[1334] 



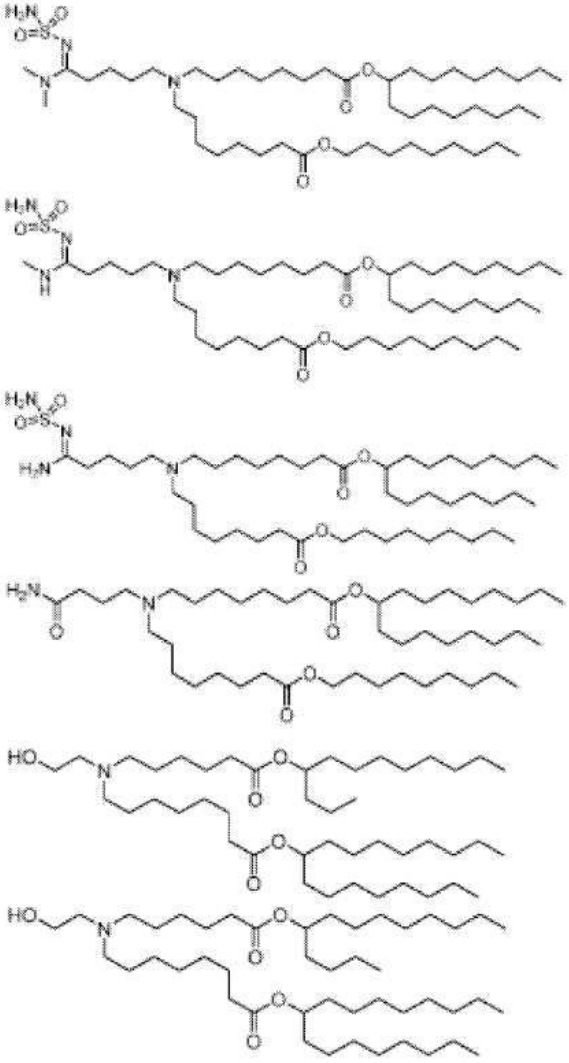
[1335]



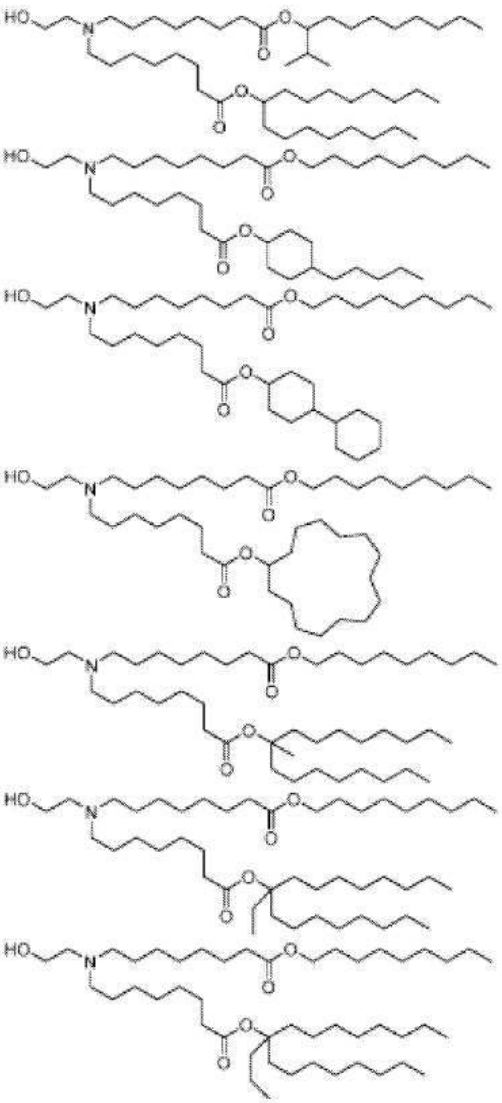
[1336]



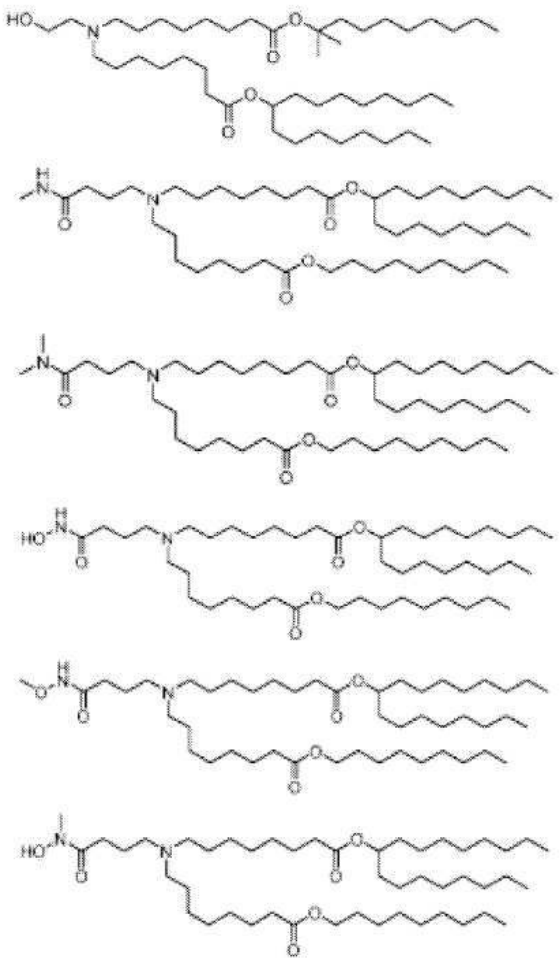
[1337]



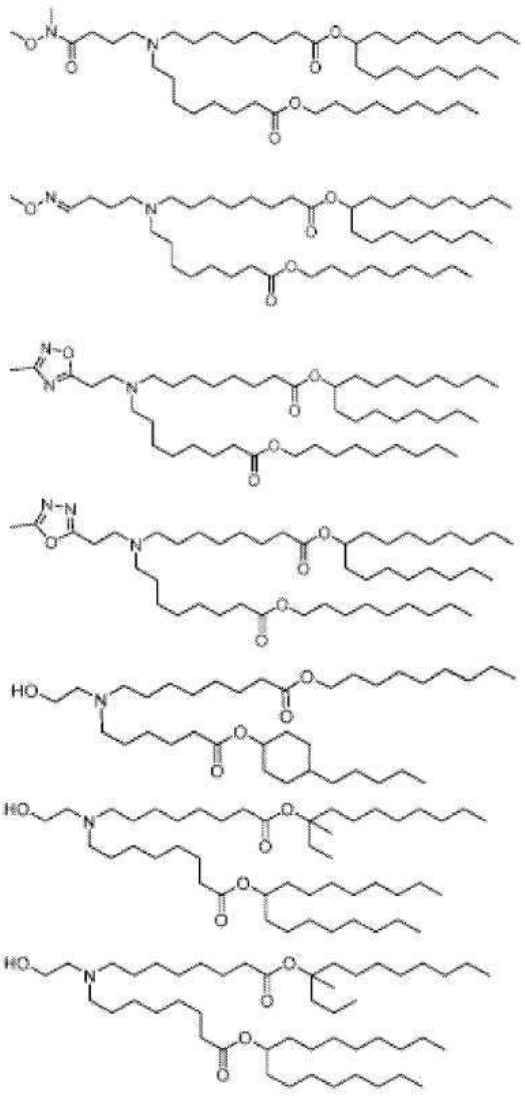
[1338]



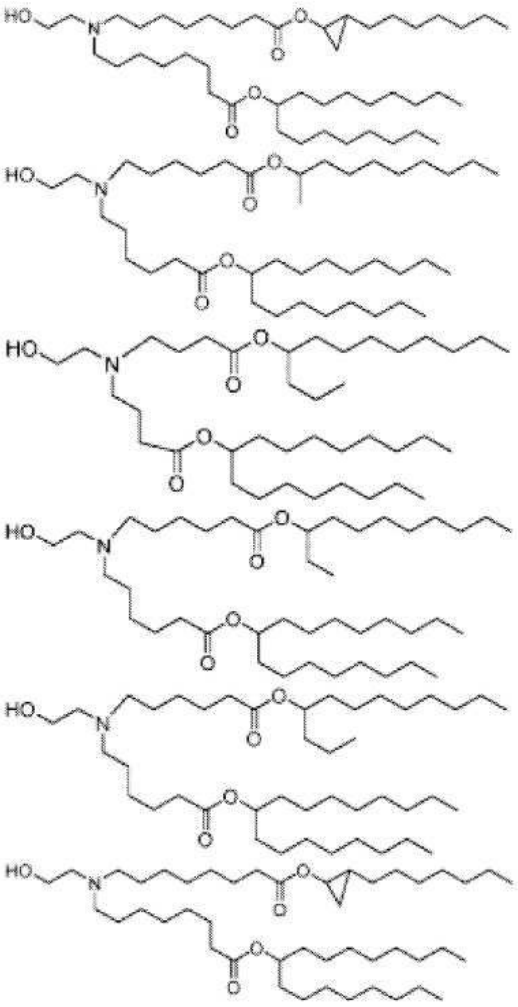
[1339]



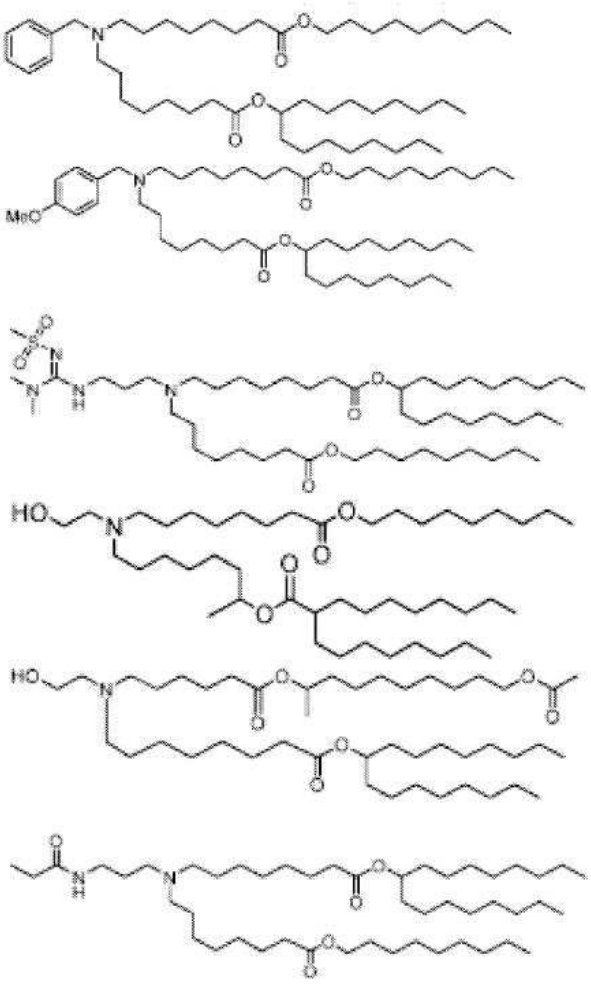
[1340]



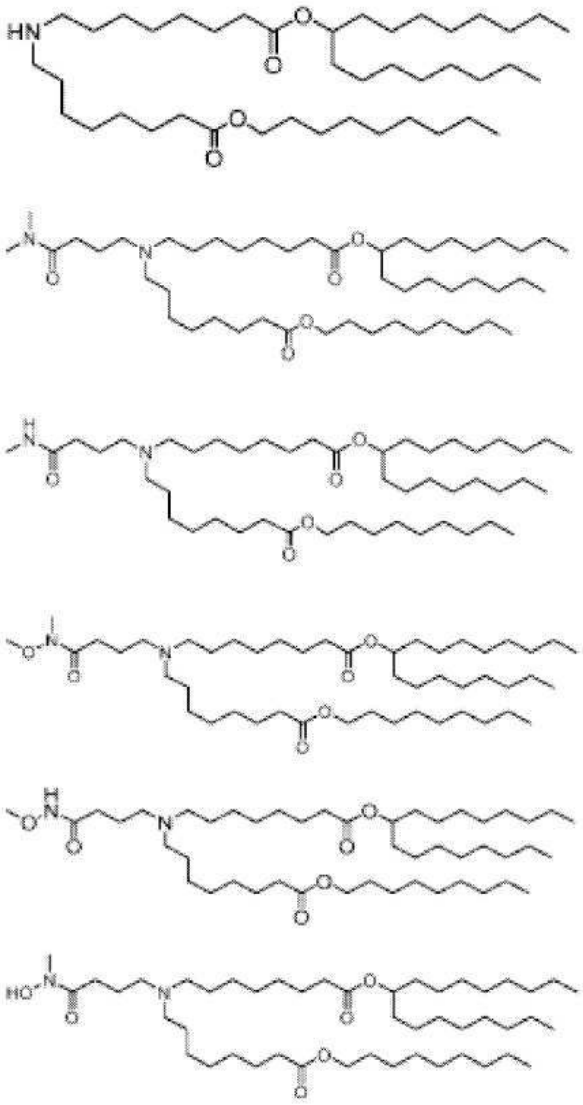
[1341]



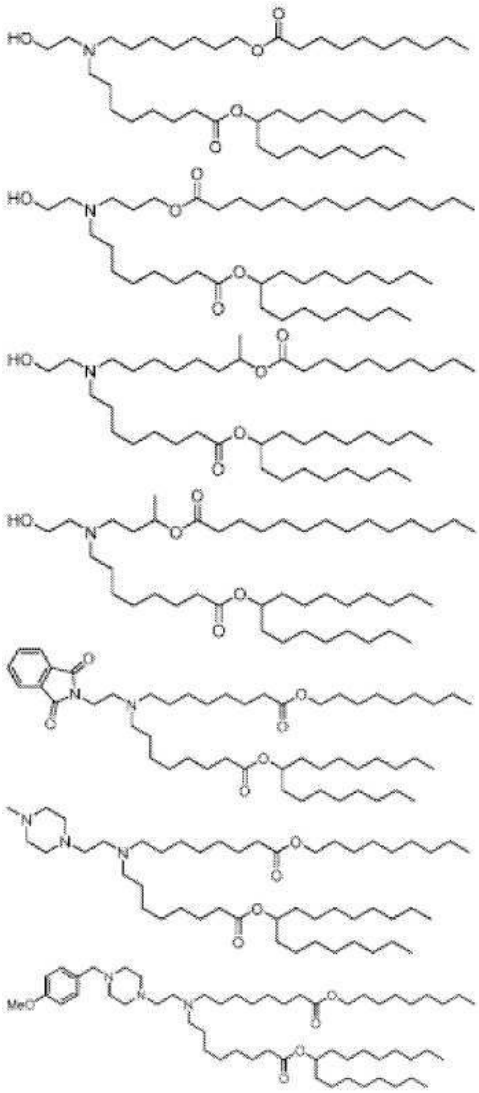
[1342]



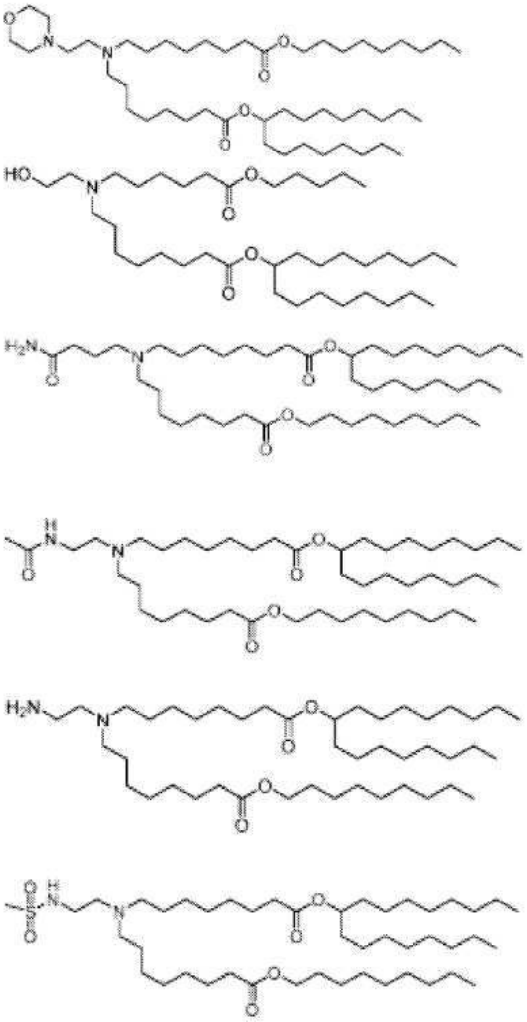
[1343]



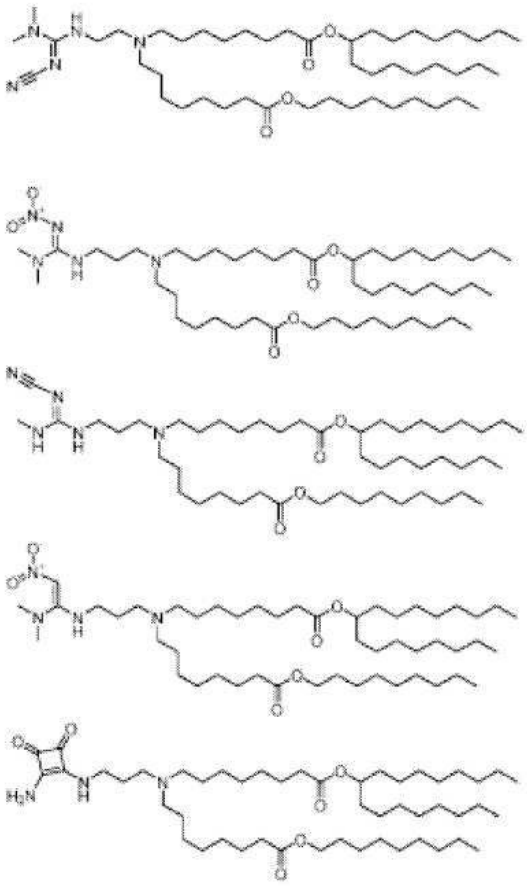
[1344]



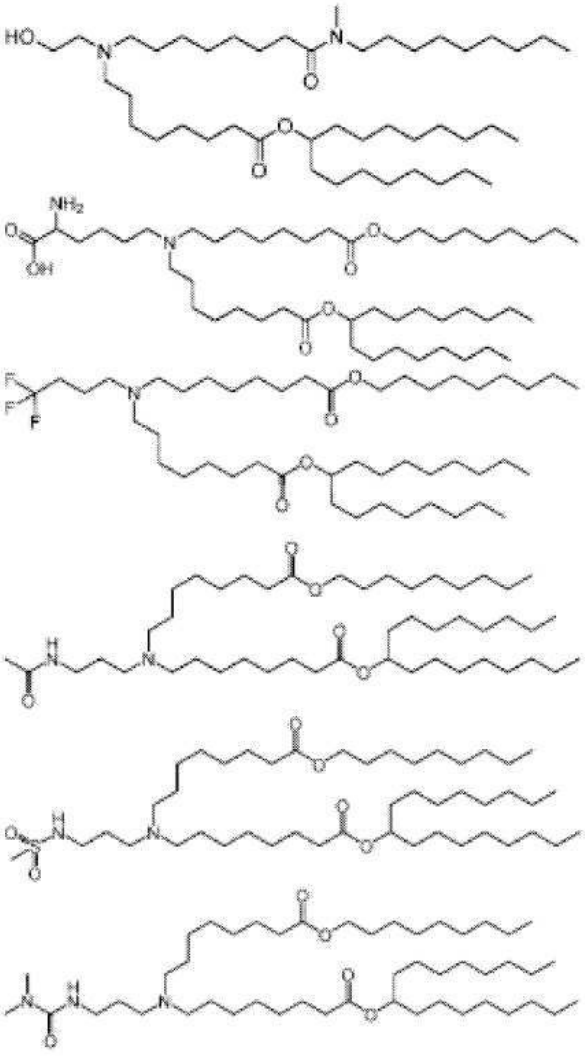
[1345]



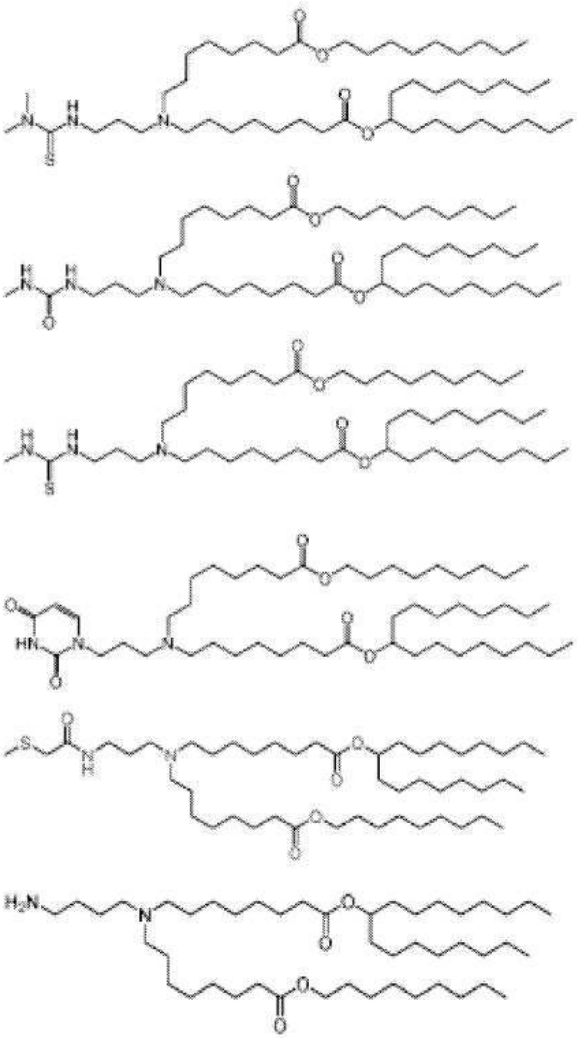
[1346]



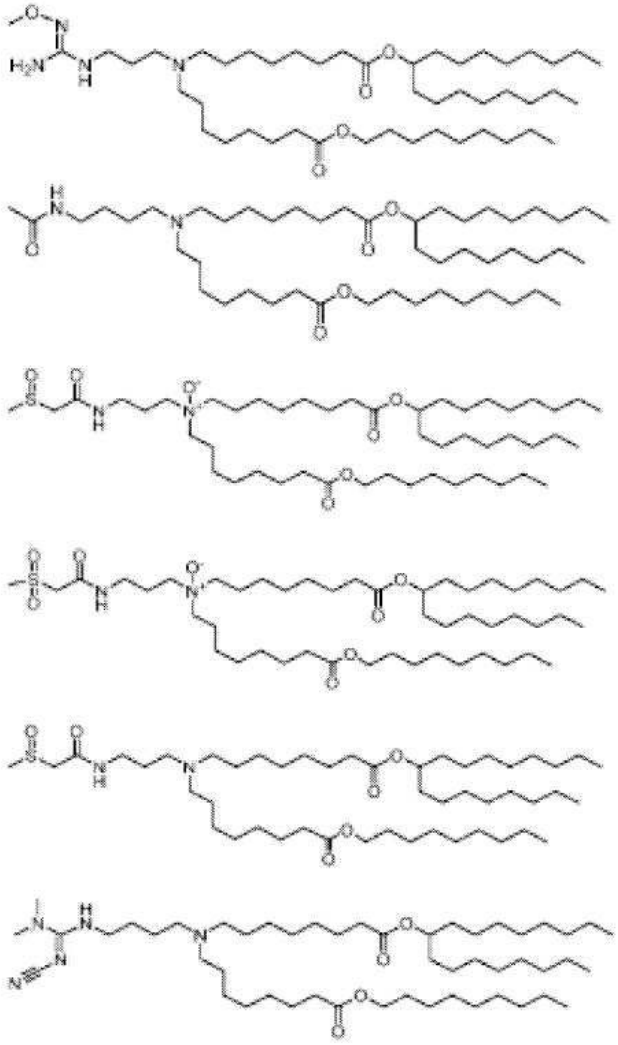
[1347]



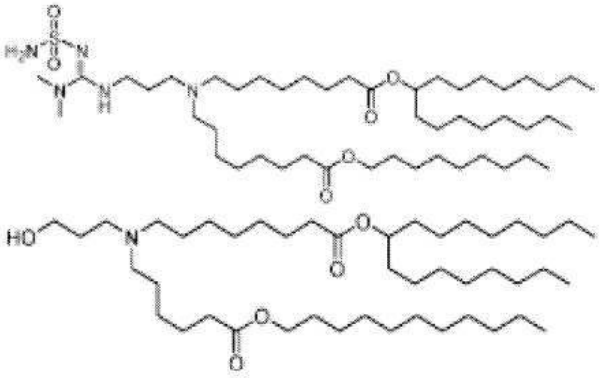
[1348]



[1349]



[1350]



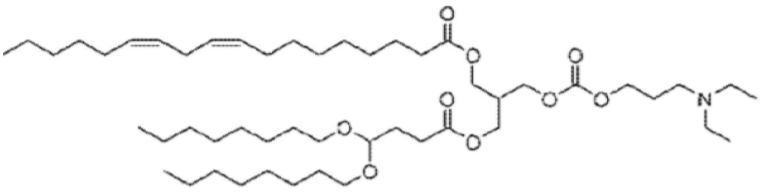
[1351]

[1352] *아민 지질*

[1353] 특정 실시형태에서, 원형 RNA의 전달을 위한 전달 비히클 조성물은 아민 지질을 포함한다. 특정 실시형태에서, 이온화 가능한 지질은 아민 지질이다. 일부 실시형태에서, 아민 지질은 국제 특허 출원 제PCT/US2018/053569호에 기재된 바와 같다.

[1354] 일부 실시형태에서, 아민 지질은 (9Z, 12Z)-3-((4,4-비스(옥틸옥시)부탄오일)옥시)-2-((((3-(다이에틸아미노)프로폭시)카보닐)옥시)메틸)프로필 옥타데카-9, 12-다이엔오에이트인 지질 E이다.

[1355] 지질 E는 하기로서 도시될 수 있다:

[1356] 

[1357] 지질 E는 국제 공개 제WO2015/095340호(예를 들어, 84 내지 86 페이지)에 따라서 합성될 수 있다. 특정 실시형태에서, 아민 지질은 지질 E에 대한 등가물이다.

[1358] 특정 실시형태에서, 아민 지질은 지질 E의 유사체이다. 특정 실시형태에서, 지질 E 유사체는 지질 E의 아세탈 유사체이다. 특정 전달 비히클 조성물에서, 아세탈 유사체는 C4-C12 아세탈 유사체이다. 일부 실시형태에서, 아세탈 유사체는 C5-C12 아세탈 유사체이다. 추가 실시형태에서, 아세탈 유사체는 C5-C10 아세탈 유사체이다. 추가 실시형태에서, 아세탈 유사체는 C4, C5, C6, C7, C9, C10, C11 및 C12 아세탈 유사체로부터 선택된다.

[1359] 전달 비히클, 예를 들어, 본 명세서에 기재된 지질 나노입자에서 사용하기에 적합한 아민 지질 및 다른 생분해성 지질은 생체내에서 생분해성이다. 본 명세서에 기재된 아민 지질은 낮은 독성을 갖는다(예를 들어, 10mg/kg 이상의 양에서 유해 효과 없이 동물 모델에서 허용됨). 특정 실시형태에서, 아민 지질을 포함하는 전달 비히클은 아민 지질 중 적어도 75%가 8, 10, 12, 24 또는 48시간, 또는 3, 4, 5, 6, 7 또는 10일 이내에 혈장으로부터 제거되는 것을 포함한다.

[1360] 생분해성 지질은 예를 들어, WO2017/173054, WO2015/095340 및 WO2014/136086의 생분해성 지질을 포함한다.

[1361] 지질 제거율은 당업자에게 공지된 방법에 의해서 측정될 수 있다. 예를 들어, 문헌[Maier, M.A., *et al.* Biodegradable Lipids Enabling Rapidly Eliminated Lipid Nanoparticles for Systemic Delivery of RNAi Therapeutics. Mol. Ther. 2013, 21(8), 1570-78] 참조.

[1362] 아민 지질을 포함하는 전달 비히클 조성물은 증가된 제거율을 초래할 수 있다. 일부 실시형태에서, 제거율은 지질 제거율, 예를 들어, 지질이 혈액, 혈청 또는 혈장으로부터 제거되는 속도이다. 일부 실시형태에서, 제거율은 RNA 제거율, 예를 들어, circRNA가 혈액, 혈청 또는 혈장으로부터 제거되는 속도이다. 일부 실시형태에서, 제거율은 전달 비히클이 혈액, 혈청 또는 혈장으로부터 제거되는 속도이다. 일부 실시형태에서, 제거율은 조직, 예컨대, 간 조직 또는 비장 조직으로부터 전달 비히클이 제거되는 속도이다. 특정 실시형태에서, 높은 제거율은 실질적인 유해 효과가 없는 안전성 프로파일을 유도한다. 아민 지질 및 생분해성 지질은 순환계 및 조직에서 전달 비히클 축적을 감소시킬 수 있다. 일부 실시형태에서, 순환계 및 조직에서의 전달 비히클 축적의 감소는 실질적인 유해 효과 없이 안전성 프로파일을 유도한다.

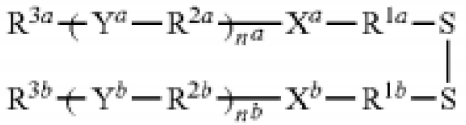
[1363] 지질은 이들이 있는 매질의 pH에 따라 이온화될 수 있다. 예를 들어, 약산성 매질에서, 아민 지질과 같은 지질은 양성자화되어 양전하를 가질 수 있다. 반대로, pH가 대략 7.35인 혈액과 같은 약 염기성인 매질에서는 아민 지질과 같은 지질이 양성자화되지 않아 전하를 띠지 않을 수 있다.

[1364] 전하를 갖는 지질의 능력은 이의 고유 pKa와 관련이 있다. 일부 실시형태에서, 본 개시내용의 아민 지질은 각각 독립적으로 약 5.1 내지 약 7.4 범위의 pKa를 가질 수 있다. 일부 실시형태에서, 본 개시내용의 생체 이용 가능한 지질은 각각 독립적으로 약 5.1 내지 약 7.4 범위의 pKa를 가질 수 있다. 예를 들어, 본 개시내용의 아민 지질은 각각 독립적으로 약 5.8 내지 약 6.5 범위의 pKa를 가질 수 있다. 약 5.1 내지 약 7.4 범위의 pKa를 갖는 지질은 생체내, 예를 들어, 간으로의 카고 전달에 효과적이다. 또한, 약 5.3 내지 약 6.4 범위의 pKa를 갖는 지질이 생체내, 예를 들어, 종양으로의 전달에 효과적인 것으로 밝혀져 있다. 예를 들어, 국제 특허 제WO2014/136086호 참조.

[1365] ***이황화 결합을 함유하는 지질***

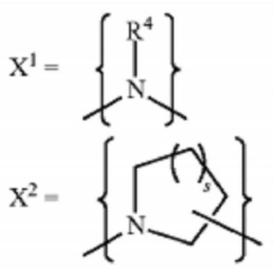
[1366] 일부 실시형태에서, 이온화 가능한 지질은 미국 특허 제9,708,628호에 기재되어 있다.

[1367] 본 발명은 구조식 (XXII)에 의해 표현된 지질을 제공한다:



[1368]

[1369] 구조식 (XXII)에서, X^a 및 X^b는 각각 독립적으로 하기에 도시된 X^1 또는 X^2이다.



[1370]

[1371] X^1에서 R^4는 1 내지 6개의 탄소 원자를 갖는 알킬기이며, 선형, 분지형 또는 사이클릭일 수 있다. 알킬기는 바람직하게는 1 내지 3의 탄소수를 갖는다. 1 내지 6개의 탄소 원자를 갖는 알킬기의 구체예는 메틸기, 에틸기, 프로필기, 아이소프로필기, n-부틸기, sec-부틸기, 아이소부틸기, tert-부틸기, 펜틸기, 아이소펜틸기, 네오펜틸기, t-펜틸기, 1,2-다이메틸프로필기, 2-메틸부틸기, 2-메틸펜틸기, 3-메틸펜틸기, 2,2-다이메틸부틸기, 2,3-다이메틸부틸기, 사이클로헥실기 등을 포함한다. R^4는 바람직하게는 메틸기, 에틸기, 프로필기 또는 아이소프로필기이고, 가장 바람직하게는 메틸기이다.

[1372] X^2에서 s는 1 또는 2이다. s가 1인 경우, X^2는 피롤리디늄기이고, s가 2인 경우, X^2는 피페리디늄기이다. s는 바람직하게는 2이다. X^2의 결합 방향은 제한되지 않지만, X^2의 질소 원자는 바람직하게는 R^{1a} 및 R^{1b}에 결합한다.

[1373] X^a는 X^b와 동일하거나 상이할 수 있고, X^a는 바람직하게는 X^b와 동일한 기이다.

[1374] n^a 및 n^b는 각각 독립적으로 0 또는 1, 바람직하게는 1이다. n^a가 1인 경우, R^{3a}는 Y^a 및 R^{2a}를 통해 X^a에 결합하고, n^a가 0인 경우에는 R^{3a}-X^a-R^{1a}-S-의 구조를 취한다. 유사하게, n^b가 1일 경우, R^{3b}는 Y^b 및 R^{2b}를 통해 X^b에 결합하고, n^b가 0인 경우, R^{3b}-X^b-R^{1b}-S-의 구조를 취한다.

[1375] n^a는 n^b와 동일하거나 상이할 수 있고, n^a는 바람직하게는 n^b와 동일하다.

[1376] R^{1a} 및 R^{1b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이며, 선형 또는 분지형일 수 있으며, 바람직하게는 선형일 수 있다. 1 내지 6개의 탄소 원자를 갖는 알킬렌기의 구체예는 메틸렌기, 에틸렌기, 트라이메틸렌기, 아이소프로필렌기, 테트라메틸렌기, 아이소부틸렌기, 펜타메틸렌기, 네오펜틸렌기 등을 포함한다. R^{1a} 및 R^{1b}는 각각 바람직하게는 메틸렌기, 에틸렌기, 트라이메틸렌기, 아이소프로필렌기 또는 테트라메틸렌기이고, 에틸렌기가 가장 바람직하다.

[1377] R^{1a}는 R^{1b}와 동일하거나 상이할 수 있고, R^{1a}는 바람직하게는 R^{1b}와 동일한 기이다.

[1378] R^{2a} 및 R^{2b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이며, 선형 또는 분지형일 수 있으며, 바람직하게는 선형일 수 있다. 1 내지 6개의 탄소 원자를 갖는 알킬렌기의 예는 R^{1a} 또는 R^{1b}에 대해서 1 내지 6개의 탄소 원자를 갖는 알킬렌기의 예로서 열거된 것을 포함한다. R^{2a} 및 R^{2b}는 각각 바람직하게는 메틸렌기, 에틸렌기, 트라이메틸렌기, 아이소프로필렌기 또는 테트라메틸렌기이다.

[1379] X^a 및 X^b가 각각 X^1인 경우, R^{2a} 및 R^{2b}는 바람직하게는 트라이메틸렌기이다. X^a 및 X^b가 각각 X^2인 경우, R^{2a} 및 R^{2b}는 바람직하게는 에틸렌기이다.

[1380] R^{2a}는 R^{2b}와 동일하거나 상이할 수 있고, R^{2a}는 바람직하게는 R^{2b}와 동일한 기이다.

[1381] Y^a 및 Y^b는 각각 독립적으로 에스터 결합, 아마이드 결합, 카바메이트 결합, 에터 결합 또는 유레아 결합이고, 바람직하게는 에스터 결합, 아마이드 결합 또는 카바메이트 결합이고, 가장 바람직하게는 에스터 결합이다. Y^a 및 Y^b의 결합 방향은 제한되지 않지만, Y^a가 에스터 결합인 경우에는 R^{3a}-CO-O-R^{2a}-의 구조가 바람직하고, Y^b가 에스터 결합인 경우에는 R^{3b}-CO-O-R^{2b}-의 구조가 바람직하다.

[1382] Y^a는 Y^b와 동일하거나 상이할 수 있고, Y^a는 바람직하게는 Y^b와 동일한 기이다.

[1383] R^{3a} 및 R^{3b}는 각각 독립적으로 스테롤 잔기, 지용성 비타민 잔기 또는 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소기, 바람직하게는 지용성 비타민 잔기 또는 12 내지 22의 탄소 원자를 갖는 지방족 탄화수소기, 가장 바람직하게는 지용성 비타민 잔기이다.

[1384] 스테롤 잔기의 예는 콜레스테릴기(콜레스테롤 잔기), 콜레스테릴기(콜레스탄올 잔기), 스티그마스테릴기(스티그마스테롤 잔기), β-시토스테릴기(β-시토스테롤 잔기), 라노스테릴기(라노스테롤 잔기), 및 에르고스테릴기(에르고스테롤 잔기) 등을 포함한다. 스테롤 잔기는 바람직하게는 콜레스테릴기 또는 콜레스테아릴기이다.

[1385] 지용성 비타민 잔기로서, 지용성 비타민으로부터 유래된 잔기뿐만 아니라 지용성 비타민의 작용기인 하이드록실기, 알데하이드 또는 카복실산을 다른 반응성 작용기로 적절히 전환시켜 얻어지는 유도체로부터 유래된 잔기가 사용될 수 있다. 하이드록실기를 갖는 지용성 비타민에 대해서, 예를 들어, 하이드록실기는 석신산 무수물, 글루타르산 무수물 등과 반응시킴으로써 카복실산으로 전환될 수 있다. 지용성 비타민의 예는 레티노산, 레티놀, 레티날, 에르고스테롤, 7-데하이드로콜레스테롤, 칼시페롤, 콜레칼시페롤, 다이하이드로에르고칼시페롤, 다이하이드로타키스테롤, 토코페롤, 토코트라이엔올 등을 포함한다. 지용성 비타민의 바람직한 예는 레티노산 및 토코페롤을 포함한다.

[1386] 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소기는 선형 또는 분지형일 수 있으며, 바람직하게는 선형일 수 있다. 지방족 탄화수소기는 포화 또는 불포화될 수 있다. 불포화 지방족 탄화수소기의 경우, 지방족 탄화수소기는 일반적으로 1 내지 6개, 바람직하게는 1 내지 3개, 보다 더 바람직하게는 1 내지 2개의 불포화 결합을 함유한다. 불포화 결합은 탄소-탄소 이중 결합 및 탄소-탄소 삼중 결합을 포함하지만, 바람직하게는 탄소-탄소 이중 결합이다. 지방족 탄화수소기는 바람직하게는 12 내지 18개, 가장 바람직하게는 13 내기 17개의 탄소수를 갖는다. 지방족 탄화수소기는 알킬기, 알켄일기, 알킨일기 등을 포함하지만, 알킬기 또는 알켄일기가 바람직하다. 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소기의 구체적인 예는 도데실기, 트라이데실기, 테트라데실기, 펜타데실기, 헥사데실기, 헵타데실기, 옥타데실기, 노나데실기, 아이코실기, 헤니코실기, 도코실기, 도데센일기, 트라이데센일기, 테트라데센일기, 펜타데센일기, 헥사데센일기, 헵타데센일기, 옥타데센일기, 노나데센일기, 아이코센일기, 헤니코센일기, 도코센일기, 데카다이엔일기, 트라이데카다이엔일기, 테트라데카다이엔일기, 펜타데카다이엔일기, 헥사데카다이엔일기, 헵타데카다이엔일기, 옥타데카다이엔일기, 노나데카다이엔일기, 아이코사다이엔일기, 헤니코사다이엔일기, 도코사다이엔일기, 옥타데카트라이엔일기, 아이코사트라이엔일기, 아이코사테트라엔일기, 아이코사펜타엔일기, 도코사헥사엔일기, 아이소스테아릴기 등을 포함한다. 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소기는 바람직하게는 트라이데실기, 테트라데실기, 헵타데실기, 옥타데실기, 헵타데카다이엔일기 또는 옥타데카다이엔일기, 특히 바람직하게는 트라이데실기, 헵타데실기 또는 헵타데카다이엔일기이다.

[1387] 일 실시형태에서, 지방산, 지방족 알코올 또는 지방족 아민으로부터 유래된 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소기가 사용된다. R^{3a}(또는 R^{3b})가 지방산으로부터 유래된 경우, Y^a(또는 Y^b)는 에스터 결합 또는 아마이드 결합이고, 지방산-유래된 카보닐 탄소가 Y^a(또는 Y^b)에 포함된다. 예를 들어, 리놀레산 이 사용되는 경우, R^{3a}(또는 R^{3b})는 헵타데카다이엔일기이다.

[1388] R^{3a}는 R^{3b}와 동일하거나 상이할 수 있고, R^{3a}는 바람직하게는 R^{3b}와 동일한 기이다.

[1389] 일 실시형태에서, X^a는 X^b와 동일하고, n^a는 n^b와 동일하고, R^{1a}는 R^{1b}와 동일하고, R^{2a}는 R^{2b}와 동일하고, R^{3a}는 R^{3b}와 동일하고, Y^a는 Y^b와 동일하다.

[1390] 일 실시형태에서,

[1391] X^a 및 X^b는 각각 독립적으로 X1이고,

[1392] R^4는 1 내지 3개의 탄소 원자를 갖는 알킬기이고, n^a 및 n^b는 각각 1이고,

[1393] R^{1a} 및 R^{1b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1394] R^{2a} 및 R^{2b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1395] Y^a 및 Y^b는 각각 에스터 결합 또는 아마이드 결합이고,

[1396] R^{3a} 및 R^{3b}는 각각 독립적으로 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소이다.

[1397] 일 실시형태에서,

[1398] X^a 및 X^b는 각각 X1이고,

[1399] R^4는 1 내지 3개의 탄소 원자를 갖는 알킬기이고, n^a 및 n^b는 각각 1이고,

[1400] R^{1a} 및 R^{1b}는 각각 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1401] R^{2a} 및 R^{2b}는 각각 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1402] Y^a 및 Y^b는 각각 에스터 결합 또는 아마이드 결합이고,

[1403] R^{3a} 및 R^{3b}는 각각 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소이고,

[1404] X^a는 X^b와 동일하고,

[1405] R^{1a}는 R^{1b}와 동일하고,

[1406] R^{2a}는 R^{2b}와 동일하고,

[1407] R^{3a}는 R^{3b}와 동일하다.

[1408] 일 실시형태에서,

[1409] X^a 및 X^b는 각각 X^1이고,

[1410] R^4는 메틸기이고, n^a 및 n^b는 각각 1이고,

[1411] R^{1a} 및 R^{1b}는 각각 에틸렌기이고,

[1412] R^{2a} 및 R^{2b}는 각각 트라이메틸렌기이고,

[1413] Y^a 및 Y^b는 각각 -CO-O-이고,

[1414] R^{3a} 및 R^{3b}는 각각 독립적으로 13 내지 17개의 탄소 원자를 갖는 알킬기 또는 알켄일기이다.

[1415] 일 실시형태에서,

[1416] X^a 및 X^b는 각각 X^1이고,

[1417] R^4는 메틸기이고, n^a 및 n^b는 각각 1이고,

[1418] R^{1a} 및 R^{1b}는 각각 에틸렌기이고,

[1419] R^{2a} 및 R^{2b}는 각각 트라이메틸렌기이고,

[1420] Y^{a} 및 Y^{b}는 각각 -CO-O-이고,

[1421] R^{3a} 및 R^{3b}는 각각 13 내지 17개의 탄소 원자를 갖는 알킬기 또는 알켄일기이다.

[1422] R^{3a}는 R^{3b}와 동일하다.

[1423] 일 실시형태에서,

[1424] X^{a} 및 X^{b}는 각각 독립적으로 X^{1}이고,

[1425] R^{4}는 1 내지 3개의 탄소 원자를 갖는 알킬기이고, n^{a} 및 n^{b}는 각각 1이고,

[1426] R^{1a} 및 R^{1b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1427] R^{2a} 및 R^{2b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1428] Y^{a} 및 Y^{b}는 각각 에스터 결합 또는 아마이드 결합이고,

[1429] R^{3a} 및 R^{3b}는 각각 독립적으로 지용성 비타민 잔기(예를 들어, 레틴산 잔기, 토코페롤 잔기)이다.

[1430] 일 실시형태에서,

[1431] X^{a} 및 X^{b}는 각각 X^{1}이고,

[1432] R^{4}는 1 내지 3개의 탄소 원자를 갖는 알킬기이고, n^{a} 및 n^{b}는 각각 1이고,

[1433] R^{1a} 및 R^{1b}는 각각 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1434] R^{2a} 및 R^{2b}는 각각 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1435] Y^{a} 및 Y^{b}는 각각 에스터 결합 또는 아마이드 결합이고,

[1436] R^{3a} 및 R^{3b}는 각각 지용성 비타민 잔기(예를 들어, 레틴산 잔기, 토코페롤 잔기)이다.

[1437] X^{a}는 X^{b}와 동일하고,

[1438] R^{1a}는 R^{1b}와 동일하고,

[1439] R^{2a}는 R^{2b}와 동일하고,

[1440] R^{3a}는 R^{3b}와 동일하다.

[1441] 일 실시형태에서,

[1442] X^{a} 및 X^{b}는 각각 X^{1}이고,

[1443] R^{4}는 메틸기이고, n^{a} 및 n^{b}는 각각 1이고,

[1444] R^{1a} 및 R^{1b}는 각각 에틸렌기이고,

[1445] R^{2a} 및 R^{2b}는 각각 트라이메틸렌기이고,

[1446] Y^{a} 및 Y^{b}는 각각 -CO-O-이고,

[1447] R^{3a} 및 R^{3b}는 각각 독립적으로 지용성 비타민 잔기(예를 들어, 레틴산 잔기, 토코페롤 잔기)이다.

[1448] 일 실시형태에서,

[1449] X^a 및 X^b는 각각 X^1이고,

[1450] R^4는 메틸기이고, n^a 및 n^b는 각각 1이고,

[1451] R^{1a} 및 R^{1b}는 각각 에틸렌기이고,

[1452] R^{2a} 및 R^{2b}는 각각 트라이메틸렌기이고,

[1453] Y^a 및 Y^b는 각각 -CO-O-이고,

[1454] R^{3a} 및 R^{3b}는 각각 지용성 비타민 잔기(예를 들어, 레틴산 잔기, 토코페롤 잔기)이고,

[1455] R^{3a}는 R^{3b}와 동일하다.

[1456] 일 실시형태에서,

[1457] X^a 및 X^b는 각각 독립적으로 X^2이고,

[1458] t는 2이고,

[1459] R^{1a} 및 R^{1b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1460] R^{2a} 및 R^{2b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1461] Y^a 및 Y^b는 각각 에스터 결합이고,

[1462] R^{3a} 및 R^{3b}는 각각 독립적으로 지용성 비타민 잔기(예를 들어, 레틴산 잔기, 토코페롤 잔기) 또는 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소기(예를 들어, 12 내지 22개의 탄소 원자를 갖는 알킬기)이다.

[1463] 일 실시형태에서,

[1464] X^a 및 X^b는 각각 독립적으로 X^2이고,

[1465] t는 2이고,

[1466] R^{1a} 및 R^{1b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1467] R^{2a} 및 R^{2b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1468] Y^a 및 Y^b는 각각 에스터 결합이고,

[1469] R^{3a} 및 R^{3b}는 각각 독립적으로 지용성 비타민 잔기(예를 들어, 레틴산 잔기, 토코페롤 잔기) 또는 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소기(예를 들어, 12 내지 22개의 탄소 원자를 갖는 알킬기)이고,

[1470] X^a는 X^b와 동일하고,

[1471] R^{1a}는 R^{1b}와 동일하고,

[1472] R^{2a}는 R^{2b}와 동일하고,

[1473] R^{3a}는 R^{3b}와 동일하다.

[1474] 일 실시형태에서,

[1475] X^a 및 X^b는 각각 독립적으로 X^2이고,

[1476] t는 2이고,

[1477] R^{1a} 및 R^{1b}는 각각 에틸렌기이고,

[1478] R^{2a} 및 R^{2b}는 각각 독립적으로 1 내지 6개의 탄소 원자를 갖는 알킬렌기이고,

[1479] Y^{a} 및 Y^{b}는 각각 에스터 결합이고,

[1480] R^{3a} 및 R^{3b}는 각각 독립적으로 지용성 비타민 잔기(예를 들어, 레틴산 잔기, 토코페롤 잔기) 또는 12 내지 22개의 탄소 원자를 갖는 지방족 탄화수소기(예를 들어, 12 내지 22개의 탄소 원자를 갖는 알킬기)이고,

[1481] X^{a}는 X^{b}와 동일하고,

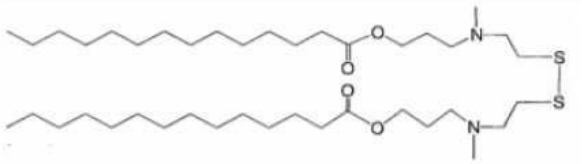
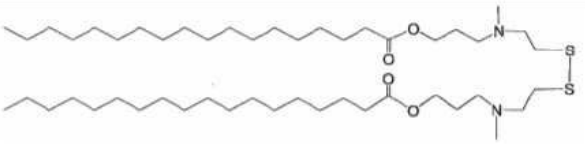
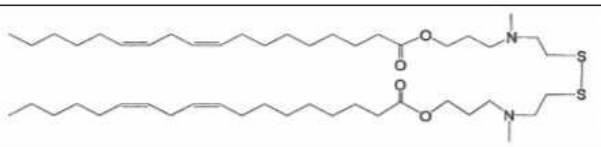
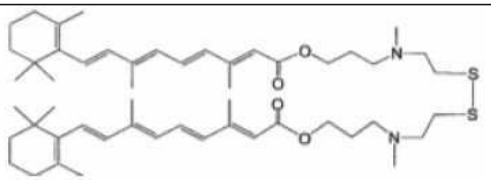
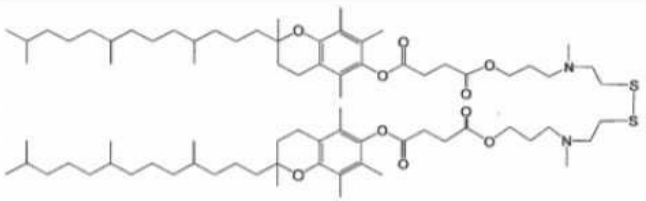
[1482] R^{2a}는 R^{2b}와 동일하고,

[1483] R^{3a}는 R^{3b}와 동일하다.

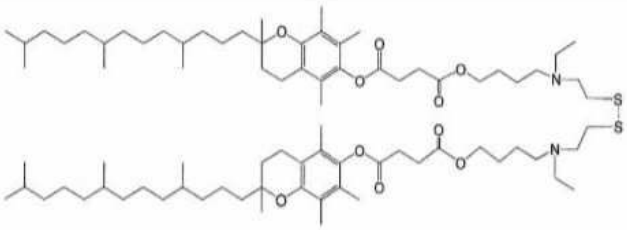
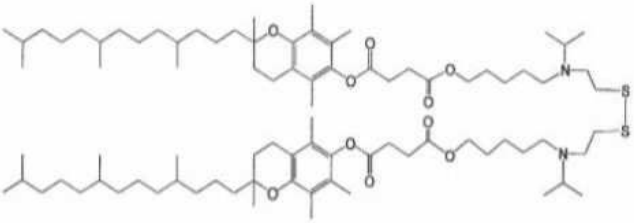
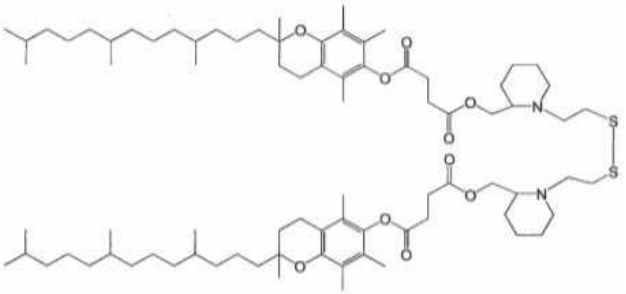
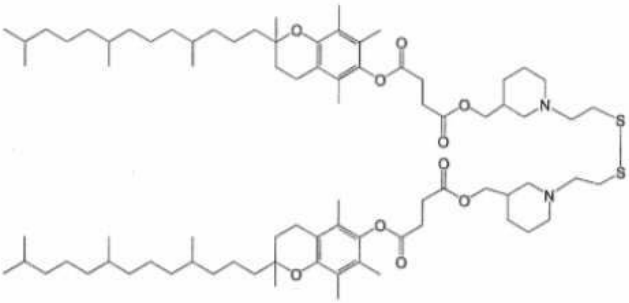
[1484] 일부 실시형태에서, 이온화 가능한 지질은 하기 표 15b에 제시된 구조 중 하나를 갖는다.

[1485] [표 15b]

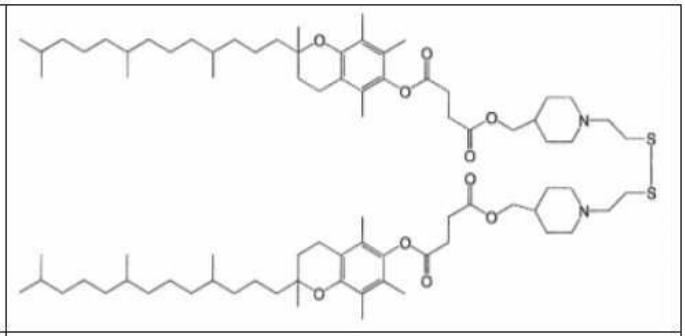
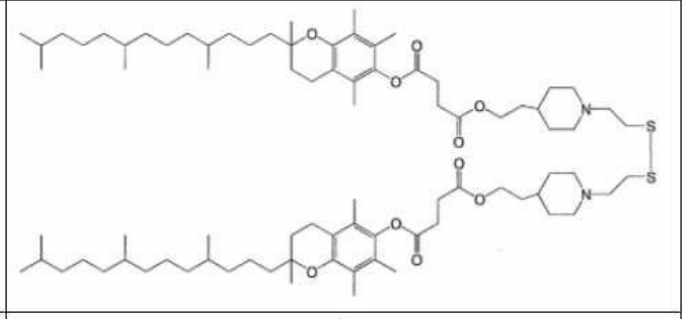
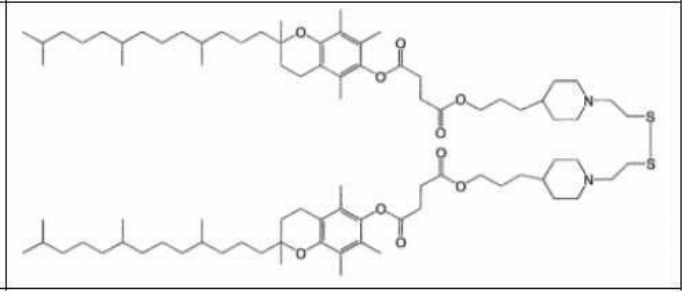
[1486]

번호	구조
1	
2	
3	
4	
5	

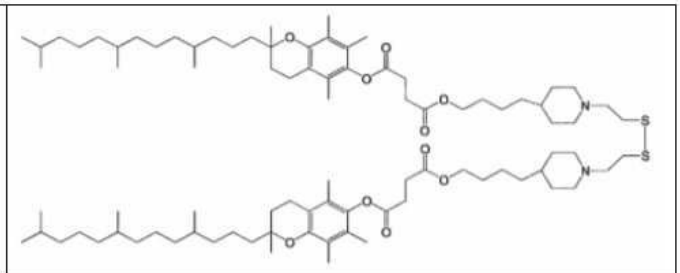
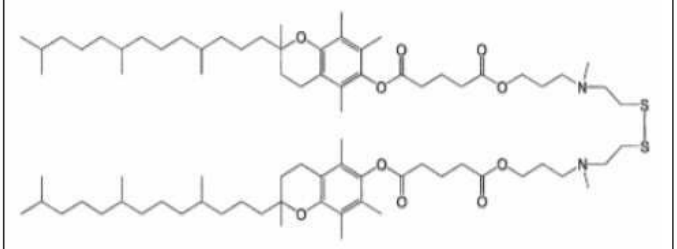
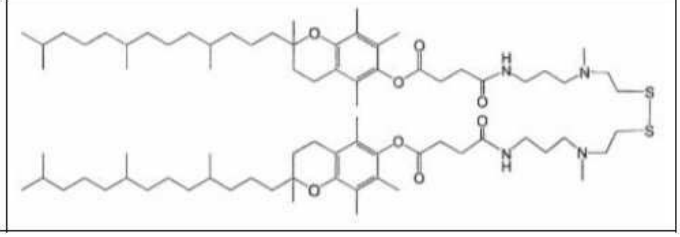
[1487]

6	
7	
8	
9	

[1488]

10	
11	
12	

[1489]

13	
14	
15	

[1490] 본 발명의 지질은 -S-S-(이황화) 결합을 가질 수 있다. 상기 화합물의 제조 방법은 예를 들어,

[1491] $R^{3a}-(Y^{a}-R^{2a})n^{a}-X^{a}-R^{1a}-SH$, 및

[1492] $R^{3b}-(Y^{b}-R^{2b})n^{b}-X^{b}-R^{1b}-SH$

[1493] 를 제조하는 단계, 및 이들을 산화(커플링)시켜 -S-S-를 함유하는 화합물을 제공하는 단계를 포함하는 방법 및 -S-S- 결합을 함유하는 화합물에 필요한 부분을 순차적으로 결합시켜 최종적으로 본 발명의 화합물을 얻는 단계를 포함하는 방법 등을 포함한다. 바람직한 것은 후자 방법이다.

[1494] 후자 방법의 구체적인 예는 하기에 제시되어 있지만, 이것은 제한으로 간주되지 않는다.

[1495] 출발 화합물의 예는 -S-S- 결합을 함유하는 2개의 말단 카복실산, 2개의 말단 카복실레이트, 2개의 말단 아민, 2개의 말단 아이소사이아네이트, 2개의 말단 알코올, 이탈기, 예컨대, MsO(메실레이트기) 등을 갖는 2개의 말단 알코올 및 이탈기, 예컨대, pNP(p-나이트로페닐카보네이트기)를 갖는 2개의 말단 카보네이트 등을 포함한다.

[1496] 예를 들어, X^{a} 및 X^{b}의 경우 X^{1} 또는 X^{2}를 함유하는 화합물이 생성되는 경우, -S-S- 결합을 함유하는 화합물 (1)의 2개의 말단 작용기를 말단에 하나의 작용기 및 -NH-기를 갖는 화합물 (2)의 -NH-기와 반응시키고, 그 반응에 기여하지 않은 화합물 (2)의 말단의 작용기를 R^{3}을 함유한 화합물 (3)의 작용기와 반응시키고, 이에 의해서 -S-S- 결합, R^{1a} 및 R^{1b}, X^{a} 및 X^{b}, R^{2a} 및 R^{2b}, Y^{a} 및 Y^{b}, 및 R^{3a} 및 R^{3b}를 함유하는 본 발명의 화합물을 얻을 수 있다.

[1497] 화합물 (1)과 화합물 (2)의 반응에서, 알칼리 촉매, 예컨대, 탄산칼륨, 탄산나트륨, 칼륨 t-부톡사이드 등을 촉매로 사용할 수 있거나, 촉매 없이 반응을 수행할 수 있다. 바람직하게는, 탄산칼륨 또는 탄산나트륨이 촉매로 사용된다.

[1498] 촉매의 양은 화합물 (1)에 대해 0.1 내지 100몰 당량, 바람직하게는 0.1 내지 20몰 당량, 보다 바람직하게는 0.1 내지 5몰 당량이다. 충전될 화합물 (2)의 양은 화합물 (1)에 대해 1 내지 50몰 당량, 바람직하게는 1 내지 10몰 당량이다.

[1499] 화합물 (1)과 화합물 (2)의 반응에 사용하는 용매는 반응을 저해하지 않는 용매 또는 수용액이면 특별히 제한되지 않는다. 예를 들어, 에틸 아세테이트, 다이클로로메탄, 클로로폼, 벤젠, 톨루엔 등을 언급할 수 있다. 그 중에서도 톨루엔, 클로로폼이 바람직하다.

[1500] 반응 온도는 -20 내지 200℃, 바람직하게는 0 내지 80℃, 보다 바람직하게는 20 내지 50℃이고, 반응 시간은 1 내지 48시간, 바람직하게는 2 내지 24시간이다.

[1501] 화합물 (1)과 화합물 (2)의 반응 생성물을 화합물 (3)과 반응시키는 경우, 알칼리 촉매, 예컨대, 탄산칼륨, 탄산나트륨, 칼륨 t-부톡사이드 등 또는 산 촉매, 예컨대, PTS(p-톨루엔설폰산), MSA(메탄설폰산) 등, 예컨대, 화합물 (1)과 화합물 (2)의 반응에 사용되는 촉매를 사용할 수 있거나, 촉매 없이 반응을 수행할 수 있다.

[1502] 또한, 화합물 (1) 및 화합물 (2)의 반응 생성물을 축합제, 예컨대, DCC(다이사이클로헥실카보다이이미드), DIC(다이아이소프로필카보다이이미드), EDC(1-에틸 -3-(3-다이메틸아미노프로필)카보다이이미드 염산염) 등을 사용함으로써 화합물 (3)과 직접 반응시킬 수 있다. 대안적으로는, 화합물 (3)을 축합제로 처리하여 일단 무수물 등으로 전환시키고, 그 후 그것을 화합물 (1)과 화합물 (2)의 반응 생성물과 반응시킬 수 있다.

[1503] 충전될 화합물 (3)의 양은 화합물 (1)과 화합물 (2)의 반응 생성물에 대해 1 내지 50몰 당량, 바람직하게는 1 내지 10몰 당량이다.

[1504] 사용될 촉매는 반응될 작용기에 따라서 적절하게 선택된다.

[1505] 촉매의 양은 화합물 (1)에 대해 0.05 내지 100몰 당량, 바람직하게는 0.1 내지 20몰 당량, 보다 바람직하게는 0.2 내지 5몰 당량이다.

[1506] 화합물 (1)과 화합물 (2)의 반응 생성물과 화합물 (3)의 반응에 사용하는 용매는 반응을 저해하지 않는 용매 또는 수용액이면 특별히 제한되지 않는다. 예를 들어, 에틸 아세테이트, 다이클로로메탄, 클로로폼, 벤젠, 톨루엔 등을 언급할 수 있다. 그 중에서도 톨루엔, 클로로폼이 바람직하다.

[1507] 반응 온도는 0 내지 200℃, 바람직하게는 0 내지 120℃, 보다 바람직하게는 20 내지 50℃이고, 반응 시간은 1시간 내지 48시간, 바람직하게는 2 내지 24시간이다.

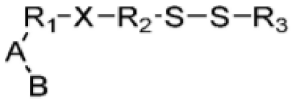
[1508] 상기에 언급된 반응에 의해 수득된 반응 생성물은 일반적인 정제 방법, 예를 들어, 물로의 세척, 실리카겔 칼럼

크로마토그래피, 결정화, 재결정화, 액체-액체 추출, 재침전, 이온 교환 칼럼 크로마토그래피 등에 의해 적절하게 정제될 수 있다.

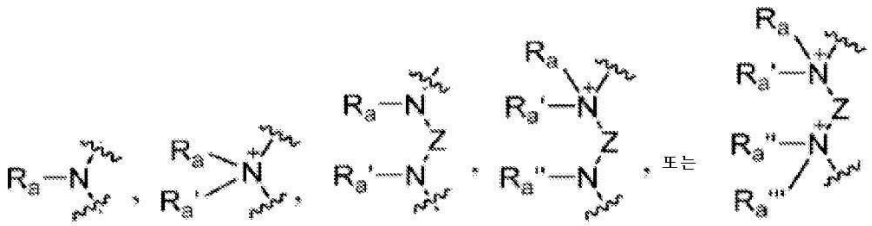
[1509] *구조식 XXIII 지질*

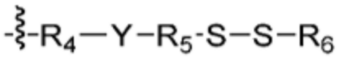
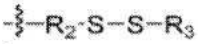
[1510] 일부 실시형태에서, 이온화 가능한 지질은 미국 특허 제9,765,022호에 기재되어 있다.

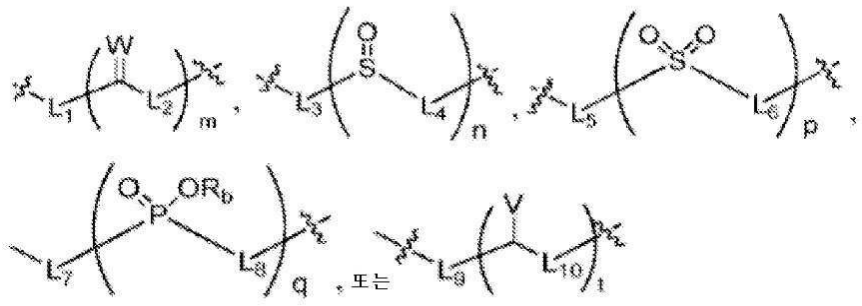
[1511] 본 발명은 구조식 (XXIII)에 의해 표현된 화합물을 제공한다:

[1512] 

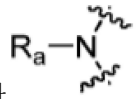
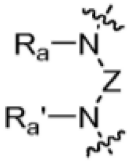
[1513] 구조식 XXIII에서, 친수성 및 선택적으로 양으로 하전된 헤드는 다음과 같다:

[1514] 

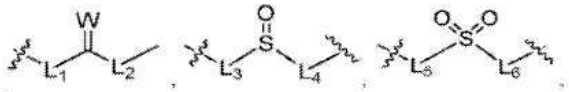
[1515] 식 중, R_a, R_a', R_a'' 및 R_a''' 각각은 독립적으로, H, C_1-C_{20} 1가 지방족 라디칼, C_1-C_{20} 1가 헤테로지방족 라디칼, 1가 아릴 라디칼 또는 1가 헤테로아릴 라디칼이고, Z는 C_1-C_{20} 2가 지방족 라디칼, C_1-C_{20} 2가 헤테로지방족 라디칼, 2가 아릴 라디칼 또는 2가 헤테로아릴 라디칼이고; B는 C_1-C_{24} 1가 지방족 라디칼, C_1-C_{24} 1가 헤테로지방족 라디칼, 1가 아릴 라디칼, 1가 헤테로아릴 라디칼 또는  이고, R_1 및 R_4는 독립적으로, 결합, C_1-C_{10} 2가 지방족 라디칼, C_1-C_{10} 2가 헤테로지방족 라디칼, 2가 아릴 라디칼 또는 2가 헤테로아릴 라디칼이고; R_2 및 R_5 각각은 독립적으로, 결합, C_1-C_{20} 2가 지방족 라디칼, C_1-C_{20} 2가 헤테로지방족 라디칼, 2가 아릴 라디칼 또는 2가 헤테로아릴 라디칼이고; R_3 및 R_6 각각은 독립적으로, C_1-C_{20} 1가 지방족 라디칼, C_1-C_{20} 1가 헤테로지방족 라디칼, 1가 아릴 라디칼 또는 1가 헤테로아릴 라디칼이고;  소수성 테일 및  또한 소수성 테일 각각은 8 내지 24개의 탄소 원자를 갖고; X, 링커 및 Y, 또한 링커 각각은 독립적으로 하기이다:

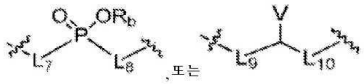
[1516] 

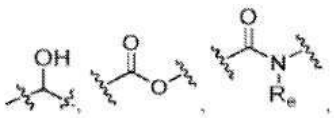
[1517] 식 중, m, n, p, q 및 t 각각은, 독립적으로, 1 내지 6이고; W는 O, S 또는 NR_c이고; R_1, R_2, R_4 또는 R_5에 직접 연결된 L_1, L_3, L_5, L_7 및 L_9는 독립적으로 결합, O, S 또는 NR_d이고; L_2, L_4, L_6, L_8 및 L_{10} 각각은 독립적으로 결합, O, S 또는 NR_e이고; V는 OR_f, SR_g 또는 NR_hR_i이고; R_b, R_c, R_d, R_e, R_f, R_g, R_h 및 R_i 각각은 독립적으로 H, OH, C_{1-10} 옥시지방족 라디칼, C_1-C_{10} 1가 지방족 라디칼, C_1-C_{10} 1가 헤테로지방족 라디칼, 1가 아릴 라디칼 또는 1가 헤테로아릴 라디칼이다.

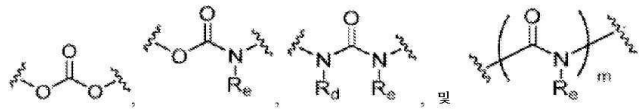
[1518] 상기에 기재된 지질-유사 화합물의 하위세트는 A가  또는  이고, R_a 및 R_a' 각각이 독립적으로, C_1-C_{10} 1가 지방족 라디칼, C_1-C_{10} 1가 헤테로지방족 라디칼, 1가 아릴 라디칼 또는 1가 헤테로아릴 라디칼이고; Z가 C_1-C_{10} 2가 지방족 라디칼, C_1-C_{10} 2가 헤테로지방족 라디칼, 가 아릴 라디칼, 또는 2가 헤테로아릴 라디칼인 것을 포함한다.

[1519] 본 발명의 일부 지질-유사 화합물은 R_1 및 R_4 각각이 독립적으로 C_1-C_6(예를 들어, C_1-C_4) 2가 지방족 라디칼 또는 C_1-C_6(예를 들어, C_1-C_4) 2가 헤테로지방족 라디칼이고, R_2 및 R_3에 대한 총 탄소 수가 12 내지 20(예를 들어, 14 내지 18)개이고, R_5 및 R_6의 총 탄소 수가 또한 12 내지 20(예를 들어, 14 내지 18)개이고, X 및 Y 각각이 독립적으로

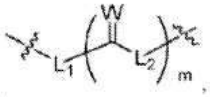
[1520] 

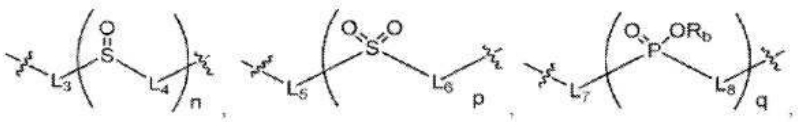
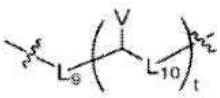
[1521]  인 것을 특징으로 한다.

[1522] X 및 Y의 구체적인 예는 

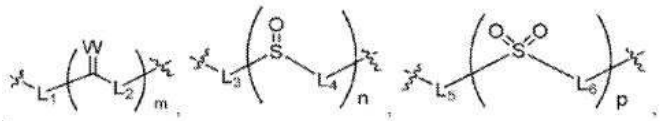
[1523]  을 포함하고, m은 2 내지 6이다.

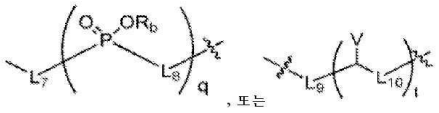
[1524] 추가로 본 발명의 범주 내에는 단백질 및 생체환원성(bioreducible) 화합물로 형성된 나노복합체를 함유하는 약제학적 조성물이 있다. 이러한 약제학적 조성물에서, 나노복합체는 50 내지 500㎚의 입자 크기를 갖고; 생체환원성 화합물은 이황화 소수성 모이어티, 친수성 모이어티 및 이황화 소수성 모이어티와 친수성 모이어티를 연결하는 링커를 함유하고; 단백질은 비-공유 상호작용, 공유 결합 또는 둘 다를 통해 생체 환원성 화합물에 결합한다.

[1525] 특정 실시형태에서, 이황화 소수성 모이어티는 하나 이상의 -S-S-기 및 8 내지 24개의 탄소 원자를 함유하는 헤테로지방족 라디칼이고; 친수성 모이어티는 1개 이상의 친수성 기 및 1 내지 20개의 탄소 원자를 함유하는 지방족 또는 헤테로지방족 라디칼이며, 친수성 기 각각은 아미노, 알킬아미노, 다이알킬아미노, 트라이알킬아미노, 테트라알킬암모늄, 하이드록시아미노, 하이드록실, 카복실, 카복실레이트, 카바메이트, 카바마이드, 카보네이트, 포스페이트, 포스파이트, 설페이트, 설파이트 또는 티오설페이트이고; 링커는 O, S, Si, C_1-C_6 알킬렌, 

[1526]  또는  이고, 식 중 변수는 상기에 정의되어 있다.

[1527] X 및 Y의 구체적인 예는 O, S, Si, C_1-C_6 알킬렌,

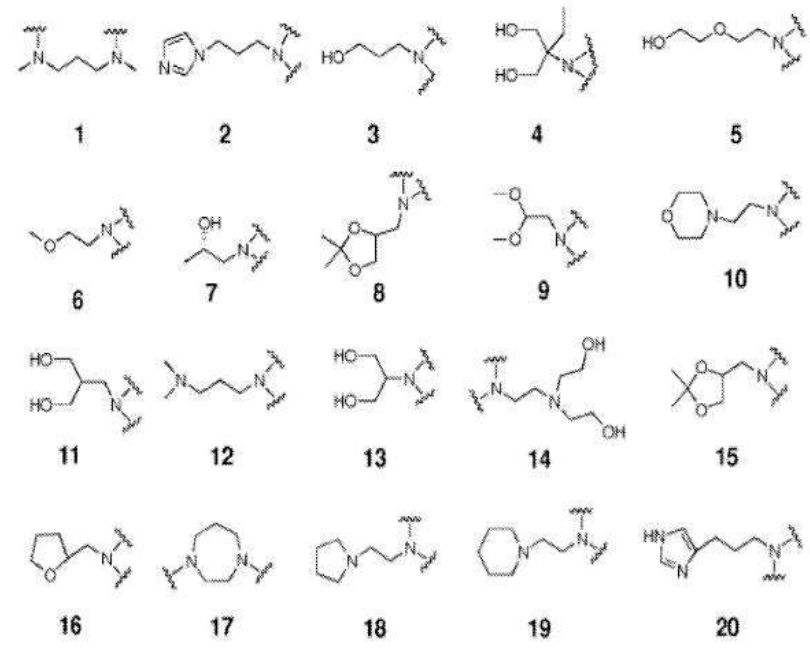
[1528] 

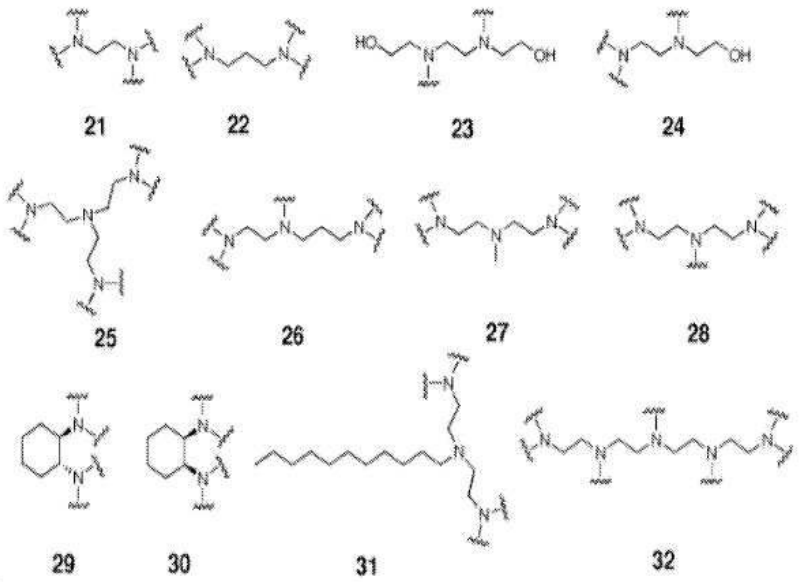
[1529] 
, 또는
을 포함한다.

[1530] 일부 실시형태에서, 상기 구조식 XXIII에 도시된 바와 같은 본 발명의 지질-유사 화합물은 (i) 친수성 헤드, A; (ii) 소수성 테일, R_2-S-S-R_3; 및 (iii) 링커, X를 포함한다. 선택적으로, 이들 화합물은 제2 소수성 테일, R_5-S-S-R_6 및 제2 링커, Y를 함유한다.

[1531] 구조식 XXIII의 친수성 헤드는 하나 이상의 친수성 작용기, 예를 들어, 하이드록실, 카보닐, 카복실, 아미노, 설프하이드릴, 포스페이트, 아마이드, 에스터, 에터, 카바메이트, 카보네이트, 카바마이드 및 포스포다이에스터를 함유한다. 이러한 기는 수소 결합을 형성할 수 있고, 선택적으로 양으로 또는 음으로 하전된다.

[1532] 친수성 헤드의 예는 하기를 포함한다:

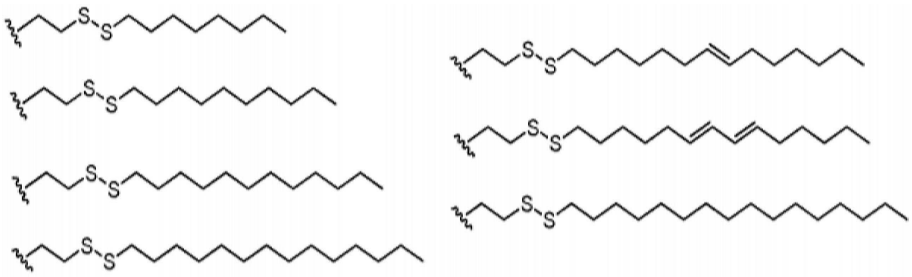
[1533] 

[1534] 

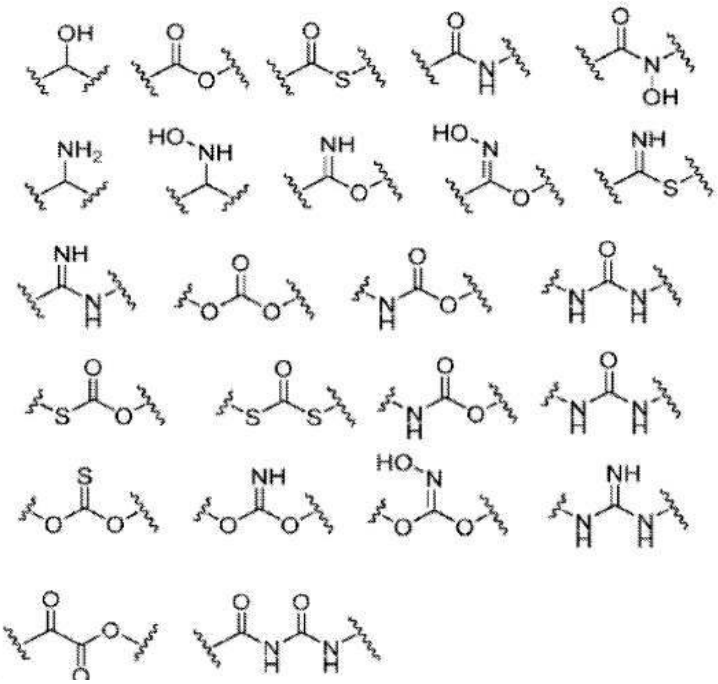
[1535] 다른 예는 문헌[Akinc et al., Nature Biotechnology, 26, 561-69 (2008)] 및 Mahon 등의 미국 특허 출원 공개 제2011/0293703호에 기재된 것을 포함한다.

[1536] 구조식 XXIII의 소수성 테일은 이황화 결합 및 8 내지 24개의 탄소 원자를 함유하는 포화 또는 불포화, 선형 또는 분지형, 비사이클릭 또는 사이클릭, 방향족 또는 비방향족 탄화수소 모이어티이다. 탄소 원자 중 하나 이상은 헤테로원자, 예컨대, N, O, P, B, S, Si, Sb, Al, Sn, As, Se 및 Ge로 대체될 수 있다. 테일은 상기에 기재된 하나 이상의 기로 선택적으로 치환된다. 이러한 이황화 결합을 함유하는 지질-유사 화합물은 생체 환원성일 수 있다.

[1537] 예는 하기를 포함한다:

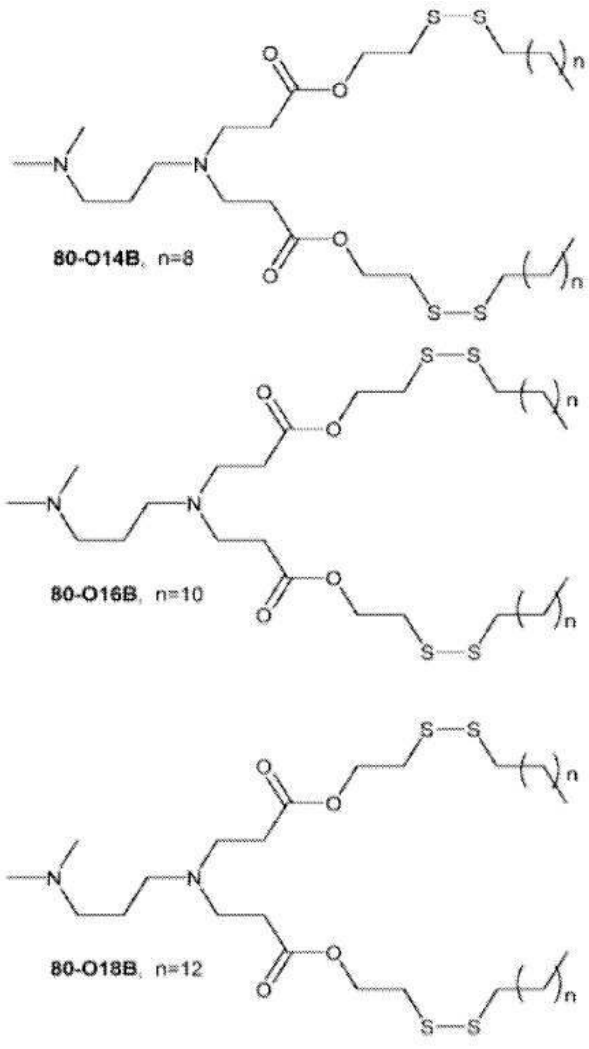
[1538] .

[1539] 구조식 XXIII의 링커는 친수성 헤드 및 소수성 테일을 연결한다. 링커는 친수성 또는 소수성, 극성 또는 비극성, 예를 들어, O, S, Si, 아미노, 알킬렌, 에스터, 아마이드, 카바메이트, 카바마이드, 카보네이트, 포스페이트, 포스파이트, 설페이트, 설파이트 및 티오설페이트인 임의의 화학기일 수 있다. 예는 하기를 포함한다:

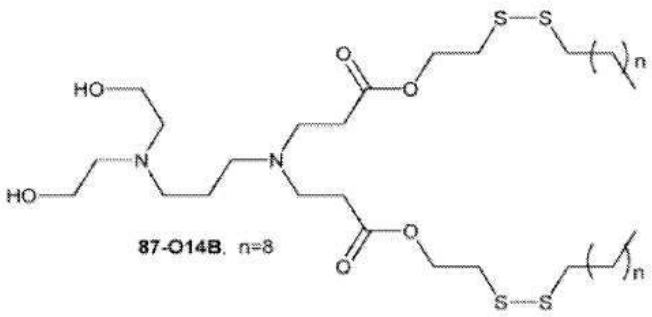


[1540]

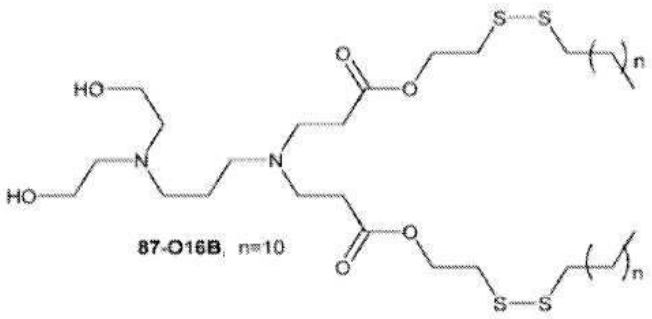
[1541] 본 발명의 예시적인 지질-유사 화합물이 하기에 제시되어 있다:



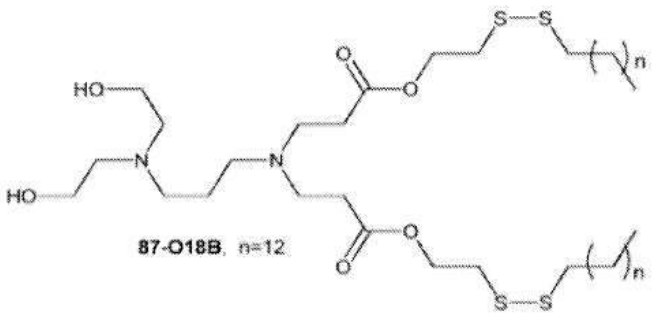
[1542]



87-O14B, n=8

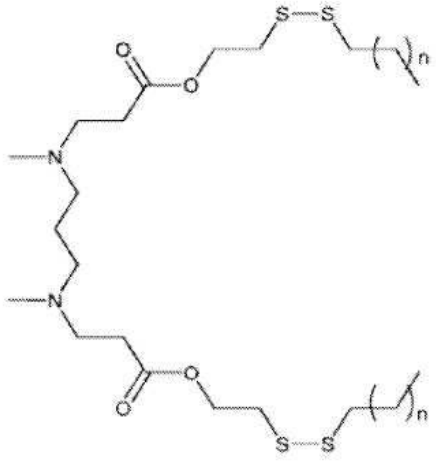


87-O16B, n=10

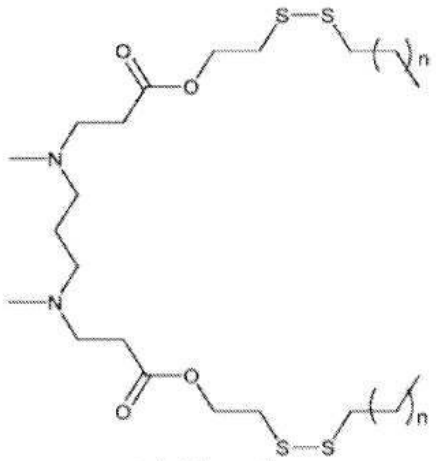


87-O18B, n=12

[1543]



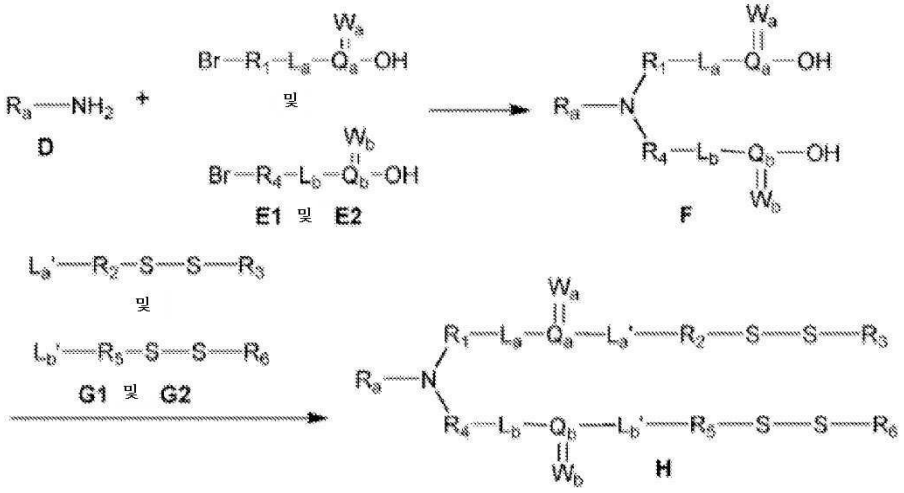
1-O16B, n=10



1-O18B, n=12

[1544] .

[1545] 구조식 XXIII의 지질-유사 화합물은 당업계에 널리 공지된 방법에 의해서 제조될 수 있다. 문헌[Wang et al., ACS Synthetic Biology, 1, 403-07 (2012); Manoharan 등의 국제 특허 출원 공개 제WO 2008/042973호; 및 Zugates 등의 미국 특허 제8,071,082호 참조. 하기에 도시된 경로는 이들 지질-유사 화합물의 합성을 예시한다:



[1546]

[1547] L_a, L_a', L 및 L' 각각은 L_1 내지 L_{10} 중 하나일 수 있고; W_a 및 W_b 각각은 독립적으로 W 또는 V이고; R_a 및 R_1 내지 R_6뿐만 아니라 L_1 내지 L_{10}, W 및 V는 상기에 정의되어 있다.

[1548] 이러한 예시적인 합성 경로에서, 아민 화합물, 즉, 화합물 D를 브로마이드 E1 및 E2와 반응시켜 화합물 F를 형

성하고, 그 다음 이것을 G1 및 G2 둘 다와 커플링시켜 최종 생성물 즉, 화합물 H를 얻는다. 이 화합물(상기에 도시됨)의 이중 결합 중 하나 또는 둘 다를 1개 또는 2개의 단일 결합으로 환원시켜 구조식 XXIII의 상이한 지질-유사 화합물을 얻을 수 있다.

[1549] 본 발명의 다른 지질-유사 화합물은 상기에 기재된 합성 경로 및 당업계에 공지된 다른 경로를 통해서 다른 적합한 출발 물질로부터 제조될 수 있다. 상기에 언급된 방법은 지질-유사 화합물의 합성을 궁극적으로 허용하기 위해서 적합한 보호기를 첨가하거나 제거하는 추가 단계(들)를 포함할 수 있다. 또한, 다양한 합성 경로는 대안적인 순서로 또는 목적하는 물질을 제공하기 위해서 수행될 수 있다. 사용 가능한 지질-유사 화합물을 합성하는데 유용한 합성 화학 변형 및 보호기 방법론은 예를 들어, 문헌[R. Larock, Comprehensive Organic Transformations (2nd Ed., VCH Publishers 1999); P. G. M. Wuts and T. W. Greene, Greene's Protective Groups in Organic Synthesis (4th Ed., John Wiley and Sons 2007); L. Fieser and M. Fieser, Fieser and Fieser' s Reagents for Organic Synthesis (John Wiley and Sons 1994); 및 L. Paquette, ed., Encyclopedia of Reagents for Organic Synthesis (2nd ed., John Wiley and Sons 2009)] 및 이들의 후속 판을 비롯하여 당업계에 공지되어 있다. 특정 지질-유사 화합물은 비-방향족 이중 결합 및 하나 이상의 비대칭 중심을 함유할 수 있다. 따라서, 이들은 라세미체 및 라세미 혼합물, 단일 거울상이성질체, 개별 부분입체이성질체, 부분입체이성질체 혼합물 및 시스 또는 트랜스-이성질체 형태로 발생할 수 있다. 이러한 모든 이성질체 형태가 고려된다.

[1550] 상기에 언급된 바와 같이, 이러한 지질-유사 화합물은 약제의 전달에 유용하다. 이것은 시험관내 검정에 의해 약제 전달 효능에 대해 사전 스크리닝된 후 동물 실험 및 임상 시험을 통해 확인될 수 있다. 다른 방법도 당업자에게 명백할 것이다.

[1551] 어떠한 이론에도 얽매이고자 함 없이, 구조식 XXIII의 지질-유사 화합물은 복합체, 예를 들어, 나노복합체 및 마이크로입자를 형성함으로써 약제의 전달을 촉진한다. 양으로 또는 음으로 하전된 이러한 지질-유사 화합물의 친수성 헤드는 반대로 하전된 약제의 모이어티에 결합하고, 이의 소수성 모이어티는 약제의 소수성 모이어티에 결합한다. 결합은 공유 또는 비공유일 수 있다.

[1552] 상기에 기재된 복합체는 간행물, 예컨대, [Wang et al., ACS Synthetic Biology, 1, 403-07 (2012)]에 기재된 절차를 사용하여 제조될 수 있다. 일반적으로, 이들은 완충액, 예컨대, 아세트산나트륨 완충액 또는 인산염 완충 식염수("PBS")에서 지질-유사 화합물 및 약제를 인큐베이션시킴으로써 얻어진다.

[1553] ***친수성 기***

[1554] 특정 실시형태에서, 선택된 친수성 작용기 또는 모이어티는 (예를 들어, 화합물이 성분인 지질 나노입자의 형질주입 효율을 개선시킴으로써) 화합물 또는 이러한 화합물이 성분인 전달 비히클에 특성을 변경하거나 달리 부여할 수 있다. 예를 들어, 본 명세서에 개시된 화합물에서 친수성 헤드-기로서 구아니디늄의 혼입은 하나 이상의 표적 세포의 세포막과 이러한 화합물(또는 이러한 화합물이 성분인 전달 비히클)의 융합원성을 촉진할 수 있고, 이에 의해서 예를 들어, 상기 화합물의 형질주입 효율을 향상시킬 수 있다. 친수성 구아니디늄 모이어티의 질소는 상호작용에 안정성을 부여하여 캡슐화된 물질의 세포 흡수를 허용하는 6-원 고리 전이 상태를 형성한다고 가정되었다(문헌[Wender, *et al.*, Adv. Drug Del. Rev. (2008) 60: 452-472]). 유사하게, 하나 이상의 아미노기 또는 모이어티를 개시된 화합물(예를 들어, 헤드-기로서)에 혼입하는 것은 이러한 아미노기의 융합원성을 이용하여 표적 세포의 엔도솜/리소솜 막의 파괴를 추가로 촉진할 수 있다. 이것은 조성물의 아미노기의 pKa뿐만 아니라 육각형 상 전이를 겪고 표적 세포 표면, 즉, 소포막과 융합하는 아미노기의 능력에 기초한다 (문헌[Koltover, *et al.* Science (1998) 281: 78-81]). 그 결과는 소포막의 파괴 및 표적 세포로의 지질 나노입자 내용물의 방출을 촉진하는 것으로 여겨진다.

[1555] 유사하게, 특정 실시형태에서, 예를 들어, 본 명세서에 개시된 화합물에서 친수성 헤드-기로서 이미다졸의 혼입은 예를 들어, 전달 비히클에 캡슐화된 내용물(예를 들어, 지질 나노입자)의 엔도솜 또는 리소좀 방출을 촉진하는 역할을 할 수 있다. 이러한 향상된 방출은 양성자-스폰지 매개 파괴 기전 및/또는 향상된 융합원성 기전 중 하나 또는 둘 모두에 의해 달성될 수 있다. 양성자-스폰지 기전은 엔도솜의 산성화를 완충하는 화합물, 특히 화합물의 작용성 모이어티 또는 기의 능력에 기초한다. 이는 화합물의 pKa 또는 이러한 화합물을 포함하는 하나 이상의 작용기(예를 들어, 이미다졸)의 pKa에 의해 조작되거나 달리 제어될 수 있다. 따라서, 특정 실시형태에서, 예를 들어, 본 명세서에 개시된 이미다졸계 화합물(예를 들어, HGT4001 및 HGT4004)의 융합원성은 엔도솜 파괴 특성과 관련되는데, 이는 다른 전통적인 이온화 가능한 지질에 비해 더 낮은 pKa를 갖는 이러한 이미다졸기에 의해 촉진된다. 이러한 엔도솜 파괴 특성은 차례로 삼투성 팽창 및 리포솜 막의 파괴를 촉진하고, 그 다음 그 안에 로딩되거나 캡슐화된 폴리뉴클레오타이드 물질의 표적 세포 내로의 형질주입 또는 세포내 방출이 뒤따

른다. 이 현상은 이미다졸 모이어티 외에도 바람직한 pKa 프로파일을 갖는 다양한 화합물에 적용할 수 있다. 이러한 구체예는 또한 폴리아민, 폴리펩타이드(히스티딘), 및 질소계 수지상 구조와 같은 다중 질소계 작용기를 포함한다.

[1556] 예시적인 이온화 가능한 및/또는 양이온성 지질은 전문이 참조에 의해 본 명세서에 포함된 국제 PCT 특허 공개 제WO2015/095340호, 제WO2015/199952호, 제WO2018/011633호, 제WO2017/049245호, 제WO2015/061467호, 제WO2012/040184호, 제WO2012/000104호, 제WO2015/074085호, 제WO2016/081029호, 제WO2017/004143호, 제WO2017/075531호, 제WO2017/117528호, 제WO2011/022460호, 제WO2013/148541호, 제WO2013/116126호, 제WO2011/153120호, 제WO2012/044638호, 제WO2012/054365호, 제WO2011/090965호, 제WO2013/016058호, 제WO2012/162210호, 제WO2008/042973호, 제WO2010/129709호, 제WO2010/144740호, 제WO20 12/099755호, 제WO2013/049328호, 제WO2013/086322호, 제WO2013/086373호, 제WO2011/071860호, 제WO2009/132131호, 제WO2010/048536호, 제WO2010/088537호, 제WO2010/054401호, 제WO2010/054406호, 제WO2010/054405호, 제WO2010/054384호, 제WO2012/016184호, 제WO2009/086558호, 제WO2010/042877호, 제WO2011/000106호, 제WO2011/000107호, 제WO2005/120152호, 제WO2011/141705호, 제WO2013/126803호, 제WO2006/007712호, 제WO2011/038160호, 제WO2005/121348호, 제WO2011/066651호, 제WO2009/127060호, 제WO2011/141704호, 제WO2006/069782호, 제WO2012/031043호, 제WO2013/006825호, 제WO2013/033563호, 제WO2013/089151호, 제WO2017/099823호, 제WO2015/095346호 및 제WO2013/086354호, 및 미국 특허 공개 호US2016/0311759호, 제US2015/0376115호, 제US2016/0151284호, 제US2017/0210697호, 제US2015/0140070호, 제US2013/0178541호, 제US2013/0303587호, 제US2015/0141678호, 제US2015/0239926호, 제US2016/0376224호, 제US2017/0119904호, 제US2012/0149894호, 제US2015/0057373호, 제US2013/0090372호, 제US2013/0274523호, 제US2013/0274504호, 제US2013/0274504호, 제US2009/0023673호, 제US2012/0128760호, 제US2010/0324120호, 제US2014/0200257호, 제US2015/0203446호, 제US2018/0005363호, 제US2014/0308304호, 제US2013/0338210호, 제US2012/0101148호, 제US2012/0027796호, 제US2012/0058144호, 제US2013/0323269호, 제US2011/0117125호, 제US2011/0256175호, 제US2012/0202871호, 제US2011/0076335호, 제US2006/0083780호, 제US2013/0123338호, 제US2015/0064242호, 제US2006/0051405호, 제US2013/0065939호, 제US2006/0008910호, 제US2003/0022649호, 제US2010/0130588호, 제US2013/0116307호, 제US2010/0062967호, 제US2013/0202684호, 제US2014/0141070호, 제US2014/0255472호, 제US2014/0039032호, 제US2018/0028664호, 제US2016/0317458호 및 제US2013/0195920에 기재되어 있다. 국제 특허 출원 제WO 2019/131770호 또한 전문이 참조에 의해 본 명세서에 포함된다.

[1557] **8. PEG 지질**

[1558] 바람직하게는 본 명세서에 개시된 화합물 중 하나 이상 및 지질과 조합하여, 본 명세서에 기재된 리포솜 및 약제학적 조성물에서 폴리에틸렌 글리콜(PEG)-변형된 인지질 및 유도체화된 지질, 예컨대 유도체화된 세라마이드(PEG-CER), 예컨대, N-옥탄일-스핑고신-1-[석신일(메톡시 폴리에틸렌 글리콜)-2000](C8 PEG-2000 세라마이드)을 사용하고, 이를 포함시키는 것이 고려된다. 고려되는 PEG-변형된 지질은 C6-C20 길이의 알킬 쇄(들)를 갖는 지질에 공유 부착된 최대 5kDa 길이의 폴리에틸렌 글리콜 쇄를 포함하지만 이들로 제한되지 않는다. 일부 실시형태에서, 본 발명의 조성물 및 방법에 사용되는 PEG-변형된 지질은 1,2-다이미리스토일-sn-글리세롤, 메톡시폴리에틸렌 글리콜(2000 MW PEG) "DMG-PEG2000"이다. PEG-변형된 지질을 지질 전달 비히클에 첨가하는 것은 복합체 응집을 방지할 수 있고, 또한 원형화 수명을 증가시키고, 표적 조직으로의 지질-폴리뉴클레오타이드 조성물의 전달을 증가시키기 위한 수단을 제공할 수 있거나(문헌[Klibanov *et al.* (1990) FEBS Letters, 268 (1): 235-237]) 또는 이들은 생체내에서 제형 외부로 신속하게 교환되도록 선택될 수 있다(미국 특허 제5,885,613호 참조). 특히 유용한 교환 가능한 지질은 더 짧은 아실 쇄를 갖는 PEG-세라마이드(예를 들어, C14 또는 C18)이다. 본 발명의 PEG-변형된 인지질 및 유도체화된 지질은 리포솜 지질 나노입자에 존재하는 총 지질의 약 0% 내지 약 20%, 약 0.5% 내지 약 20%, 약 1% 내지 약 15%, 약 4% 내지 약 10% 또는 약 2%의 몰비를 차지할 수 있다.

[1559] 실시형태에서, PEG-변형된 지질은 전문이 참조에 의해 본 명세서에 포함된 국제 특허 출원 제PCT/US2019/015913호에 기재되어 있다. 일 실시형태에서, 전달 비히클은 하나 이상의 PEG-변형된 지질을 포함한다.

[1560] PEG-변형된 지질의 비제한적인 예는 PEG-변형된 포스파티딜에탄올아민 및 포스파티드산, PEG-세라마이드 접합체(예를 들어, PEG-CerC14 또는 PEG-CerC20), PEG-변형된 다이알킬아민 및 PEG-변형된 1,2-다이아실옥시프로판-3-아민을 포함한다. 일부 추가 실시형태에서, PEG-변형된 지질은 예를 들어, PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC 또는 PEG-DSPE일 수 있다.

[1561] 일부 더 추가의 실시형태에서, PEG-변형된 지질은 1,2-다이미리스토일-sn-글리세롤 메톡시폴리에틸렌 글리콜(PEG-DMG), 1,2-다이스테아로일-sn-글리세로-3-포스포에탄올아민-N-[아미노(폴리에틸렌 글리콜)](PEG-DSPE), PEG-다이스테릴 글리세롤(PEG-DSG), PEG-다이팔메톨레일, PEG-다이올레일, PEG-다이스테아릴, PEG-다이아실글리카마이드(PEG-DAG), PEG-다이팔미토일 포스파티딜에탄올아민(PEG-DPPE) 또는 PEG-1,2-다이미리스틸옥실프로필-3-아민(PEG-c-DMA)을 포함하지만 이들로 제한되지 않는다.

[1562] 다양한 실시형태에서, PEG-변형된 지질은 또한 "PEGylated 지질" 또는 "PEG-지질"이라고 지칭될 수 있다.

[1563] 일 실시형태에서, PEG-지질은 PEG-변형된 포스파티딜에탄올아민, PEG-변형된 포스파티드산, PEG-변형된 세라마이드, PEG-변형된 다이알킬아민, PEG-변형된 다이아실글리세롤, PEG-변형된 다이알킬글리세롤 및 이들의 혼합물로 이루어진 군으로부터 선택된다.

[1564] 일부 실시형태에서, PEG-지질의 지질 모이어티는 약 C_{14} 내지 약 C_{22}, 예컨대 약 C_{14} 내지 약 C_{16}의 길이를 갖는 것을 포함한다. 일부 실시형태에서, PEG 모이어티, 예를 들어, mPEG-NH2는 약 1000, 약 2000, 약 5000, 약 10,000, 약 15,000 또는 약 20,000달톤의 크기를 갖는다. 일 실시형태에서, PEG-지질은 PEG2k-DMG이다.

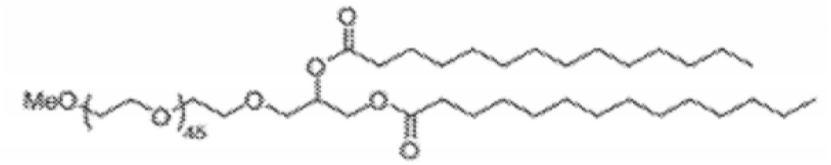
[1565] 일 실시형태에서, 본 명세서에 기재된 지질 나노입자는 비확산성 PEG로 변형된 지질을 포함할 수 있다. 비확산성 PEG의 비제한적인 예는 PEG-DSG 및 PEG-DSPE를 포함한다.

[1566] PEG-지질은 당업계에 공지되어 있고, 예컨대, 미국 특허 제8,158,601호 및 국제 공개 제WO2015/130584 A2호에 기재된 것이고, 이들은 전문이 참조에 의해 본 명세서에 포함된다.

[1567] 다양한 실시형태에서, 본 명세서에 기재된 지질(예를 들어, PEG-지질)은 전문이 참조에 의해 본 명세서에 포함된 국제 특허 공개 제PCT/US2016/000129호에 기재된 바와 같이 합성될 수 있다.

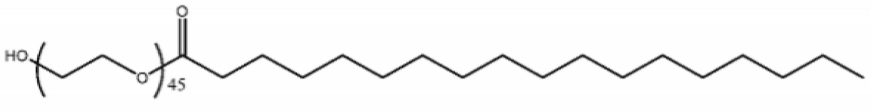
[1568] 지질 나노입자 조성물의 지질 성분은 폴리에틸렌 글리콜, 예컨대 PEG 또는 PEG-변형된 지질을 포함하는 하나 이상의 분자를 포함할 수 있다. 이러한 종은 대안적으로 PEG화된 지질이라고 지칭될 수 있다. PEG 지질은 폴리에틸렌 글리콜로 변형된 지질이다. PEG 지질은 PEG-변형된 포스파티딜에탄올아민, PEG-변형된 포스파티드산, PEG-변형된 세라마이드, PEG-변형된 다이알킬아민, PEG-변형된 다이아실글리세롤, PEG-변형된 다이알킬글리세롤 및 이들의 혼합물을 포함하는 비제한적인 군으로부터 선택될 수 있다. 예를 들어, PEG 지질은 PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC 또는 PEG-DSPE 지질일 수 있다.

[1569] 일부 실시형태에서 PEG-변형된 지질은 PEG-DMG의 변형된 형태이다. PEG-DMG는 하기 구조를 갖는다:

[1570] 

.

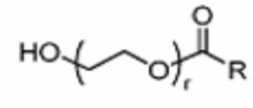
[1571] 일부 실시형태에서 PEG-변형된 지질은 PEG-C18 또는 PEG-1의 변형된 형태이다. PEG-1은 하기 구조를 갖는다:

[1572] 

.

[1573] 일 실시형태에서, 본 발명에 유용한 PEG 지질은 국제 공개 제WO2012099755호에 기재된 PEG화된 지질일 수 있고, 이의 내용은 전문이 참조에 의해 본 명세서에 포함된다. 본 명세서에 기재된 이러한 예시적인 PEG 지질 중 임의의 것이 PEG 쇄 상에 하이드록실기를 포함하도록 변형될 수 있다. 특정 실시형태에서, PEG 지질은 PEG-OH 지질이다. 특정 실시형태에서, PEG-OH 지질은 PEG 쇄 상에 하나 이상의 하이드록실기를 포함한다. 특정 실시형태에서, PEG-OH 또는 하이드록시-PEG화된 지질은 PEG 쇄의 말단에 -OH기를 포함한다. 각각의 가능성은 본 발명의 개별 실시형태를 나타낸다.

[1574] 일부 실시형태에서, PEG 지질은 화학식 (P1)의 화합물 또는 이의 염 또는 이성질체이다:

[1575]  (P1)

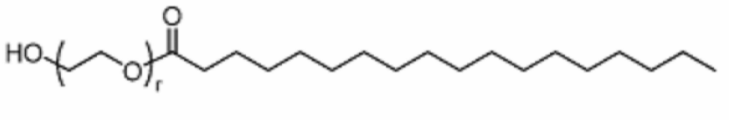
[1576] 식 중,

[1577] r은 1 내지 100의 정수이고;

[1578] R은 C_{10-40} 알킬, C_{10-40} 알켄일 또는 C_{10-40} 알킨일; 선택적으로 R의 하나 이상의 메틸렌기는 C_{3-10} 카보사이클릴렌, 4 내지 10원 헤테로사이클릴렌, C_{6-10} 아릴렌, 4 내지 10원 헤테로아릴렌, $-N(R^N)-$, $-O-$, $-S-$, $-C(O)-$, $-C(O)N(R^N)-$, $-NR^NC(O)-$, $-NR^NC(O)N(R^N)-$, $-C(O)O-$, $-OC(O)-$, $-OC(O)O-$, $-OC(O)N(R^N)-$, $-NR^NC(O)O-$, $-C(O)S-$, $-SC(O)-$, $-C(=NR^N)-$, $-C(=NR^N)N(R^N)-$, $-NR^NC(=NR^N)-$, $-NR^NC(=NR^N)N(R^N)-$, $-C(S)-$, $-C(S)N(R^N)-$, $-NR^NC(S)-$, $-NR^NC(S)N(R^N)-$, $-S(O)-$, $-OS(O)-$, $-S(O)O-$, $-OS(O)O-$, $-OS(O)_2-$, $-S(O)_2O-$, $-OS(O)_2O-$, $-N(R^N)S(O)-$, $-S(O)N(R^N)-$, $-N(R^N)S(O)N(R^N)-$, $-OS(O)N(R^N)-$, $-N(R^N)S(O)O-$, $-S(O)2-$, $-N(R^N)S(O)2-$, $-S(O)_2N(R^N)-$, $-N(R^N)S(O)_2N(R^N)-$, $-OS(O)_2N(R^N)-$ 또는 $-N(R^N)S(O)_2O-$로 독립적으로 대체되고;

[1579] R^N의 각각의 경우는 독립적으로 수소, C_{1-6} 알킬 또는 질소 보호기이다.

[1580] 예를 들어, R은 C17 알킬이다. 예를 들어, PEG 지질은 화학식 (P1-a)의 화합물 또는 이의 염 또는 이성질체이다:

[1581] 

[1582] 식 중, r은 1 내지 100의 정수이다.

[1583] 예를 들어, PEG 지질은 하기 화학식의 화합물이다:

[1584] 

[1585] **9. 헬퍼 지질**

[1586] 일부 실시형태에서, 본 명세서에 기재된 전달 비히클(예를 들어, LNP)은 하나 이상의 비-양이온성 헬퍼 지질을 포함한다. 일부 실시형태에서, 헬퍼 지질은 인지질이다. 일부 실시형태에서, 헬퍼 지질은 인지질 치환체 또는 대체물이다. 일부 실시형태에서, 인지질 또는 인지질 치환체는 예를 들어, 하나 이상의 치환된 또는 (폴리)비치환된 인지질 또는 인지질 치환체 또는 이들의 조합물일 수 있다. 일반적으로, 인지질은 인지질 모이어티 및 하나 이상의 지방산 모이어티를 포함한다.

[1587] 인지질 모이어티는 예를 들어, 포스파티딜 콜린, 포스파티딜 에탄올아민, 포스파티딜 글리세롤, 포스파티딜 세린, 포스파티드산, 2-리소포스파티딜 콜린, 및 스핑고미엘린으로 이루어진 비제한적인 군으로부터 선택될 수 있다.

[1588] 지방산 모이어티는 예를 들어, 라우르산, 미리스트산, 미리스트올레산, 팔미트산, 팔미트올레산, 스테아르산, 올레산, 리놀레산, 알파-리놀렌산, 에루스산, 피탄산, 아라키드산, 아라키돈산, 에이코사펜타엔산, 베헨산, 도코사펜타엔산 및 도코사헥사엔산으로 이루어진 비제한적인 군으로부터 선택될 수 있다.

[1589] 인지질은 글리세로인지질, 예컨대, 포스파티딜콜린, 포스파티딜에탄올아민, 포스파티딜세린, 포스파티딜이노시톨, 포스파티딜 글리세롤 및 포스파티드산을 포함하지만 이들로 제한되지 않는다. 인지질은 또한 포스포스핑고지질, 예컨대, 스핑고미엘린을 포함한다.

[1590] 일부 실시형태에서, 헬퍼 지질은 1,2-다이스테아로일-177-글리세로-3-포스포콜린(DSPC) 유사체, DSPC 치환체,

올레산 또는 올레산 유사체이다.

[1591] 일부 실시형태에서, 헬퍼 지질은 비-포스파티딜 콜린(PC) 쯔비터이온성 지질, DSPC 유사체, 올레산, 올레산 유사체 또는 DSPC 치환체이다.

[1592] 일부 실시형태에서, 헬퍼 지질은 PCT/US2018/053569에 기재되어 있다. 본 개시내용의 지질 조성물에 사용하기에 적합한 헬퍼 지질은 예를 들어, 다양한 중성, 비하전 또는 쯔비터이온성 지질을 포함한다. 이러한 헬퍼 지질은 본 명세서에 개시된 화합물 중 하나 이상 및 지질과 조합하여 바람직하게 사용된다. 헬퍼 지질의 예는 5-헵타데실벤젠-1,3-다이올(레소르시놀), 다이팔미토일포스파티딜콜린(DPPC), 다이스테아로일포스파티딜콜린(DSPC), 포스포콜린(DOPC), 다이미리스토일포스파티딜콜린(DMPC), 포스파티딜콜린(PLPC), 1,2-다이스테아로일sn-글리세로-3-포스포콜린(DAPC), 포스파티딜에탄올아민(PE), 난 포스파티딜콜린(EPC), 다이라우릴로일포스파티딜콜린(DLPC), 다이미리스토일포스파티딜콜린(DMPC), 1-미리스토일-2-팔미토일 포스파티딜콜린(MPPC), 1-팔미토일-2-미리스토일 포스파티딜콜린(PMPC), 1-팔미토일-2-스테아로일 포스파티딜콜린(PSPC), 1,2-다이아라키도일-sn-글리세로-3-포스포콜린(DBPC), 1-스테아로일-2-팔미토일 포스파티딜콜린(SPPC), 1,2-다이에이코센오일-sn-글리세로-3-포스포콜린(DEPC), 팔미토일올레오일 포스파티딜콜린(POPC), 리소포스파티딜 콜린, 다이올레오일 포스파티딜에탄올 아민(DOPE) 다이리놀레오일포스파티딜콜린 다이스테아로일포스파티딜에탄올아민(DSPE), 다이미리스토일 포스파티딜에탄올아민 (DMPE), 다이팔미토일 포스파티딜에탄올아민(DPPE), 팔미토일올레오일 포스파티딜에탄올아민 (POPE), 리소포스파티딜에탄올아민 및 이들의 조합물을 포함하지만 이들로 제한되지 않는다. 일 실시형태에서, 헬퍼 지질은 다이스테아로일포스파티딜콜린(DSPC) 또는 다이미리스토일 포스파티딜 에탄올아민(DMPE)일 수 있다. 또 다른 실시형태에서, 헬퍼 지질은 다이스테아로일포스파티딜콜린(DSPC)일 수 있다. 헬퍼 지질은 전달 비히클을 안정화시키고, 이의 가공을 개선시키는 기능을 한다. 이러한 헬퍼 지질은 다른 부형제, 예를 들어, 본 명세서에 개시된 이온화 가능한 지질과 조합하여 바람직하게 사용된다. 일부 실시형태에서, 이온화 가능한 지질과 조합하여 사용되는 경우, 헬퍼 지질은 지질 나노입자에 존재하는 총 지질의 5% 내지 약 90% 또는 약 10% 내지 약 70%의 몰비를 차지할 수 있다.

[1593] **10. 구조 지질**

[1594] 일 실시형태에서, 구조 지질은 국제 특허 출원 제PCT/US2019/015913호에 기재되어 있다.

[1595] 본 명세서에 기재된 전달 비히클은 하나 이상의 구조 지질을 포함한다. 지질 나노입자에 구조 지질을 혼입하는 것은 입자 내의 다른 지질의 응집을 완화시키는 데 도움이 될 수 있다. 구조 지질은 콜레스테롤, 페코스테롤, 에르고스테롤, 바시카스테롤, 토마티딘, 토마틴, 우르솔릭, 알파-토코페롤 및 이들의 혼합물을 포함할 수 있지만 이들로 제한되지 않는다. 특정 실시형태에서, 구조 지질은 콜레스테롤이다. 특정 실시형태에서, 구조 지질은 콜레스테롤 및 코티코스테로이드(예컨대, 예를 들어, 프레드니솔론, 덱사메타손, 프레드니손 및 하이드로코티손) 또는 이들의 조합물을 포함한다.

[1596] 일부 실시형태에서, 구조 지질은 스테롤이다. 특정 실시형태에서, 구조 지질은 스테로이드이다. 특정 실시형태에서, 구조 지질은 콜레스테롤이다. 특정 실시형태에서, 구조 지질은 콜레스테롤의 유사체이다. 특정 실시형태에서, 구조 지질은 알파-토코페롤이다.

[1597] 본 명세서에 기재된 전달 비히클은 하나 이상의 구조 지질을 포함한다. 전달 비히클, 예를 들어, 지질 나노입자에 구조 지질을 혼입하는 것은 입자 내의 다른 지질의 응집을 완화시키는 데 도움이 될 수 있다. 특정 실시형태에서, 구조 지질은 콜레스테롤 및 코티코스테로이드(예컨대, 예를 들어, 프레드니솔론, 덱사메타손, 프레드니손 및 하이드로코티손) 또는 이들의 조합물을 포함한다.

[1598] 일부 실시형태에서, 구조 지질은 스테롤이다. 구조 지질은 스테롤(예를 들어, 피토스테롤 또는 주스테롤)을 포함할 수 있지만 이들로 제한되지 않는다.

[1599] 특정 실시형태에서, 구조 지질은 스테로이드이다. 예를 들어, 스테롤은 콜레스테롤, β-시토스테롤, 페코스테롤, 에르고스테롤, 시토스테롤, 캄페스테롤, 스티그마스테롤, 브라시카스테롤, 에르고스테롤, 토마티딘, 토마틴, 우르솔산 또는 알파-토코페롤을 포함할 수 있지만 이들로 제한되지 않는다.

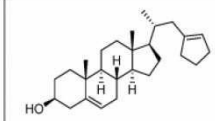
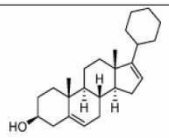
[1600] 일부 실시형태에서, 전달 비히클은 전달 비히클, 예를 들어, 지질 나노입자에 존재하는 경우, 세포 회합/흡수, 내재화, 세포내 전달 및/또는 가공 및/또는 엔도솜 탈출을 향상시킴으로써 기능할 수 있고/있거나 면역 세포 전달 강화 지질이 없는 전달 비히클에 비해서 면역 세포에 의한 인식 및/또는 면역 세포에 대한 결합을 향상시킬 수 있는 면역 세포 전달 강화 지질, 예를 들어, 콜레스테롤 유사체 또는 아미노 지질 또는 이들의 조합물을 유효량으로 포함한다. 따라서, 어떠한 특정 기전 또는 이론에도 얽매이고자 함 없이, 일 실시형태에서, 본 개시내

용의 구조 지질 또는 다른 면역 세포 전달 강화 지질은 C1q에 결합하거나, 이러한 지질을 포함하는 전달 비히클의 C1q에 대한 결합을 촉진시킨다. 따라서, 면역 세포로의 핵산 분자의 전달을 위해서 본 개시내용의 전달 비히클을 시험관내에서 사용하는 경우, C1q를 포함하는 배양 조건이 사용된다(예를 들어, 혈청을 포함하거나 무혈청 배지에 외인성 C1q를 첨가한 배양 배지 사용). 본 개시내용의 전달 비히클의 생체내 사용의 경우, C1q의 요건은 내인성 C1q에 의해서 공급된다.

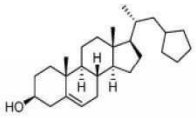
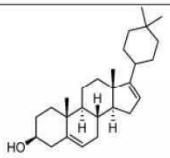
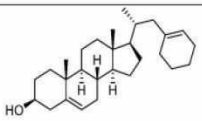
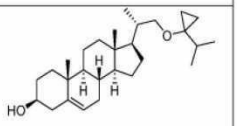
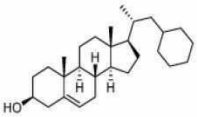
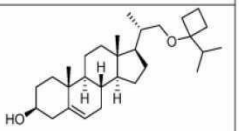
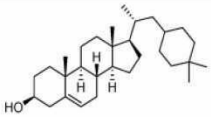
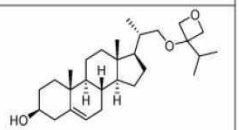
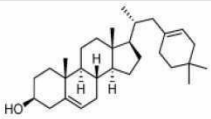
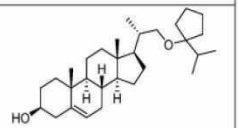
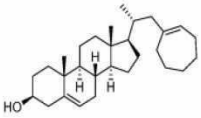
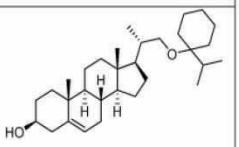
[1601] 특정 실시형태에서, 구조 지질은 콜레스테롤이다. 특정 실시형태에서, 구조 지질은 콜레스테롤의 유사체이다. 일부 실시형태에서, 구조 지질은 표 16의 지질이다:

표 16

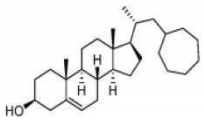
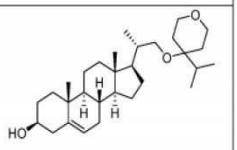
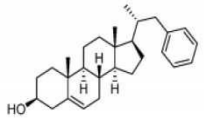
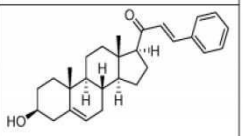
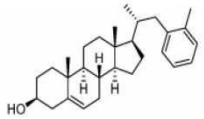
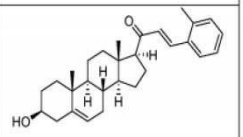
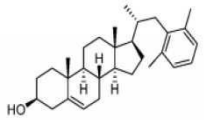
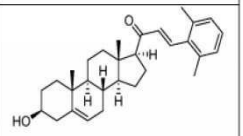
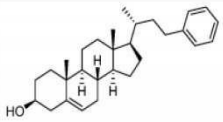
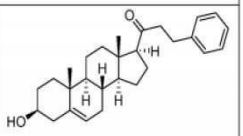
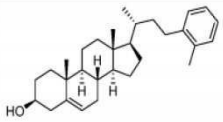
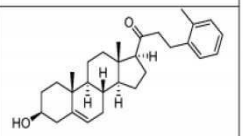
[1602]

화합물 번호 S-	구조	화합물 번호 S-	구조
1		22	

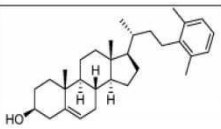
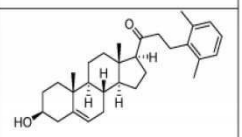
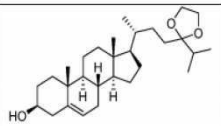
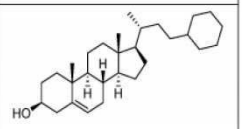
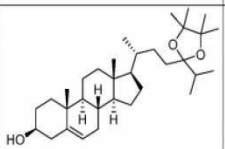
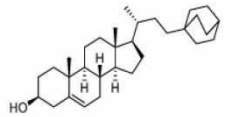
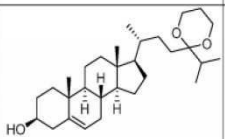
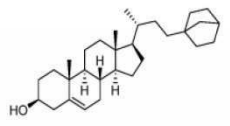
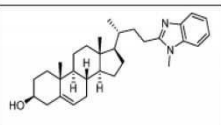
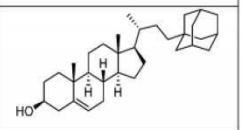
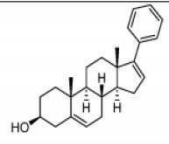
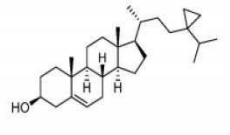
[1603]

화합물 번호 S-	구조	화합물 번호 S-	구조
2		23	
3		24	
4		25	
5		26	
6		27	
7		28	

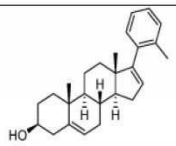
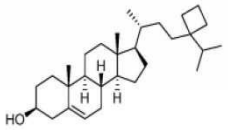
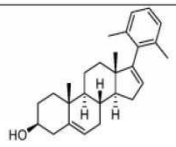
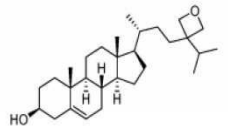
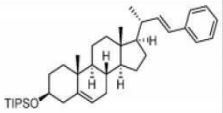
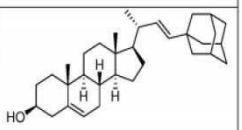
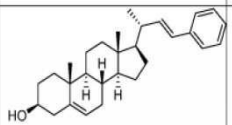
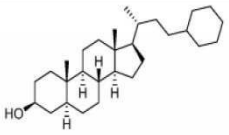
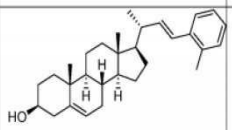
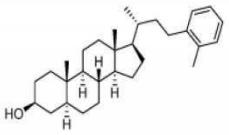
[1604]

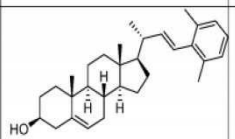
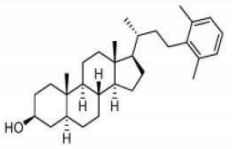
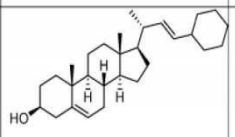
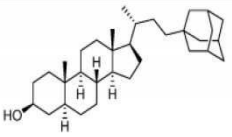
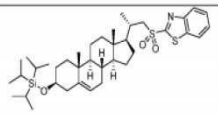
화합물 번호 S-	구조	화합물 번호 S-	구조
8		29	
9		30	
10		31	
11		32	
12		33	
13		34	

[1605]

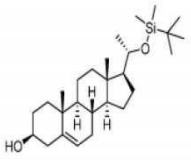
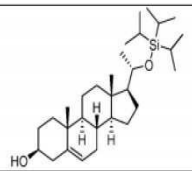
화합물 번호 S-	구조	화합물 번호 S-	구조
14		35	
15		36	
16		37	
17		38	
18		39	
19		40	

[1606]

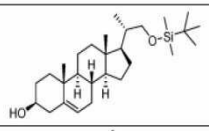
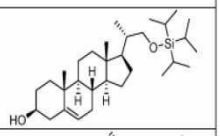
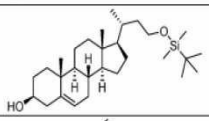
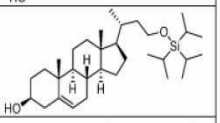
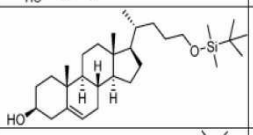
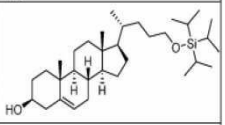
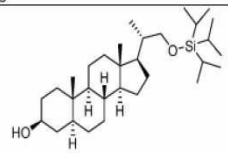
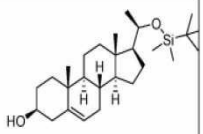
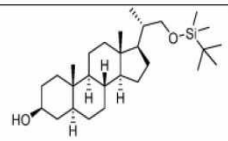
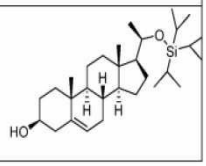
화합물 번호 S-	구조	화합물 번호 S-	구조
20		41	
21		42	
150		165	
154		169	
162		170	

화합물 번호 S-	구조	화합물 번호 S-	구조
163		171	
164		172	
184			

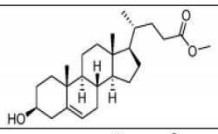
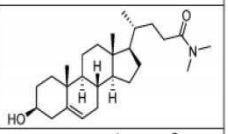
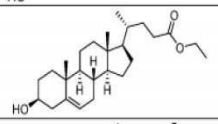
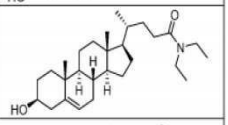
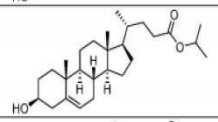
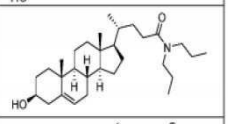
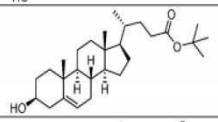
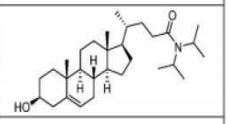
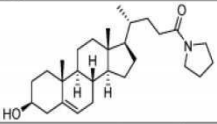
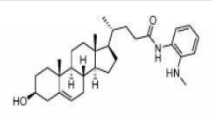
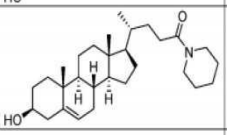
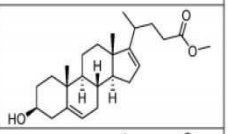
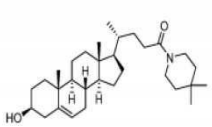
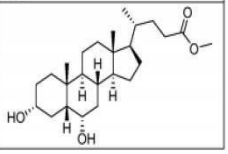
[1607]

화합물 번호 S-	구조	화합물 번호 S-	구조
43		47	

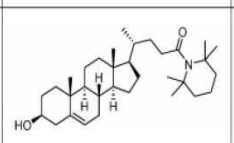
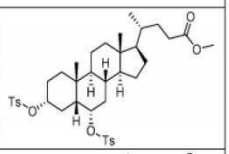
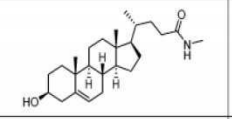
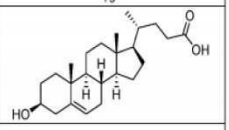
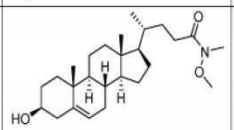
[1608]

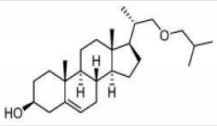
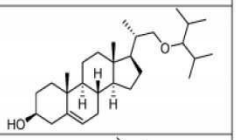
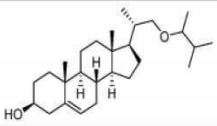
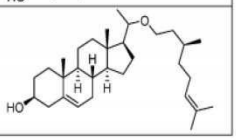
화합물 번호 S-	구조	화합물 번호 S-	구조
44		48	
45		49	
46		50	
175		177	
176		178	

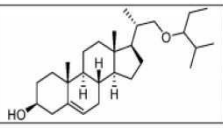
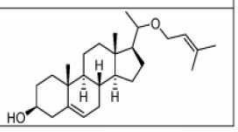
[1609]

화합물 번호 S-	구조	화합물 번호 S-	구조
51		60	
52		61	
53		62	
54		63	
55		64	
56		65	
57		66	

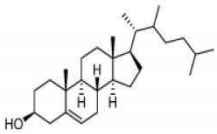
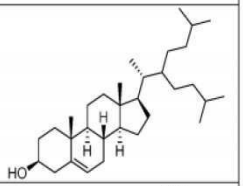
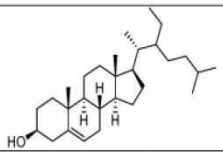
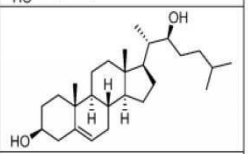
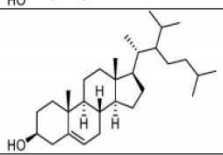
[1610]

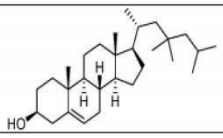
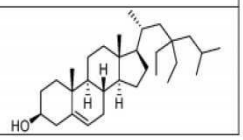
화합물 번호 S-	구조	화합물 번호 S-	구조
58		67	
59		149	
153			

화합물 번호 S-	구조	화합물 번호 S-	구조
68		71	
69		72	

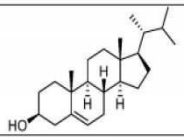
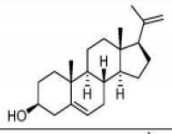
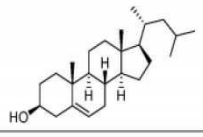
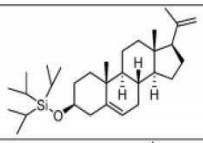
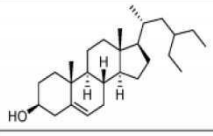
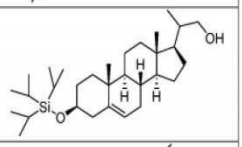
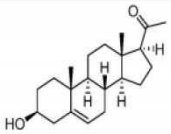
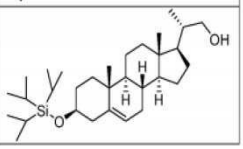
화합물 번호 S-	구조	화합물 번호 S-	구조
70		73	

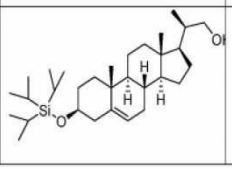
[1611]

화합물 번호 S-	구조	화합물 번호 S-	구조
74		77	
75		78	
76			

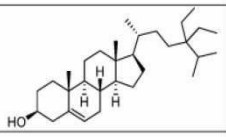
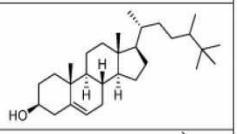
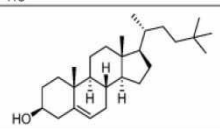
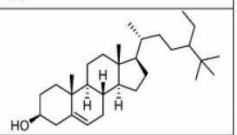
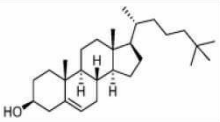
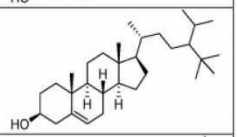
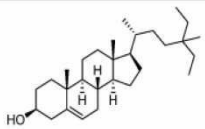
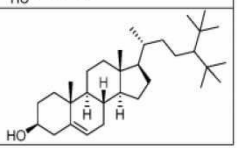
화합물 번호 S-	구조	화합물 번호 S-	구조
79		80	

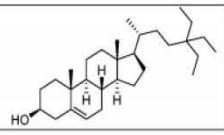
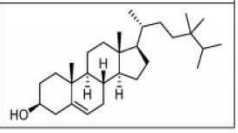
[1612]

화합물 번호 S-	구조	화합물 번호 S-	구조
81		85	
82		86	
83		87	
84		152	

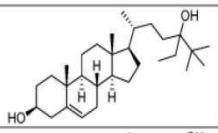
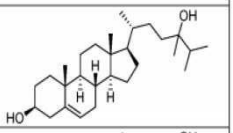
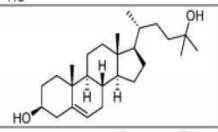
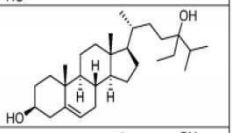
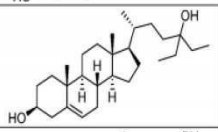
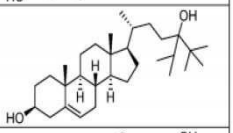
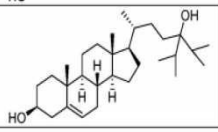
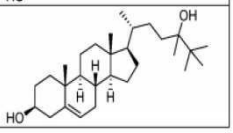
화합물 번호 S-	구조	화합물 번호 S-	구조
157			

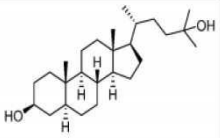
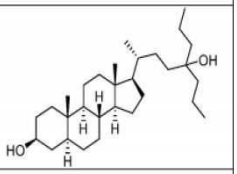
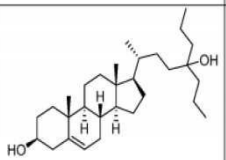
[1613]

화합물 번호 S-	구조	화합물 번호 S-	구조
88		93	
89		94	
90		95	
91		96	

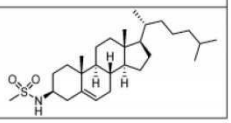
화합물 번호 S-	구조	화합물 번호 S-	구조
92		97	

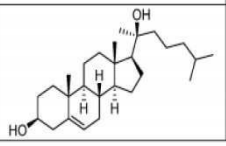
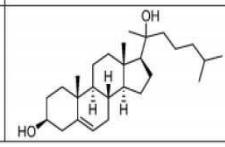
[1614]

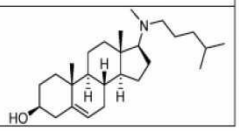
화합물 번호 S-	구조	화합물 번호 S-	구조
98		102	
99		103	
100		104	
101		105	

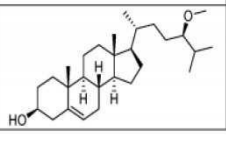
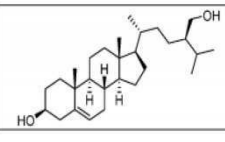
화합물 번호 S-	구조	화합물 번호 S-	구조
180		182	
181			

[1615]

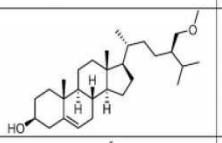
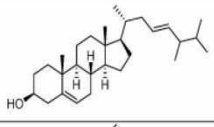
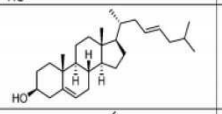
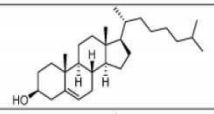
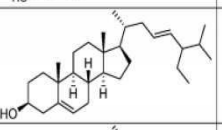
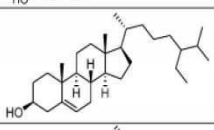
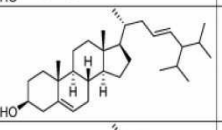
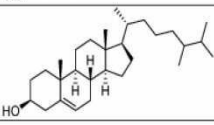
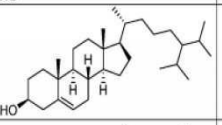
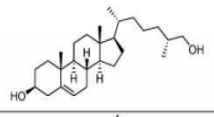
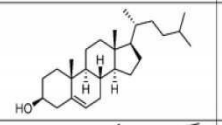
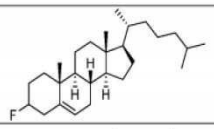
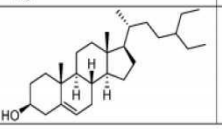
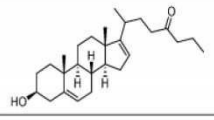
화합물 번호 S-	구조
106	

화합물 번호 S-	구조	화합물 번호 S-	구조
107		108	

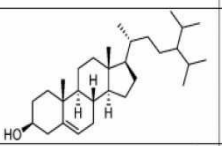
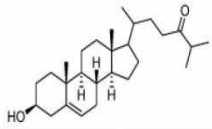
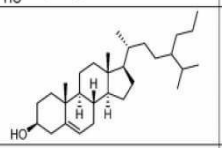
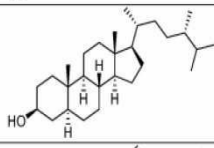
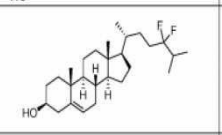
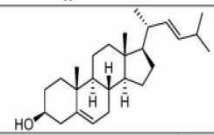
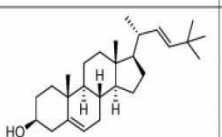
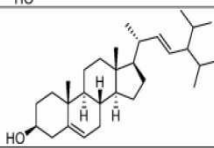
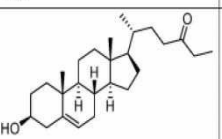
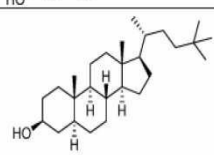
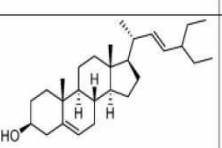
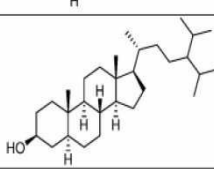
화합물 번호 S-	구조
109	

화합물 번호 S-	구조	화합물 번호 S-	구조
110		121	

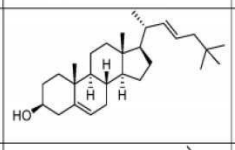
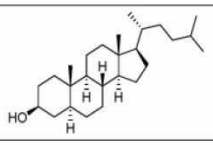
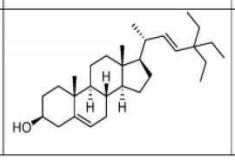
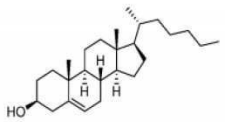
[1616]

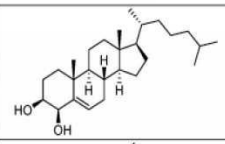
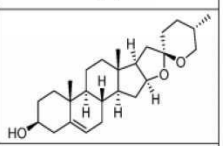
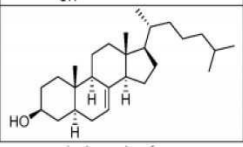
화합물 번호 S-	구조	화합물 번호 S-	구조
111		122	
112		123	
113		124	
114		125	
115		126	
116		127	
117		128	

[1617]

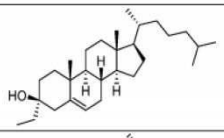
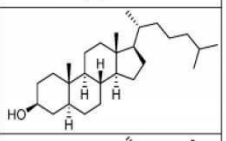
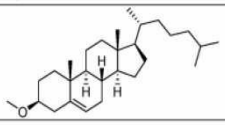
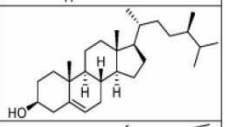
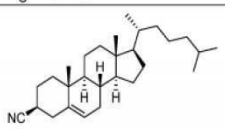
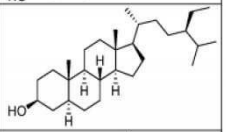
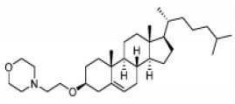
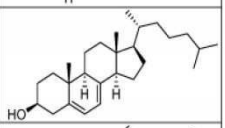
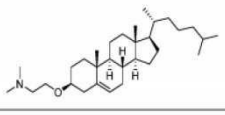
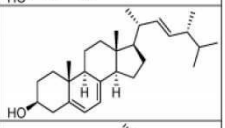
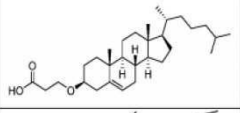
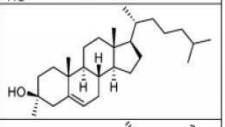
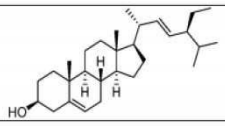
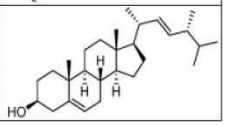
화합물 번호 S-	구조	화합물 번호 S-	구조
118		129	
119		130	
120		155	
156		167	
158		168	
160		173	

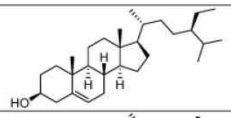
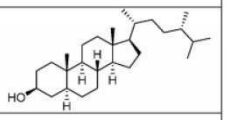
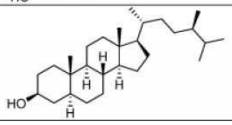
[1618]

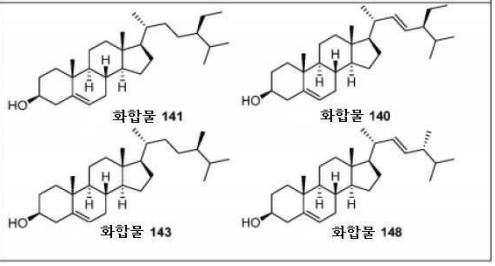
화합물 번호 S-	구조	화합물 번호 S-	구조
161		174	
166		179	

화합물 번호 S-	구조	화합물 번호 S-	구조
131		133	
132			

[1619]

화합물 번호 S-	구조	화합물 번호 S-	구조
134		142	
135		143	
136		144	
137		145	
138		146	
139		147	
140		148	

화합물 번호 S-	구조	화합물 번호 S-	구조
141		151	
159			

조성물 S-번호	구조
183	

[1620]

[1621] **11. LNP 제형**

[1622] 본 명세서에 기재된 지질 나노입자(LNP)는 당업계에 공지된 임의의 방법에 의해서 달성될 수 있다. 예를 들어, 전문이 참조에 의해 본 명세서에 포함된 미국 특허 공개 제US2012/0178702 A1호에 기재된 바와 같음. 지질 나노입자 조성물 및 이의 제조 방법의 비제한적인 예는 예를 들어, 문헌[Semple *et al.* (2010) Nat. Biotechnol. 28:172-176; Jayarama *et al.* (2012), Angew. Chem. Int. Ed., 51:8529-8533; 및 Maier *et al.* (2013) Molecular Therapy 21, 1570-1578](이들 각각의 내용은 전문이 참조에 의해 본 명세서에 포함됨)에 기재되어 있다.

[1623] 일 실시형태에서, LNP 제형은 예를 들어, 각각의 내용 전문이 참조에 의해 본 명세서에 포함된 국제 특허 공개 제WO 2011/127255호 또는 제WO 2008/103276호에 기재된 방법에 의해서 제조될 수 있다.

[1624] 일 실시형태에서, 본 명세서에 기재된 LNP 제형은 다양이온성 조성물을 포함할 수 있다. 비제한적인 예로서, 다양이온성 조성물은 내용 전문이 참조에 의해 본 명세서에 포함된 미국 특허 공개 제US2005/0222064 A1호의 화학식 1 내지 60으로부터 선택된 조성물일 수 있다.

[1625] 일 실시형태에서, 지질 나노입자는 각각의 전문이 참조에 의해 본 명세서에 포함된 미국 특허 공개 제US2013/0156845 A1호 및 국제 특허 공개 제WO2013/093648 A2호 또는 제WO2012/024526 A2호에 기재된 방법에 의해서 제형화될 수 있다.

[1626] 일 실시형태에서, 본 명세서에 기재된 지질 나노입자는 전문이 참조에 의해 본 명세서에 포함된 미국 특허 공개 제US2013/0164400 A1호에 기재된 시스템 및/또는 방법에 의해서 멸균 환경으로 제조될 수 있다.

[1627] 일 실시형태에서, LNP 제형은 나노입자, 예컨대, 전문이 참조에 의해 본 명세서에 포함된 미국 특허 제8,492,359호에 기재된 핵산-지질 입자로 제형화될 수 있다.

[1628] 나노입자 조성물은 선택적으로 1종 이상의 코팅을 포함할 수 있다. 예를 들어, 나노입자 조성물은 코팅을 갖는 캡슐, 필름 또는 정제로 제형화될 수 있다. 본 명세서에 기재된 조성물을 포함하는 캡슐, 필름 또는 정제는 임의의 유용한 크기, 인장 강도, 경도 또는 밀도를 가질 수 있다.

[1629] 일부 실시형태에서, 본 명세서에 기재된 지질 나노입자는 미세유체 혼합기를 포함하는 방법을 사용하여 합성될 수 있다. 예시적인 미세유체 혼합기는 프리시젼 나노시스템즈사(Precision Nanosystems)(캐나다 BC주 밴쿠버 소재), 마이크로이노바사(Microinnova)(Allerheiligen bei Wildon, Austria)에서 제조된 것을 포함하지만 이들로 제한되지 않는 슬릿 인터디지털(slit interdigitial) 마이크로혼합기 및/또는 스태거드 헤링본 마이크로혼합기(staggered herringbone micromixer: SHM)을 포함할 수 있지만 이들로 제한되지 않는다(전문이 참조에 의해 본 명세서에 포함된 문헌[Zhigaltsev, I.V. *et al.* (2012) Langmuir. 28:3633-40; Belliveau, N.M. *et al.* Mol. Ther. Nucleic. Acids. (2012) 1:e37; Chen, D. *et al.* J. Am. Chem. Soc. (2012) 134(16):6948-51]).

[1630] 일부 실시형태에서, SHM을 포함하는 LNP 생성 방법은 적어도 2개의 투입 스트림의 혼합을 추가로 포함하는데, 여기서 혼합은 마이크로구조-유도 카오스 이류(microstructure-induced chaotic advection: MICA)에 의해서 발

생한다. 이 방법에 따르면, 유체 스트림이 헤링본 패턴으로 존재하는 채널을 통해 유동하여 회전 유동을 유도하고, 유체를 서로 주변에서 폴딩한다. 이러한 방법은 또한 유체 혼합을 위한 표면을 포함할 수 있는데, 여기서 표면은 유체 순환 동안 배향을 변화시킨다. SHM을 사용한 LNP의 생성 방법은 전문이 참조에 의해 본 명세서에 포함된 미국 특허 공개 제US2004/0262223 A1호 및 제US2012/0276209 A1호에 개시된 것을 포함한다.

[1631] 일 실시형태에서, 지질 나노입자는 마이크로혼합기 예컨대, 비제한적으로 인스티튜트 퍼 마이크로네크닉 마인즈사(the Institut fur Mikrotechnik Mainz GmbH)(독일 마인즈 소재)로부터의 Slit Interdigital Microstructured Mixer(SIMM-V2) 또는 Standard Slit Interdigital Micro Mixer(SSIMM) 또는 Caterpillar(CPMM) 또는 Impinging-jet(IJMM)를 사용하여 제형화될 수 있다. 일 실시형태에서, 지질 나노입자는 마이크로유체 기술을 사용하여 생성된다(각각 전문이 참조에 의해 본 명세서에 포함된 문헌[Whitesides (2006) Nature. 442: 368-373; 및 Abraham *et al.* (2002) Science. 295: 647-651] 참조). 비제한적인 예로서, 제어된 마이크로유체 제형은 낮은 레이놀드 수에서 마이크로 채널에서 정상 압력-구동 유동의 스트림을 혼합하기 위한 수동 방법을 포함한다(예를 들어, 전문이 참조에 의해 본 명세서에 포함된 문헌[Abraham *et al.* (2002) Science. 295: 647651] 참조).

[1632] 일 실시형태에서, 본 발명의 circRNA는 마이크로믹서 칩, 예컨대, 비제한적으로 하버드 어패러터스사(Harvard Apparatus)(미국 매사추세츠주 홀리스톤 소재), 돌로마이트 마이크로플루이딕스사(Dolomite Microfluidics)(영국 로이스톤 소재) 또는 프리시젼 나노시스템즈사(Precision Nanosystems)(캐나다 BC주 밴쿠버 소재)로부터의 제품을 사용하여 생성된 지질 나노입자에 제형화될 수 있다. 마이크로믹서 칩은 2종 이상의 유체 스트림을 분할 및 재조합 메커니즘으로 신속하게 혼합하기 위해서 사용될 수 있다.

[1633] 일 실시형태에서, 지질 나노입자는 약 10 내지 약 100nm, 예컨대, 비제한적으로, 약 10 내지 약 20nm, 약 10 내지 약 30nm, 약 10 내지 약 40nm, 약 10 내지 약 50nm, 약 10 내지 약 60nm, 약 10 내지 약 70nm, 약 10 내지 약 80nm, 약 10 내지 약 90nm, 약 20 내지 약 30nm, 약 20 내지 약 40nm, 약 20 내지 약 50nm, 약 20 내지 약 60nm, 약 20 내지 약 70nm, 약 20 내지 약 80nm, 약 20 내지 약 90nm, 약 20 내지 약 100nm, 약 30 내지 약 40 nm, 약 30 내지 약 50nm, 약 30 내지 약 60nm, 약 30 내지 약 70nm, 약 30 내지 약 80nm, 약 30 내지 약 90nm, 약 30 내지 약 100nm, 약 40 내지 약 50nm, 약 40 내지 약 60nm, 약 40 내지 약 70nm, 약 40 내지 약 80nm, 약 40 내지 약 90nm, 약 40 내지 약 100nm, 약 50 내지 약 60nm, 약 50 내지 약 70nm 약 50 내지 약 80nm, 약 50 내지 약 90nm, 약 50 내지 약 100nm, 약 60 내지 약 70nm, 약 60 내지 약 80nm, 약 60 내지 약 90nm, 약 60 내지 약 100nm, 약 70 내지 약 80nm, 약 70 내지 약 90nm, 약 70 내지 약 100nm, 약 80 내지 약 90nm, 약 80 내지 약 100nm 및/또는 약 90 내지 약 100nm의 직경을 가질 수 있다. 일 실시형태에서, 지질 나노입자는 약 10 내지 500nm의 직경을 가질 수 있다. 일 실시형태에서, 지질 나노입자는 100nm 초과, 150nm 초과, 200nm 초과, 250 nm 초과, 300nm 초과, 350nm 초과, 400nm 초과, 450nm 초과, 500nm 초과, 550nm 초과, 600nm 초과, 650nm 초과, 700nm 초과, 750nm 초과, 800nm 초과, 850nm 초과, 900nm 초과, 950nm 초과 또는 1000nm 초과의 직경을 가질 수 있다. 각각의 가능성은 본 발명의 개별 실시형태를 나타낸다.

[1634] 일부 실시형태에서, 나노입자 (예를 들어, 지질 나노입자)는 10 내지 500nm, 20 내지 400nm, 30 내지 300nm 또는 40 내지 200nm의 평균 직경을 갖는다. 일부 실시형태에서, 나노입자 (예를 들어, 지질 나노입자)는 50 내지 150nm, 50 내지 200nm, 80 내지 100nm 또는 80 내지 200nm의 평균 직경을 갖는다.

[1635] 일부 실시형태에서, 본 명세서에 기재된 지질 나노입자는 0.1㎛ 미만 내지 최대 1mm, 예컨대, 비제한적으로 0.1 ㎛ 미만, 1.0㎛ 미만, 5㎛ 미만, 10㎛ 미만, 15㎛ 미만, 20㎛ 미만, 25㎛ 미만, 30㎛ 미만, 35㎛ 미만, 40㎛ 미만, 50㎛ 미만, 55㎛ 미만, 60㎛ 미만, 65㎛ 미만, 70㎛ 미만, 75㎛ 미만, 80㎛ 미만, 85㎛ 미만, 90㎛ 미만, 95㎛ 미만, 100㎛ 미만, 125㎛ 미만, 150㎛ 미만, 175㎛ 미만, 200㎛ 미만, 225㎛ 미만, 250㎛ 미만, 275㎛ 미만, 300㎛ 미만, 325㎛ 미만, 350㎛ 미만, 375㎛ 미만, 400㎛ 미만, 425㎛ 미만, 450㎛ 미만, 475㎛ 미만, 500㎛ 미만, 525㎛ 미만, 550㎛ 미만, 575㎛ 미만, 600㎛ 미만, 625㎛ 미만, 650㎛ 미만, 675㎛ 미만, 700㎛ 미만, 725㎛ 미만, 750㎛ 미만, 775㎛ 미만, 800㎛ 미만, 825㎛ 미만, 850㎛ 미만, 875㎛ 미만, 900㎛ 미만, 925㎛ 미만, 950㎛ 미만, 975㎛ 미만의 직경을 가질 수 있다.

[1636] 또 다른 실시형태에서, LNP는 약 1nm 내지 약 100nm, 약 1nm 내지 약 10nm, 약 1nm 내지 약 20nm, 약 1nm 내지 약 30nm, 약 1nm 내지 약 40nm, 약 1nm 내지 약 50nm, 약 1nm 내지 약 60nm, 약 1nm 내지 약 70nm, 약 1nm 내지 약 80nm, 약 1nm 내지 약 90nm, 약 5nm 내지 약 from 100nm, 약 5nm 내지 약 10nm, 약 5nm 내지 약 20nm, 약 5 nm 내지 약 30nm, 약 5nm 내지 약 40nm, 약 5nm 내지 약 50nm, 약 5nm 내지 약 60nm, 약 5nm 내지 약 70nm, 약 5nm 내지 약 80nm, 약 5nm 내지 약 90nm, 약 10 내지 약 50nm, 약 20 내지 약 50nm, 약 30 내지 약 50nm, 약

40 내지 약 50nm, 약 20 내지 약 60nm, 약 30 내지 약 60nm, 약 40 내지 약 60nm, 약 20 내지 약 70nm, 약 30 내지 약 70nm, 약 40 내지 약 70nm, 약 50 내지 약 70nm, 약 60 내지 약 70nm, 약 20 내지 약 80nm, 약 30 내지 약 80nm, 약 40 내지 약 80nm, 약 50 내지 약 80nm, 약 60 내지 약 80nm, 약 20 내지 약 90nm, 약 30 내지 약 90nm, 약 40 내지 약 90nm, 약 50 내지 약 90nm, 약 60 내지 약 90nm 및/또는 약 70 내지 약 90nm의 직경을 가질 수 있다. 각각의 가능성은 본 발명의 개별 실시형태를 나타낸다.

[1637] 나노입자 조성물은 비교적 균질할 수 있다. 다분산도 지수를 사용하여 나노입자 조성물의 균질성, 예를 들어, 나노입자 조성물의 입자 크기 분포를 나타낼 수 있다. 작은(예를 들어, 0.3 미만) 다분산도 지수는 일반적으로 좁은 입자 크기 분포를 나타낸다. 나노입자 조성물은 약 0 내지 약 0.25, 예컨대, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.20, 0.21, 0.22, 0.23, 0.24 또는 0.25의 다분산도 지수를 가질 수 있다. 일부 실시형태에서, 나노입자 조성물의 다분산도 지수는 약 0.10 내지 약 0.20일 수 있다. 각각의 가능성은 본 발명의 개별 실시형태를 나타낸다.

[1638] 나노입자 조성물의 제타 전위를 사용하여 조성물의 동전기적 전위를 나타낼 수 있다. 예를 들어, 제타 전위는 나노입자 조성물의 표면 전하를 설명할 수 있다. 양전하 또는 음전하의 비교적 낮은 전하를 갖는 나노입자 조성물이 일반적으로 바람직한데, 그 이유는 더 높게 하전된 종은 신체의 세포, 조직 및 다른 요소와 바람직하지 않게 상호작용할 수 있기 때문이다. 일부 실시형태에서, 나노입자 조성물의 제타 전위는 약 -20mV 내지 약 +20mV, 약 -20mV 내지 약 +15mV, 약 -20mV 내지 약 +10mV, 약 -20mV 내지 약 +5mV, 약 -20mV 내지 약 0mV, 약 -20mV 내지 약 -5mV, 약 -20mV 내지 약 -10mV, 약 -20mV 내지 약 -15mV 약 -20mV 내지 약 +20mV, 약 -20mV 내지 약 +15mV, 약 -20mV 내지 약 +10mV, 약 -20mV 내지 약 +5mV, 약 -20mV 내지 약 0mV, 약 0mV 내지 약 +20mV, 약 0mV 내지 약 +15mV, 약 0mV 내지 약 +10mV, 약 0mV 내지 약 +5mV, 약 +5mV 내지 약 +20mV, 약 +5mV 내지 약 +15mV 또는 약 +5mV 내지 약 +10mV일 수 있다. 각각의 가능성은 본 발명의 개별 실시형태를 나타낸다.

[1639] 치료제의 캡슐화 효율은 제공된 초기 양에 대해, 캡슐화되거나 제조 후 나노입자 조성물과 다른 방식으로 회합되는 치료제의 양을 설명한다. 캡슐화 효율은 높은 것(예를 들어, 100%에 가까운 것)이 바람직하다. 캡슐화 효율은, 예를 들어, 나노입자 조성물을 1종 이상의 유기 용매 또는 세제로 분해시키기 전과 후에 나노입자 조성물을 함유하는 용액 중의 치료제의 양을 비교함으로써 측정될 수 있다. 형광은 용액에서 유리 치료제(예를 들어, 핵산)의 양을 측정하는 데 사용될 수 있다. 본 명세서에 기재된 나노입자 조성물의 경우, 치료제의 캡슐화 효율은 적어도 50%, 예를 들어, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%일 수 있다. 일부 실시형태에서, 캡슐화 효율은 적어도 80%일 수 있다. 특정 실시형태에서, 캡슐화 효율은 적어도 90%일 수 있다. 각각의 가능성은 본 발명의 개별 실시형태를 나타낸다. 일부 실시형태에서, 지질 나노입자는 0.4 미만의 다분산도 값을 갖는다. 일부 실시형태에서, 지질 나노입자는 중성 pH에서 순 중성 전하를 갖는다. 일부 실시형태에서, 지질 나노입자는 50 내지 200nm의 평균 직경을 갖는다.

[1640] 지질 나노입자 제형의 특성은 양이온성 지질 성분의 선택, 양이온성 지질 포화도, 비-양이온성 지질 성분의 선택, 비양이온성 지질 포화도, 구조 지질 성분의 선택, PEG화의 성질, 모든 성분의 비율 및 생물물리학적 매개변수, 예컨대, 크기를 포함하지만 이들로 제한되지 않는 인자에 의해서 영향을 받을 수 있다. 본 명세서에 기재된 바와 같이, PEG 지질 성분의 순도가 또한 LNP의 특성 및 성능에 중요하다.

[1641] **12. 방법**

[1642] 일 실시형태에서, 지질 나노입자 제형은 국제 공개 제WO2011127255호 또는 제WO2008103276호에 기재된 방법에 의해 제조될 수 있으며, 이들 각각은 전문이 본 명세서에 참조에 의해 포함된다. 일부 실시형태에서, 지질 나노입자 제형은 전문이 참조에 의해 본 명세서에 포함된 국제 공개 제WO2019131770호에 기재된 바와 같을 수 있다.

[1643] 일부 실시형태에서, 원형 RNA는 미국 특허 출원 15/809,680에 기재된 공정에 따라 제형화된다. 일부 실시형태에서, 본 발명은 지질을 사전 형성된 전달 비히클(즉, RNA의 부재 하에 형성됨)로 형성한 다음 사전 형성된 전달 비히클을 RNA와 조합하는 단계를 포함하는 원형 RNA를 전달 비히클에 캡슐화하는 방법을 제공한다. 일부 실시형태에서, 신규 제형 공정은 시험관내 및 생체내 둘 다에서 지질 나노입자의 사전 형성 단계 없이(예를 들어, 지질을 직접 RNA와 배합함) 제조된 동일한 RNA 제형에 비해 잠재적으로 더 양호한 내약성을 갖는 더 높은 효력(펩타이드 또는 단백질 발현) 및 더 높은 효능(생물학적으로 관련된 종점의 개선)을 갖는 RNA 제형을 생성시킨다.

[1644] RNA의 특정 양이온성 지질 나노입자 제형의 경우, RNA의 높은 캡슐화를 달성하기 위해서, 완충액(예를 들어, 시트르산염 완충액)에서 RNA를 가열시켜야 한다. 이러한 공정 또는 방법에서, 가열은 제형화 후(나노입자의 형성 후) 가열이 지질 나노입자에서 RNA의 캡슐화 효율을 증가시키지 않기 때문에, 제형 공정 이전에 일어날 필요가

있다(즉, 개별 성분 가열). 이에 비해서, 본 발명의 신규한 방법의 일부 실시형태에서, RNA의 가열 순서는 RNA 캡슐화 백분율에 영향을 미치는 것으로 보이지 않는다. 일부 실시형태에서, 사전 형성된 지질 나노입자를 포함하는 용액, RNA를 포함하는 용액 및 지질 나노입자 캡슐화된 RNA를 포함하는 혼합 용액 중 하나 이상의 가열은 제형 공정 이전에 또는 이후에 수행될 필요가 없다(즉, 주위 온도에서 유지).

[1645] RNA는 RNA가 지질 나노입자에 캡슐화될 수 있도록 지질 용액과 혼합될 용액에 제공될 수 있다. 적합한 RNA 용액은 다양한 농도로 캡슐화될 RNA를 함유하는 임의의 수용액일 수 있다 예를 들어, 적합한 RNA 용액은 약 0.01mg/㎖, 0.05mg/㎖, 0.06mg/㎖, 0.07mg/㎖, 0.08mg/㎖, 0.09mg/㎖, 0.1mg/㎖, 0.15mg/㎖, 0.2mg/㎖, 0.3mg/㎖, 0.4mg/㎖, 0.5mg/㎖, 0.6mg/㎖, 0.7mg/㎖, 0.8mg/㎖, 0.9mg/㎖ 또는 1.0mg/㎖ 초과의 농도로 RNA를 함유할 수 있다. 일부 실시형태에서, 적합한 RNA 용액은 약 0.01 내지 1.0mg/㎖, 0.01 내지 0.9mg/㎖, 0.01 내지 0.8mg/㎖, 0.01 내지 0.7mg/㎖, 0.01 내지 0.6mg/㎖, 0.01 내지 0.5mg/㎖, 0.01 내지 0.4mg/㎖, 0.01 내지 0.3mg/㎖, 0.01 내지 0.2mg/㎖, 0.01 내지 0.1mg/㎖, 0.05 내지 1.0mg/㎖, 0.05 내지 0.9mg/㎖, 0.05 내지 0.8mg/㎖, 0.05 내지 0.7mg/㎖, 0.05 내지 0.6mg/㎖, 0.05 내지 0.5mg/㎖, 0.05 내지 0.4mg/㎖, 0.05 내지 0.3mg/㎖, 0.05 내지 0.2mg/㎖, 0.05 내지 0.1mg/㎖, 0.1 내지 1.0mg/㎖, 0.2 내지 0.9mg/㎖, 0.3 내지 0.8mg/㎖, 0.4 내지 0.7mg/㎖ 또는 0.5 내지 0.6mg/㎖의 범위의 농도로 RNA를 함유할 수 있다.

[1646] 전형적으로, 적합한 RNA 용액은 또한 완충제 및/또는 염을 함유할 수 있다. 일반적으로, 완충액은 HEPES, Tris, 황산암모늄, 중탄산나트륨, 시트르산나트륨, 아세트산나트륨, 인산칼륨 또는 인산나트륨을 포함할 수 있다. 일부 실시형태에서, 완충제의 적합한 농도는 약 0.1mM 내지 100mM, 0.5mM 내지 90mM, 1.0mM 내지 80mM, 2mM 내지 70mM, 3mM 내지 60mM, 4mM 내지 50mM, 5mM 내지 40mM, 6mM 내지 30mM, 7mM 내지 20mM, 8mM 내지 15 mM 또는 9 내지 12 mM의 범위일 수 있다.

[1647] 예시적인 염은 염화나트륨, 염화마그네슘 및 염화칼륨을 포함할 수 있다. 일부 실시형태에서, RNA 용액 중의 염의 적합한 농도는 약 1mM 내지 500mM, 5mM 내지 400mM, 10mM 내지 350mM, 15mM 내지 300mM, 20mM 내지 250mM, 30mM 내지 200mM, 40mM 내지 190mM, 50mM 내지 180mM, 50mM 내지 170mM, 50mM 내지 160mM, 50mM 내지 150mM 또는 50mM 내지 100 mM의 범위일 수 있다.

[1648] 일부 실시형태에서, 적합한 RNA 용액은 약 3.5 내지 6.5, 3.5 내지 6.0, 3.5 내지 5.5, 3.5 내지 5.0, 3.5 내지 4.5, 4.0 내지 5.5, 4.0 내지 5.0, 4.0 내지 4.9, 4.0 내지 4.8, 4.0 내지 4.7, 4.0 내지 4.6 또는 4.0 내지 4.5의 범위의 pH를 가질 수 있다.

[1649] 본 발명에 적합한 RNA 용액을 제조하기 위해 다양한 방법이 사용될 수 있다. 일부 실시형태에서, RNA는 본 명세서에 기재된 완충 용액에 직접 용해될 수 있다. 일부 실시형태에서, RNA 용액은 캡슐화를 위해 지질 용액과 혼합하기 전에 RNA 스톡 용액을 완충 용액과 혼합함으로써 생성될 수 있다. 일부 실시형태에서, RNA 용액은 캡슐화를 위해 지질 용액과 혼합하기 직전에 RNA 스톡 용액을 완충 용액과 혼합함으로써 생성될 수 있다.

[1650] 본 발명에 따르면, 지질 용액은 RNA의 캡슐화를 위한 전달 비히클을 형성하기에 적합한 지질의 혼합물을 함유한다. 일부 실시형태에서, 적합한 지질 용액은 에탄올 기재이다. 예를 들어, 적합한 지질 용액은 순수한 에탄올(즉, 100% 에탄올)에 용해된 목적하는 지질의 혼합물을 함유할 수 있다. 또 다른 실시형태에서, 적합한 지질 용액은 아이소프로필 알코올 기재이다. 또 다른 실시형태에서, 적합한 지질 용액은 다이메틸설폭사이드 기재이다. 또 다른 실시형태에서, 적합한 지질 용액은 에탄올, 아이소프로필 알코올 및 다이메틸설폭사이드를 포함하지만 이들로 제한되지 않는 적합한 용매의 혼합물이다.

[1651] 적합한 지질 용액은 다양한 농도에서 목적하는 지질의 혼합물을 함유할 수 있다. 일부 실시형태에서, 적합한 지질 용액은 약 0.1 내지 100mg/㎖, 0.5 내지 90mg/㎖, 1.0 내지 80mg/㎖, 1.0 내지 70mg/㎖, 1.0 내지 60mg/㎖, 1.0 내지 50mg/㎖, 1.0 내지 40mg/㎖, 1.0 내지 30mg/㎖, 1.0 내지 20mg/㎖, 1.0 내지 15mg/㎖, 1.0 내지 10mg/㎖, 1.0 내지 9mg/㎖, 1.0 내지 8mg/㎖, 1.0 내지 7mg/㎖, 1.0 내지 6mg/㎖ 또는 1.0 내지 5mg/㎖의 범위의 총 농도로 목적하는 지질의 혼합물을 함유할 수 있다.

[1652] **13. 표적화**

[1653] 본 발명은 또한 수동 및 능동 표적화 수단 둘 다에 의한 표적 세포 및 조직의 차별적 표적화를 고려한다. 수동 표적화 현상은 표적 세포에 의한 전달 비히클의 인식을 향상시키기 위한 수단 또는 추가 부형제의 사용에 의존하지 않고 생체내에서 전달 비히클의 자연 분포 패턴을 이용한다. 예를 들어, 세망-내피 시스템의 세포에 의해 식균 작용을 받는 전달 비히클은 간 또는 비장에 축적될 가능성이 있고, 따라서 이러한 표적 세포에 대한 조성물의 전달을 수동적으로 지시하는 수단을 제공할 수 있다.

[1654] 대안적으로, 본 발명은 특정 표적 세포 또는 표적 조직에서 이러한 전달 비히클의 국지화를 촉진하기 위해 전달 비히클에 (공유적으로 또는 비공유적으로) 결합될 수 있는 표적화 모이어티의 사용을 포함하는 능동 표적화를 고려한다. 예를 들어, 표적화는 표적 세포 또는 조직으로의 분포를 장려하기 위해 전달 비히클 내에 또는 그 상에 하나 이상의 내인성 표적화 모이어티를 포함시킴으로써 매개될 수 있다. 표적 조직에 의한 표적화 모이어티의 인식은 표적 세포 및 조직에서 전달 비히클 및/또는 그 내용물의 조직 분포 및 세포 흡수를 적극적으로 촉진시킨다(예를 들어, 전달 비히클 내에 또는 그 상에 아포지단백질-E 표적화 리간드를 포함시키는 것은 간세포에 의해 발현되는 내인성 저밀도 지단백질 수용체에 대한 전달 비히클의 인식 및 결합을 장려한다). 본 명세서에 제공된 바와 같이, 조성물은 표적 세포에 대한 조성물의 친화도를 향상시킬 수 있는 모이어티를 포함할 수 있다. 표적화 모이어티는 제형 동안 또는 제형 후에 지질 입자의 외부 이중층에 연결될 수 있다. 이들 방법은 당업계에 널리 공지되어 있다. 또한, 일부 지질 입자 제형은 PEAA, 혈구응집소, 기타 리포펩타이드(미국 특허 출원 제08/835,281호 및 제60/083,294호 참조, 본 명세서에 참조에 의해 포함됨)와 같은 융합성 중합체 및 생체내 및/또는 세포내 전달에 유용한 기타 특징을 사용할 수 있다. 다른 일부 실시형태에서, 본 발명의 조성물은 개선된 형질주입 효능을 입증하고/하거나 관심 표적 세포 또는 조직에 대한 향상된 선택성을 입증한다. 따라서 표적 세포 또는 조직에 대한 조성물의 친화도 및 이들의 핵산 함량을 향상시킬 수 있는 하나 이상의 모이어티(예를 들어, 펩타이드, 압타머, 올리고뉴클레오타이드, 비타민 또는 기타 분자)를 포함하는 조성물이 고려된다. 적합한 모이어티는 선택적으로 전달 비히클의 표면에 결합되거나 연결될 수 있다. 일부 실시형태에서, 표적화 모이어티는 전달 비히클의 표면에 걸쳐 있거나 전달 비히클 내에 캡슐화될 수 있다. 적절한 모이어티는 물리적, 화학적 또는 생물학적 특성(예를 들어, 표적 세포 표면 마커 또는 특징의 선택적 친화성 및/또는 인식)을 기반으로 선택된다. 세포-특이적 표적 부위와 이에 상응하는 표적 리간드는 매우 다양할 수 있다. 표적 세포의 고유한 특성이 이용되어, 조성물이 표적 세포와 비-표적 세포를 구별할 수 있도록 적합한 표적화 모이어티가 선택된다. 예를 들어, 본 발명의 조성물은 (예를 들어, 이러한 표면 마커의 수용체-매개 인식 및 결합에 의해) 간세포의 인식 또는 친화성을 선택적으로 향상시키는 표면 마커 (예를 들어, 아포지단백질-B 또는 아포지단백질-E)를 포함할 수 있다. 예를 들어, 표적화 모이어티로서 갈락토스의 사용은 본 발명의 조성물을 실질 간세포로 안내할 것으로 예측되거나, 대안적으로 표적 리간드로서 당 잔기를 함유하는 만노스의 사용은 본 발명의 조성물을 간내피 세포(예를 들어, 간세포에 존재하는 아시알로당단백질 수용체에 우선적으로 결합할 수 있는 당 잔기를 함유하는 만노스)로 안내할 것이라고 예측될 것이다(문헌[Hillery A M, *et al.* "Drug Delivery and Targeting: For Pharmacists and Pharmaceutical Scientists" (2002) Taylor & Francis, Inc.] 참조) 전달 비히클(예를 들어, 따라서 지질 나노입자)에 존재하는 모이어티에 접합된 이러한 표적화 모이어티의 제시는 표적 세포 및 조직에서 본 발명의 조성물의 인식 및 흡수를 촉진한다. 적합한 표적화 모이어티의 예는 하나 이상의 펩타이드, 단백질, 압타머, 비타민 및 올리고뉴클레오타이드를 포함한다.

[1655] 특정 실시형태에서, 전달 비히클은 표적화 모이어티를 포함한다. 일부 실시형태에서, 표적화 모이어티는 특이적 세포의 집단 내에서 수용체-매개된 엔도사이토시스를 세포의 특정 집단으로 선택적으로 매개한다. 일부 실시형태에서, 표적화 모이어는 T 세포 항원에 결합할 수 있다. 일부 실시형태에서, 표적화 모이어티는 NK, NKT 또는 대식세포 항원에 결합할 수 있다. 일부 실시형태에서, 표적화 모이어티는 군 CD3, CD4, CD8, PD-1, 4-1BB 및 CD2으로부터 선택된 단백질에 결합할 수 있다. 일부 실시형태에서, 표적화 모이어티는 단일 쇄 Fv(scFv) 단편, 나노바디, 펩타이드, 펩타이드-기재 마크로사이클, 미니바디, 중쇄 가변 영역, 경쇄 가변 영역 또는 이의 단편이다. 일부 실시형태에서, 표적화 모이어티는 항 T-세포 수용체 모티프 항체, T-세포 α 쇄 항체, T-세포 β 쇄 항체, T-세포 γ 쇄 항체, T-세포 δ 쇄 항체, CCR7 항체, CD3 항체, CD4 항체, CD5 항체, CD7 항체, CD8 항체, CD11b 항체, CD11c 항체, CD16 항체, CD19 항체, CD20 항체, CD21 항체, CD22 항체, CD25 항체, CD28 항체, CD34 항체, CD35 항체, CD40 항체, CD45RA 항체, CD45RO 항체, CD52 항체, CD56 항체, CD62L 항체, CD68 항체, CD80 항체, CD95 항체, CD117 항체, CD127 항체, CD133 항체, CD137(4-1BB) 항체, CD163 항체, F4/80 항체, IL-4Rα 항체, Sca-1 항체, CTLA-4 항체, GITR 항체, GARP 항체, LAP 항체, 그랜자임 B 항체, LFA-1 항체, 트랜스페린 수용체 항체, 및 이들의 단편으로부터 선택된다. 일부 실시형태에서, 표적화 모이어티는 림프구 상의 체외효소의 소분자 결합제이다. 체외효소의 소분자 결합제는 A2A 저해제 CD73 저해제, CD39 또는 아데신 수용체 A2aR 및 A2bR을 포함한다. 잠재적인 소분자는 AB928을 포함한다.

[1656] 일부 실시형태에서, 전달 비히클은 문헌[Shobaki N, Sato Y, Harashima H. Mixing lipids to manipulate the ionization status of lipid nanoparticles for specific tissue targeting. Int J Nanomedicine. 2018;13:8395-8410. Published 2018 Dec 10]에 기재된 바와 같이 제형화되고/되거나 표적화된다. 일부 실시형태에서, 전달 비히클은 3개의 지질 유형으로 구성된다. 일부 실시형태에서, 전달 비히클은 4개의 지질 유형으로 구성된다. 일부 실시형태에서, 전달 비히클은 5개의 지질 유형으로 구성된다. 일부 실시형태에서, 전달 비히클

은 6개의 지질 유형으로 구성된다.

[1657] **14. 표적 세포**

[1658] 핵산을 면역 세포로 전달하는 것이 바람직한 경우, 면역 세포가 표적 세포를 나타낸다. 일부 실시형태에서, 본 발명의 조성물은 차별적 기반으로 표적 세포를 형질주입시킨다(즉, 비-표적 세포를 형질주입시키지 않는다). 본 발명의 조성물은 또한 T 세포, B 세포, 대식세포 및 수지상 세포를 포함하지만 이들로 제한되지 않는 다양한 표적 세포를 우선적으로 표적으로 하도록 제조될 수 있다.

[1659] 일부 실시형태에서, 표적 세포는 관심 단백질 또는 효소가 결핍되어 있다. 예를 들어, 핵산을 간세포에 전달하고자 하는 경우, 간세포는 표적 세포를 나타낸다. 일부 실시형태에서, 본 발명의 조성물은 차별적 기반으로 표적 세포를 형질주입시킨다(즉, 비-표적 세포를 형질주입시키지 않는다). 본 발명의 조성물은 또한 간세포, 상피 세포, 조혈 세포, 상피 세포, 내피 세포, 폐 세포, 골 세포, 줄기 세포, 중간엽 세포, 신경 세포(예를 들어, 수막, 성상교세포, 운동 뉴런, 후근 신경절 및 전각 운동 뉴런의 세포), 광수용체 세포(예를 들어, 간상체 및 원추체), 망막 색소 상피 세포, 분비 세포, 심장 세포, 지방세포, 혈관 평활근 세포, 심근세포, 골격근 세포, 베타 세포, 뇌하수체 세포, 활액 내막 세포, 난소 세포, 고환 세포, 섬유모세포, B 세포, T 세포, 망상적혈구, 백혈구, 과립구 및 종양 세포을 포함하지만 이들로 제한되지 않는 다양한 표적 세포를 우선적으로 표적으로 하도록 제조될 수 있다.

[1660] 본 발명의 조성물은 심장, 폐, 신장, 간 및 비장과 같은 표적 세포에 우선적으로 분포되도록 제조될 수 있다. 일부 실시형태에서, 본 발명의 조성물은 간 또는 비장의 세포(예를 들어, 간세포)에 의해 그 내에 포함된 circRNA의 전달 및 후속 발현을 촉진하기 위해 간의 세포 또는 비장의 세포(예를 들어, 면역 세포) 내에 분포한다. 표적화 세포는 기능성 단백질 또는 효소를 생산하고 전신에 배출할 수 있는 생물학적 "저장부" 또는 "저장소"로 기능할 수 있다. 따라서, 본 발명의 일 실시형태에서 전달 비히클은 간세포 또는 면역 세포를 표적화하고/하거나 전달 시 간 또는 비장의 세포에 우선적으로 분포할 수 있다. 일 실시형태에서, 표적 간세포 또는 면역 세포의 형질주입 후, 비히클에 로딩된 circRNA가 번역되고, 기능성 단백질 산물이 생산되고, 배설되고, 전신에 분포된다. 다른 실시형태에서, 간세포 이외의 세포(예를 들어, 폐, 비장, 심장, 안구 또는 중추 신경계의 세포)는 단백질 생산을 위한 저장소 위치로 작용할 수 있다.

[1661] 일 실시형태에서, 본 발명의 조성물은 1종 이상의 기능성 단백질 및/또는 효소의 대상체의 내인성 생산을 촉진한다. 본 발명의 일 실시형태에서, 전달 비히클은 결핍된 단백질 또는 효소를 암호화하는 circRNA를 포함한다. 표적 조직으로의 이러한 조성물의 분포 및 이러한 표적 세포의 후속적인 형질주입 시, 전달 비히클(예를 들어, 지질 나노입자)에 로딩된 외인성 circRNA는 생체내에서 번역되어 외부에서 투여되는 circRNA(예를 들어, 대상체가 결핍된 단백질 또는 효소) 의해 암호화된 기능적 단백질 또는 효소를 생성할 수 있다. 따라서, 본 발명의 조성물은 내인적으로 번역된 단백질 또는 효소를 생산하기 위해 외인성으로 또는 재조합 방식으로 제조된 circRNA를 번역하고, 이에 의해 기능성 단백질 또는 효소를 생산(및 적용 가능한 경우 배설)하는 대상체의 능력을 이용한다. 발현되거나 번역된 단백질 또는 효소는 또한 재조합에 의해서 제조된 단백질 또는 효소에 종종 없을 수 있는 천연 번역후 변형의 생체내 포함을 특징으로 할 수 있으며, 이에 의해 번역된 단백질 또는 효소의 면역원성을 추가로 감소시킬 수 있다.

[1662] 결핍된 단백질 또는 효소를 암호화하는 circRNA의 투여는 표적 세포 내의 특정 소기관에 핵산을 전달할 필요성을 회피한다. 오히려, 표적 세포의 형질주입 및 표적 세포의 세포질로 핵산의 전달 시에, 전달 비히클의 circRNA 내용물이 번역될 수 있고 기능성 단백질 또는 효소가 발현될 수 있다.

[1663] 일부 실시형태에서, 원형 RNA는 하나 이상의 miRNA 결합 부위를 포함한다. 일부 실시형태에서, 원형 RNA는 하나 이상의 비표적 세포 또는 비표적 세포 유형(예를 들어, 쿠퍼 세포 또는 간세포)에 존재하고 하나 이상의 표적 세포 또는 표적 세포 유형(예를 들어, 간세포 또는 T 세포)에는 존재하지 않는 miRNA에 의해 인식되는 하나 이상의 miRNA 결합 부위를 포함한다. 일부 실시형태에서, 원형 RNA는 하나 이상의 표적 세포 또는 표적 세포 유형(예를 들어, 간세포 또는 T 세포)에 비해서 하나 이상의 비표적 세포 또는 비표적 세포 유형(예를 들어, 쿠퍼 세포 또는 간세포)에 증가된 농도로 존재하는 miRNA에 의해 인식되는 하나 이상의 miRNA 결합 부위를 포함한다. miRNA는 RNA 분자 내의 상보적인 서열과 짝을 이루어 기능을 하여 유전자 침묵을 초래한다고 생각된다.

[1664] 일부 실시형태에서, 본 발명의 조성물은 차별적 기반으로 표적 세포를 형질주입시키거나 표적 세포에 분포시킨다(즉, 비-표적 세포를 형질주입시키지 않는다). 본 발명의 조성물은 또한 간세포, 상피 세포, 조혈 세포, 상피 세포, 내피 세포, 폐 세포, 골 세포, 줄기 세포, 중간엽 세포, 신경 세포(예를 들어, 수막, 성상교세포, 운동

뉴런, 후근 신경절 및 전각 운동 뉴런의 세포), 광수용체 세포(예를 들어, 간상체 및 원추체), 망막 색소 상피 세포, 분비 세포, 심장 세포, 지방세포, 혈관 평활근 세포, 심근세포, 골격근 세포, 베타 세포, 뇌하수체 세포, 활액 내막 세포, 난소 세포, 고환 세포, 섬유모세포, B 세포, T 세포, 망상적혈구, 백혈구, 과립구 및 종양 세포을 포함하지만 이들로 제한되지 않는 다양한 표적 세포를 우선적으로 표적으로 하도록 제조될 수 있다.

[1665] **15. 약제학적 조성물**

[1666] 특정 실시형태에서, 본 명세서에 제공된 치료제를 포함하는 조성물 (예를 들어, 약제학적 조성물)이 본 명세서에 제공된다. 일부 실시형태에서, 치료제는 본 명세서에 제공된 원형 RNA 폴리뉴클레오타이드이다. 일부 실시형태에서 치료제는 본 명세서에서 제공되는 벡터이다. 일부 실시형태에서, 치료제는 본 명세서에 제공된 원형 RNA 또는 벡터를 포함하는 세포(예를 들어, 인간 세포, 예컨대, 인간 T 세포)이다. 특정 실시형태에서, 조성물은 약제학적으로 허용 가능한 담체를 추가로 포함한다. 일부 실시형태에서, 본 명세서에 제공된 조성물은 다른 약제학적 활성제 또는 약물, 예컨대, 항염증성 약물 또는 B 세포 항원을 표적으로 할 수 있는 항체, 예를 들어, 항-CD20 항체, 예를 들어, 리툭시맙과 조합하여 본 명세서에 제공된 치료제를 포함한다.

[1667] 약제학적 조성물과 관련하여, 약제학적으로 허용 가능한 담체는 통상적으로 사용되는 임의의 것일 수 있으며, 활성제(들)와의 용해도 및 반응성 부족과 같은 화학적-물리적 고려사항 및 투여 경로에 의해서만 제한된다. 본 명세서에 기재된 약제학적으로 허용 가능한 담체, 예를 들어, 비히클, 보조제, 부형제 및 희석제는 당업자에게 잘 알려져 있고 대중에게 용이하게 입수 가능하다. 약제학적으로 허용 가능한 담체는 치료제(들)에 대해 화학적으로 불활성이고 사용 조건 하에서 유해한 부작용 또는 독성이 없는 것이 바람직하다.

[1668] 담체의 선택은 특정 치료제 뿐만아니라 치료제를 투여하는데 사용되는 특정 방법에 의해 부분적으로 결정될 것이다. 따라서, 본 명세서에서 제공되는 약제학적 조성물의 다양한 적합한 제형이 존재한다.

[1669] 특정 실시형태에서, 약제학적 조성물은 보존제를 포함한다 특정 실시형태에서, 적합한 보존제는 예를 들어, 메틸파라벤, 프로필파라벤, 벤조산나트륨, 및 염화벤즈알코늄을 포함할 수 있다. 선택적으로, 둘 이상의 보존제의 혼합물이 사용될 수 있다. 보존제 또는 이의 혼합물은 일반적으로 전체 조성물의 약 0.0001중량% 내지 약 2중량%의 양으로 존재한다.

[1670] 일부 실시형태에서, 약제학적 조성물은 완충제를 포함한다. 일부 실시형태에서, 적합한 완충제는, 예를 들어, 시트르산, 시트르산나트륨, 인산, 인산칼륨, 및 다양한 기타 산 및 염을 포함할 수 있다. 둘 이상의 완충제의 혼합물이 선택적으로 사용될 수 있다. 완충제 또는 이들의 혼합물은 일반적으로 총 조성물의 약 0.001중량% 내지 약 4중량%의 양으로 존재한다.

[1671] 일부 실시형태에서, 약제학적 조성물 중의 치료제의 농도는 예를 들어, 중량 기준으로 약 1% 미만, 또는 적어도 약 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% 또는 약 50% 또는 그 초과이고, 선택된 특정 투여 모드에 따라서, 유체 부피 및 점도에 의해서 주로 선택될 수 있다.

[1672] 경구, 에어로졸, 비경구(예를 들어, 피하, 정맥내, 동맥내, 근육내, 피내, 복강내 및 척추강내) 및 국소 투여를 위한 하기 제형은 단지 예시적이며 어떠한 방식으로도 제한이 아니다. 하나 이상의 경로가 본 명세서에 제공된 치료제를 투여하는 데 사용될 수 있으며, 특정 경우에 특정 경로는 다른 경로보다 더 즉각적이고 효과적인 반응을 제공할 수 있다.

[1673] 경구 투여에 적합한 제형은 (a) 액체 용액, 예컨대, 물, 식염수 또는 오렌지 주스와 같은 희석제에 용해된 유효량의 치료제; (b) 캡슐, 사쉐(sachet), 정제, 로젠지 및 트로키(각각 미리 결정된 양의 활성 성분을 고체 또는 과립 형태로 함유함); (c) 분말 (d) 적절한 액체의 현탁액 및 (e) 적합한 에멀션을 포함하거나 이들로 이루어질 수 있다. 액체 제형은 약제학적으로 허용 가능한 계면활성제의 첨가 여부에 관계없이 물 및 알코올, 예를 들어, 에탄올, 벤질 알코올 및 폴리에틸렌 알코올과 같은 희석제를 포함할 수 있다. 캡슐 형태는 예를 들어, 계면 활성제, 윤활제 및 락토스, 자당, 인산 칼슘 및 옥수수 전분과 같은 불활성 충전제를 함유하는 일반적인 경질 또는 연질 쉘 젤라틴 유형일 수 있다. 정제는 하나 이상의 락토스, 수크로스, 만니톨, 옥수수 전분, 감자 전분, 알긴산, 미결정 셀룰로오스, 아카시아, 젤라틴, 구아검, 콜로이드성 이산화규소, 크로스카멜로스 소듐, 탤크, 마그네슘 스테아레이트, 칼슘 스테아레이트, 스테아르 산 아연, 스테아르산, 및 다른 보존제, 향미제, 및 약제학적으로 허용 가능한 분해제, 습윤제, 보존제, 향미제, 및 약리학적으로 호환 가능한 부형제를 포함할 수 있다. 로젠지 형태는 일반적으로 자당, 아카시아 또는 트래거캔스와 같은 향미제와 치료제를 포함할 수 있다. 알약(pastille)은 젤라틴 및 글리세린, 또는 수크로스 및 아카시아, 에멀션, 겔 등과 같은 불활성 염기를 갖는 치료제를 포함할 수 있으며, 이에 더하여 당업계에 공지된 이러한 부형제를 함유할 수 있다.

[1674] 비경구 투여에 적합한 제형은 항산화제, 완충액, 정균제, 및 제형을 의도된 수용자의 혈액과 등장성으로 만드는 용질을 함유할 수 있는 수성 및 비수성 등장성 멸균 주사 용액, 및 현탁화제, 가용화제, 증점제, 안정제 및 보존제를 포함할 수 있는 수성 및 비수성 멸균 현탁액을 포함한다. 일부 실시형태에서, 본 명세서에 제공된 치료제는 약제학적 담체, 예컨대, 멸균 액체 또는 물, 염수, 수성 덱스트로스 및 관련 당 용액을 포함하는 액체의 혼합물, 알코올, 예컨대, 에탄올 또는 헥사데실 알코올, 글리콜, 예컨대 프로필렌 글리콜 또는 폴리에틸렌 글리콜, 다이메틸설폭사이드, 글리세롤, ketals 예컨대 2,2-다이메틸-1,3-다이옥솔란-4-메탄올, 에터, 폴리(에틸렌 글리콜) 400, 오일, 지방산, 지방산 에스터 또는 글리세리드 또는 아세틸화된 지방산 글리세리드(약제학적으로 허용 가능한 계면활성제, 예컨대, 비누 또는 세제, 현탁화제, 예컨대, 펙틴, 카보머, 메틸셀룰로스, 하이드록시프로필메틸셀룰로스 또는 카복시메틸셀룰로스 또는 유화제 및 다른 약제학적 아주반트를 함유하거나 함유하지 않음) 중의 생리학적으로 허용 가능한 희석제로 투여될 수 있다.

[1675] 일부 실시형태에서 비경구 제형에 사용될 수 있는 오일은 석유, 동물성 오일, 식물성 오일, 또는 합성 오일을 포함한다. 오일의 특정 예는 땅콩, 대두, 참깨, 면실유, 옥수수, 올리브, 바셀린 및 광유를 포함한다. 비경구 제형에 사용하기에 적합한 지방산은 올레산, 스테아르산, 및 아이소스테아르산을 포함한다. 에틸 올레에이트 및 아이소프로필 미리스테이트는 적합한 지방산 에스터의 예이다.

[1676] 비경구 제형의 특정 실시형태에서 사용하기에 적합한 비누는 지방 알칼리 금속, 암모늄, 및 트라이에탄올아민 염을 포함하고, 적합한 세제에는 (a) 양이온성 세제, 예를 들어, 다이메틸 다이알킬 암모늄 할라이드 및 알킬 피리디늄 할라이드, (b) 음이온성 세제, 예를 들어, 알킬, 아릴 및 올레핀 설포네이트, 알킬, 올레핀, 에터, 및 모노글리세라이드 설페이트, 및 설포석시네이트, (c) 비이온성 세제, 예를 들어, 지방 아민 산화물, 지방산 알칸올아마이드, 및 폴리옥시에틸렌폴리프로필렌 공중합체, (d) 양쪽성 세제, 예를 들어, 알킬-β-아미노프로피오네이트, 및 2-알킬-이미다졸린 4차 암모늄 염, 및 (e) 이들의 혼합물을 포함한다.

[1677] 일부 실시형태에서, 비경구 제형은 예를 들어, 용액 중 치료제의 약 0.5 중량% 내지 약 25 중량%를 함유할 것이다. 보존제 및 완충액을 사용할 수 있다. 주사 부위의 자극을 최소화하거나 제거하기 위해, 이러한 조성물은 예를 들어, 친수성-친유성 균형(HLB)이 약 12 내지 약 17인 하나 이상의 비이온성 계면활성제를 함유할 수 있다. 이러한 제형 중 계면활성제의 양은 전형적으로, 예를 들어, 약 5 중량% 내지 약 15 중량%의 범위일 것이다. 적합한 계면활성제는 폴리에틸렌 글리콜, 소르비탄 모노올레에이트와 같은 소르비탄 지방산 에스터, 및 프로필렌 글리콜과 프로필렌 글리콜의 축합에 의해 형성된 소수성 염기를 갖는 에틸렌 옥사이드의 고분자량 부가물을 포함한다. 비경구 제형은 앰플 또는 바이알과 같은 단위 용량 또는 다중 용량 밀봉 용기로 제공될 수 있으며, 주사의 경우 사용 직전에 물과 같은 멸균 액체 부형제만 추가하면 되는 냉동 건조(동결 건조) 상태로 저장될 수 있다. 즉석 주사 용액 및 현탁액은 앞서 설명한 종류의 멸균 분말, 과립 및 정제로부터 제조할 수 있다.

[1678] 특정 실시형태에서, 주사 가능한 제형이 본 명세서에 제공된다. 주사 가능한 조성물을 위한 효과적인 약제학적 담체에 대한 요건은 당업자에게 널리 공지되어 있다(예를 들어, 문헌[Pharmaceutics and Pharmacy Practice, J.B. Lippincott Company, Philadelphia, PA, Banker and Chalmers, eds., pages 238-250 (1982), 및 ASHP Handbook on Injectable Drugs, Toissel, 4th ed, pages 622-630 (1986))] 참조).

[1679] 일부 실시형태에서, 국소 제형이 본 명세서에 제공된다. 경피 약물 방출에 유용한 것을 포함하는 국소 제형은 피부에 적용하기 위해 본 명세서에 제공된 특정 실시형태의 맥락에서 적합하다. 일부 실시형태에서, 치료제는 단독으로 또는 다른 적합한 성분과 조합하여 흡입을 통해 투여될 에어로졸 제형으로 제조될 수 있다. 이러한 에어로졸 제형은 다이클로로디플루오로메탄, 프로판, 질소 등과 같은 가압된 허용 가능한 추진제에 배치될 수 있다. 이들은 또한 네불라이저 또는 아토마이저와 같은 비압축 제제용 약제로 제형화될 수 있다. 이러한 스프레이 제형은 또한 점막을 스프레이하는 데 사용될 수 있다.

[1680] 특정 실시형태에서, 본 명세서에 제공된 치료제는 사이클로덱스트린 내포 복합체 또는 리포솜과 같은 내포 복합체로서 제형화될 수 있다. 리포솜은 특정 조직에 치료제를 표적화하는 역할을 할 수 있다. 리포솜은 또한 치료제의 반감기를 증가시키는 데 사용될 수 있다. 리포솜을 제조하기 위한 다양한 방법이 사용 가능하고, 예를 들어, 문헌[Szoka *et al.*, Ann. Rev. Biophys. Bioeng., 9, 467 (1980)] 및 미국 특허 제4,235,871호, 제4,501,728호, 제4,837,028호 및 제5,019,369호에 기재되어 있다.

[1681] 일부 실시형태에서, 본 명세서에서 제공되는 치료제는 시간-방출, 지연 방출 또는 지속 방출 전달 시스템으로 제형화되어, 조성물의 전달이 치료될 부위의 감작을 야기하기에 충분한 시간과 함께 일어나도록 한다. 이러한 시스템은 치료제의 반복 투여를 피할 수 있고, 이에 의해 대상체 및 의사에게 편의성을 증가시킬 수 있고, 본 명세서에 제공된 특정 조성물 실시형태에 특히 적합할 수 있다. 일 실시형태에서, 본 발명의 조성물은 그 내에

함유된 circRNA의 연장 방출에 적합하도록 제형화된다. 이러한 연장 방출 조성물은 연장된 투여 간격으로 대상체에게 편리하게 투여될 수 있다. 예를 들어, 일 실시형태에서, 본 발명의 조성물은 1일 2회, 일 단위로 또는 격일로 대상체에게 투여된다. 일 실시형태에서, 본 발명의 조성물은 대상체에게 주 2회, 주 1회, 10일마다, 2주마다, 3주마다, 4주마다, 1개월에 1회, 6주마다, 8주마다, 3개월마다, 4개월마다, 6개월마다, 8개월마다, 9개월마다 또는 매년 투여된다.

[1682] 일부 실시형태에서, 본 발명의 폴리뉴클레오타이드에 의해 암호화된 단백질은 지속된 시간 동안 표적 세포에 의해 생산된다. 예를 들어, 단백질은 투여 후 1시간 초과, 4시간 초과, 6시간 초과, 12시간 초과, 24시간 초과, 48시간 초과, 또는 72시간 초과 동안 생산될 수 있다. 일부 실시형태에서 폴리펩타이드는 투여 후 약 6시간 후에 최고 수준으로 발현된다. 일부 실시형태에서 폴리펩타이드의 발현은 적어도 치료 수준에서 지속된다. 일부 실시형태에서, 폴리펩타이드는 투여 후 1시간 초과, 4시간 초과, 6시간 초과, 12시간 초과, 24시간 초과, 48시간 초과, 또는 72시간 초과 동안 적어도 치료 수준으로 발현된다. 일부 실시형태에서, 폴리펩타이드는 환자 조직(예를 들어, 간 또는 폐)에서 치료 수준에서 검출 가능하다. 일부 실시형태에서, 검출 가능한 폴리펩타이드의 수준은 투여 후 1시간 초과, 4시간 초과, 6시간 초과, 12시간 초과, 24시간 초과, 48시간 초과, 또는 72시간 초과의 기간에 걸쳐 circRNA 조성물로부터의 연속 발현으로부터 유래한다.

[1683] 특정 실시형태에서, 본 발명의 폴리뉴클레오타이드에 의해 암호화되는 단백질은 정상 생리학적 수준보다 높은 수준으로 생산된다. 단백질 수준은 대조군에 비해 증가할 수 있다. 일부 실시형태에서, 대조군은 정상 개체 또는 정상 개체 집단에서 폴리펩타이드의 기준선 생리학적 수준이다. 다른 실시형태에서, 대조군은 관련 단백질 또는 폴리펩타이드가 결핍된 개체 또는 관련 단백질 또는 폴리펩타이드가 결핍된 개체 집단에서 폴리펩타이드의 기준선 생리학적 수준이다. 일부 실시형태에서, 대조군은 조성물이 투여되는 개체에서 관련 단백질 또는 폴리펩타이드의 정상 수준일 수 있다. 다른 실시형태에서, 대조군은 다른 치료적 개입 시, 예를 들어, 하나 이상의 유사한 시점에서 상응하는 폴리펩타이드의 직접 주사 시 폴리펩타이드의 발현 수준이다.

[1684] 특정 실시형태에서, 본 발명의 폴리뉴클레오타이드에 의해 암호화되는 단백질의 수준은 투여 후 3일, 4일, 5일, 또는 1주에 또는 그보다 더 긴 시간 후에 검출 가능하다. 증가된 수준의 단백질은 조직(예를 들어, 간 또는 폐)에서 관찰될 수 있다.

[1685] 일부 실시형태에서, 방법은 본 발명의 폴리뉴클레오타이드에 의해 암호화되는 단백질의 지속적인 순환 반감기를 산출한다. 예를 들어, 단백질은 단백질 또는 단백질을 암호화하는 mRNA의 피하 주사를 통해 관찰된 반감기보다 더 긴 시간 또는 수일 동안 검출될 수 있다. 일부 실시형태에서, 단백질의 반감기는 1일, 2일, 3일, 4일, 5일, 또는 1주 이상이다.

[1686] 많은 유형의 방출 전달 시스템이 이용 가능하고 당업자에게 공지되어 있다. 이들은 중합체계 시스템, 예컨대, 폴리(락티드-글리콜리드), 코폴리옥살레이트, 폴리카프로락톤, 폴리에스터아마이드, 폴리오쏘에스터, 폴리하이드록시부티르산 및 폴리무수물을 포함한다. 약물을 함유하는 상기 중합체의 마이크로캡슐은 예를 들어, 미국 특허 제5,075,109호에 기재되어 있다. 전달 시스템은 또한 지질, 예를 들어, 스테롤, 예컨대, 콜레스테롤, 콜레스테롤 에스터 및 지방산 또는 중성 지방, 예컨대, 모노-다이- 및 트라이-글리세리드인 비중합체 시스템; 하이드로겔 방출 시스템; 시알스틱(sylastic) 시스템; 펩타이드 기반 시스템: 왁스 코팅; 통상적인 결합제 및 부형제를 사용하는 압축 정제; 부분적으로 융합된 임플란트 등을 포함한다. 특정 예는 (a) 활성 조성물이 미국 특허 제4,452,775호, 제4,667,014호, 제4,748,034호 및 제5,239,660호에 기재된 것과 같은 매트릭스 내에 형태로 함유되어 있는 침식 시스템 및 (b) 활성 성분이 미국 특허 제3,832,253호 및 제3,854,480호에 기술된 바와 같이 중합체로부터 제어된 속도로 침투하는 확산 시스템을 포함하지만 이들로 제한되지 않는다. 또한, 펌프 기반 하드웨어 전달 시스템을 사용할 수 있고, 그 중 일부는 이식에 적합하다.

[1687] 일부 실시형태에서, 치료제는 표적화 모이어티에 연결 모이어티를 통해 직접적으로 또는 간접적으로 접합될 수 있다. 치료제를 표적화 모이어티에 접합시키는 방법은 당업계에 공지되어 있다. 예를 들어, 문헌[Wadwa *et al.*, J, Drug Targeting 3:111 (1995)] 및 미국 특허 제5,087,616호 참조.

[1688] 일부 실시형태에서, 본 명세서에 제공된 치료제는 치료제가 투여되는 신체 내로 방출되는 방식이 신체 내 시간 및 위치에 대해 제어되도록 저장소 형태로 제형화된다(예를 들어, 미국 특허 제4,450,150호 참조). 치료제의 저장소 형태는 예를 들어, 치료제 및 중합체와 같은 다공성 또는 비다공성 물질을 포함하는 이식 가능한 조성물일 수 있고, 여기서 치료제는 물질 및/또는 비-다공성 물질의 분해에 의해서 캡슐화거나 그 전체에 확산된다. 그 다음 저장소는 신체 내 원하는 위치에 이식되고 치료제는 미리 결정된 속도로 임플란트에서 방출된다.

[1689] **16. 치료 방법**

[1690] 특정 양상에서, 본 명세서에는 병태, 예를 들어, 자가면역 장애 또는 암을 치료 및/또는 예방하는 방법이 제공된다.

[1691] 특정 실시형태에서, 본 명세서에 제공된 치료제는 (예를 들어, 동일한 약제학적 조성물로 또는 별도의 약제학적 조성물로) 하나 이상의 추가 치료제와 공동 투여된다. 일부 실시형태에서, 본 명세서에 제공된 치료제가 먼저 투여될 수 있고 하나 이상의 추가 치료제가 두 번째로 투여될 수 있거나, 그 반대의 경우도 마찬가지이다. 대안적으로, 본 명세서에 제공된 치료제 및 하나 이상의 추가 치료제는 동시에 투여될 수 있다.

[1692] 일부 실시형태에서, 대상체는 포유동물이다. 일부 실시형태에서, 본 명세서에 언급된 포유동물은 설치목의 포유동물, 예컨대, 마우스 및 햄스터 또는 토끼와 같은 토끼목의 포유동물, 예컨대, 토끼를 포함하는 임의의 포유동물일 수 있다. 포유동물은 고양이과(고양이)와 개과(개)를 포함하여 육식 동물 목에서 유래할 수 있다. 포유동물은 소류(소) 및 돼지류(돼지)를 포함하는 우제류목 또는 말류(말)을 포함하는 말목에 속할 수 있다. 포유동물은 영장류, 세보이드(Ceboids) 또는 시모이드(Simoids(원숭이)) 또는 진원류목(order Anthropoids)(인간 및 원숭이)일 수 있다. 바람직하게는, 포유동물은 인간이다.

[1693] **17. 서열**

표 17

IRES 서열.

서열번호	IRES	서열
1	EMCV-A	ccccccctctccctccccccctaacgttactggccgaagccgcttggaataaggccggtgtgcgttt gtctatatgttatttttccaccatattgccgtcttttggcaatgtgagggcccggaaacctggccctgtct tcttgacgagcattcctaggggtctttcccctctcgccaaaggaatgcaaggtctgttgaatgtcgtg aaggaagcagttcctctggaagcttcttgaagacaaacaacgtctgtagcgacccctttgcaggcag cggaaccccccacctggcgacaggtgcctctgcggccaaaagccacgtgtataagatacacctg caaaggcggcacaaccccagtgccacgttgtgagttggatagttgtggaaagagtcaaatggctc tcctcaagcgtattcaacaaggggctgaaggatgcccagaaggtaccccattgtatgggatctgat ctggggcctcggtgcacatgctttacatgtgtttagtcgaggttaaaaaacgtctaggcccccccgaa ccacggggacgtggttttcctttgaaaaacacgatgataatatggccacaacc
2	EMCV-B	ctccccctcccccccctttactatactggccgaagccacttggaataaggccggtgtgcgtttgtcta catgctattttctaccgcattaccgtcttatggtaatgtgagggtccagaacctgaccctgtcttcttga cgaacactcctaggggtctttcccctctcgacaaaggagtgtaaggtctgttgaatgtcgtgaagga agcagttcctctggaagcttcttaaagacaaacaacgtctgtagcgacccctttgcaggcagcggaa ccccccacctggtgacaggtgcctctgcggccaaaagccacgtgtataagatacacctgcaaagg gcggcacaaccccagtgccacgttgtgagttggatagttgtggaaagagtcaaatggctctcctca agcgtattcaacaaggggctgaaggatgcccagaaggtaccccattgtatgggatctgatctggg gcctcggtgcacgtgctttacacgtgttgagtcgaggtgaaaaaacgtctaggcccccgaacca cggggacgtggttttcctttgaaaaccacgattacaat
3	EMCV-Bf	ttgccagtctgctcgatatcgcaggctgggtccgtgactaccccactcccccctttcaacgtgaaggct acgatagtgccagggcgggtactgccgtaagtgccaccccaaacaacaacaacaaaacaaactc ccctccccccccctttactatactggccgaagccacttggaataaggccggtgtgcgtttgtctacat gctattttctaccgcattaccgtcttatggtaatgtgagggtccagaacctgaccctgtcttcttgacg aacactcctaggggtctttcccctctcgacaaaggagtgtaaggtctgttgaatgtcgtgaaggaag cagttcctctggaagcttcttaaagacaaacaacgtctgtagcgacccctttgcaggcagcggaacc ccccacctggtgacaggtgcctctgcggccaaaagccacgtgtataagatacacctgcaaaggc ggcacaaccccagtgccacgttgtgagttggatagttgtggaaagagtcaaatggctctcctcaag cgtattcaacaaggggctgaaggatgcccagaaggtaccccattgtatgggatctgatctggggc ctcggtgcacgtgctttacacgtgttgagtcgaggtgaaaaaacgtctaggcccccgaaccacg gggacgtggttttcctttgaaaaccacgattacaat
4	EMCV-Cf	ttgccagtctgctcgatatcgcaggctgggtccgtgactaccccactcccccctttcaacgtgaaggct acgatagtgccagggcgggtactgccgtaagtgccaccccaaaacaacaacaaccccccctctc cctccTccccccctaacgttactggccgaagccgcttggaataaggccggtgtgcgtttgtctatat gttatttttccaccatattgccgtcttttggcaatgtgagggcccggaaacctggccctgtcttcttgac gagcattcctaggggtctttcccctctcgccaaaggaatgcaaggtctgttgaatgtcgtgaaggaa gcagttcctctggaagcttcttgaagacaaacaacgtctgtagcgacccctttgcaggcagcggaac ccccccacctggcgacaggtgcctctgcggccaaaagccacgtgtataagatacacctgcaaagg cggcacaaccccagtgccacgttgtgagttggatagttgtggaaagagtcaaatggctctcctcaa gcgtattcaacaaggggctgaaggatgcccagaaggtaccccattgtatgggatctgatctgggg cctcggtgcacatgctttacatgtgtttagtcgaggttaaaaaacgtctaggcccccgaaccacgg ggacgtggttttcctttgaaaaacacgatgataat
5	EMCV pEC9	ccccccccctaacgttactggccgaagccgcttggaataaggccggtgtgcgtttgtctatatgttat tttccaccatattgccgtcttttggcaatgtgagggcccggaaacctggccctgtcttcttgacgagc attcctaggggtctttcccctctcgccaaaggaatgcaaggtctgttgaatgtcgtgaaggaagcag

[1694]

		ttcctctggaagcttcttgaagacaaacaacgtctgtagcgacccctttgcaggcagcggaaccccc cacctggcgacaggtgcctctgcggccaaaagccacgtgtataagatacacctgcaaaggcggc acaaccccagtgccacgttgtgagttggatagttgtggaaagagtcaaatggctctcctcaagcgt attcaacaaggggctgaaggatgcccagaaggtaccccattgtatgggatctgatctggggcctc ggtgcacatgctttacatgtgtttagtcgaggttaaaaaacgtctaggccccccgaaccacgggga cgtggttttccttttgaaaaacacgatgataat
6	파코비르나바이러스	gtaaattaaatgctatttacaaaatttaaacagaaaggagagatgttatgaaccggttttacaaggttt catacatcgaaaatagcactacctggggcagccgacacactaacatcgtctgtttaaccagaagtg ttactgaaaggaggttattta
7	HCV QC64	acctgcccctaataggggcgacactccgccatgaatcactcccctgtgaggaactactgtcttcac gcagaaagcgtctagccatggcgttagtatgagtgtcgtacagcctccaggcccccccctcccgg gagagccatagtggtctgcggaaccggtgagtacaccggaattgccgggaagactgggtcctttc ttggataaacccactctatgcccggacatttgggcgtgcccccgcaagactgctagccgagtagc gttgggttgcgaaaggccttgtggtactgcctgatagggtgcttgcgagtgcccgggaggtctcg tagaccgtgcatc
8	인간 코사바이러스 E/D	ctacaagctttgtgtaaacaaacttttgtttggcttttctcaagcttctctcacatcaggccccaaagat gtcctgaaggtaccccgtgtatctgaggatgagcaccatcgactaccggacctgcaaaattttgc aaacgcatgtggtatcccagccccctcctctcggggaggggcttgctcactcagcacaggatct gatcaggagatccacctccggtgctttacaccggggcgtggatttaaaaattgcccaaggcctggc gcacaacctaggggactaggttttccttatattttaaagctgtcaat
9	인간 코사바이러스 F	gtcttaggacgacgcatgtggtatcccagcccccgcctacattggcgggggctttgaagcacca gacactggatctgatcaggaggagggtagctgctttacagcccctcttaaaaattgcccaaggtcc ggccacccaacctaggggactaggttttccttttatttttaaattgtcatt
10	인간 코사바이러스 JMY	acatgggggagactgcatgtggcagtcttgaaacgtgtggtttgacgtctaccttatatggcagtgg gtggagtactgcaaagatgtcaccgtgctttacacggtttttgaaccccacaccggctgtttgacgct cgtagggcagcaggtttatttttcattaaaattcttacttttctagctgcatgagttctattcatgcagacgg agtgatactcccgttccttcttggacaggttgcctccacgccctttgtggatcttaaggtgaccaagtc actggtgttggaggtgaagatagagagtcctcttgggaatgtcatgtggctgtgccaggggttgta gcgatgccattcgtgtgtgcggatttcctctcgtggtgacacgagcctcacaggccaaaagccccg tccgaaaggacccgaatggtggagtgaccctgactccccccctgcatagttttgtgattaggaacttg aggaatttctgtcataaatctctatcacatcaggccccaaagatgtcctgaaggtaccctgtgtatctg aggatgagcaccaccgactacccggacttgcattagcagacacatgtggttgcccagccccacct cttcagaggtggggctttgctcactcagcacaggatctgatcaggagccccgctcgtgtgctttaca ctcgacgcggggttaaaaattgcccaaggcctggcacaacaacctaggggactaggttttcctattt ttgtaaattatgtcaat
11	리노바이러스 NAT001	gtgacaatcagccagattgttaacggtcaagcacttctgtttccccggtacccttgtatacgcttcacc cgaggcgaaaagtgaggttatcgttatccgcaaagtgcctacgagaagcctagtagcacttttgaa gcctatggctggtcgctcaactgtttacccagcagtagacctggcagatgaggctagatgttcccc accagcgatggtgatctagcctgcgtggctgcctgcacactctattgagtgtgaagccagaaagtg gacaaggtgtgaagagcctattgtgctcactttgagtcctccggcccctgaatgtggctaatcctaa ccccgtagctgttgcatgtaatccaacatgtctgcagtcgtaatgggcaactatgggatggaaccaa ctactttgggtgtccgtgtttcttgttttttctttatgcttgcttatggtgacaactgtagttattacatttgttacc
12	HRV14	ttaaaacagcggatgggtatcccaccattcgacccattgggtgtagtactctggtactatgtacctttg tacgcctgtttctccccaaccaccctccttaaaattcccacccatgaaacgttagaagcttgacatta aagtacaataggtggcgccatatccaatggtgtctatgtacaagcacttctgtttcccaggagcgag gtataggctgtacccactgccaaaagcctttaaccgttatccgccaaccaactacgtaacagttagt

[1695]

		accatcttgttcttgactggacgttcgatcaggtggatttttccctccactagtttggtcgatgaggcta ggaattccccacgggtgaccgtgtcctagcctgcgtggcggccaacccagcttatgctgggacgc ccttttaaggacatggtgtgaagactcgcatgtgcttggttgtgagtcctccggcccctgaatgcgg ctaaccttaaccctagagccttatgccacgatccagtggttgtaaggtcgtaatgagcaattccggg acgggaccgactactttgggtgtccgtgtttctcatttttcttcatattgtcttatggtcacagcatatata tacatatactgtgatc
13	HRV89	ttaaaactgggagtgggttgttcccactcactccacccatgcggtgttgtactctgttattacggtaac tttgtacgccagttttcccaccctttccccataatgtaacttagaagtttgtacaatatgaccaataggt gacaatcatccagactgtcaaaggtcaagcacttctgtttccccggtcaatgaggatatgctttaccc aaggcaaaaaccttagagatcgttatccccacactgcctacacagagcccagtaccatttttgatat aattgggttggtcgctccctgcaaacccagcagtagacctggcagatgaggctggacattcccca ctggcgacagtggtccagcctgcgtggctgcctgctcacccttcttgggtgagaagcctaattattg acaaggtgtgaagagccgcgtgtgctcagtgtgcttcctccggcccctgaatgtggctaaccttaa ccctgcagccgttgcccataatccaatgggtttgcggtcgtaatgcgtaagtgcgggatgggacca actactttgggtgtccgtgtttccttgttttctttgattgcattttatggtgacaatttatagtgtatagattg tcatc
14	HRVC-02	ttaaaactgggtacaggttgttcccacctgtatcacccacgtggtgtggtgctcttgtattccggtaca cttgcacgccagtttgccacccctcacccgtcgtaacttagaagctaacaactcgaccaacaggcg gtggtaaaccataccacttacggtcaagcactcctgtttcccggtatgcgaggaatagactcctac agggttgaagcctcaagtatcgttatccgcattggtactacgcaaagcttagtagtgccttgaaagtc ccttggttggtcgctccgctagtttcccctagtagacctggcagatgaggcaggacactccccact ggcgacagtggtcctgcctgcgtggctgcctgcgcacccttagggtgcgaagccaagtgacag acaaggtgtgaagagccccgtgtgctaccaatgagtcctccggcccctgaatgcggctaatccaa ccccacagctattgcacacaagccagtgtgtatgtagtcgtaatgagcaattgtgggacggaaccg actactttgggtgtccgtgtttccttttattcttatcattctgcttatggtgacaatactgtgaaatagtgtt gttacc
15	HRV-A21	taaaactggatccaggttgttcccacctggatctcctattgggagttgtactctattattccggtaatttt gtacgccagttttatcttcccctccccaattgtaacttagaaggttatcaatacgaccaataggtggt agttagccaaactaccaaaggtcaagcacttctgtttccccggtcaaagttgatatgctccaacagg gcaaaaacaactgagatcgttatccgcaaagtgcctacgcaaagcctagtaacacctttgaagattt atggttggtcgttccgctatttcccatagtagacctggcagatgaggctagaaatcccccactggcg acagtgctctagcctgcgtggctgcctgcgcacccctttgggtgcgaagccatacattggacaagg tgtgaagagccccgtgtgctcactttgagtcctccggcccctgaatgtggctaaccttaaccctgca gctagtgcatgtaatccaacatgttgctagtcgtaatgagtaattgcgggacgggaccaactactttg ggtgtccgtgtttcacttttttcctttttaatattgcttatggtgacaatatatatagctatatatattgacacc
16	살리바이러스 ASH1	ttcccctgcaaccattacgcttactcgcatgtgcattgagtggtgcatgtgttgaacaaacagctaca ctcacatgggggcgggttttcccgccctacggcttctcgcgaggcccaccctcccctttctcccat aactacagtgctttggtaggtaagcatcctgatcccccgcggaagctgctcacgtggcaactgtgg ggacccagacaggttatcaaaggcacccggtctttccgccttcaggagtatccctgctagcgaatt ctagtagggctctgcttggtgccaacctcccccaaatgcgcgctgcgggagtgctcttccccaact caccctagtatcctctcatgtgtgtgcttggtcagcatatctgagacgatgttccgctgtcccagacc agtccagtaatggacgggccagtgtgcgtagtcgtcttccggcttgtccggcgcatgttttggtgaac cggtggggtaaggttggtgtgcccaacgcccgtactcaggggatacctcaaggcacccaggaat gccagggaggtaccccgcttcacagcgggatctgaccctggggtaaatgtctgcgggggtcttc ttggcccacttctcagtacttttcagg
17	살리바이러스 FHB	acatgggggtctgcggacggcttcggcccacccgcgacaagaatgccgtcatctgtcctcatta cccgtattccttcccttccccccgcaaccaccacgcttactcgcgcacgtgttgagtggcacgtgcgt

[1696]

		tgtccaaacagctacacccacacccttcggggcgggtttgtcccgccctcgggttcctcgcggaa ccccccccctccctctctctttctatccgccctcacttcccataactacagtgctttggtaggtgagc accctgaccccccgcggaagctgctaacgtggcaactgtggggatccaggcaggtttatcaaagg cacccggtctttccgccttcaggagtatctctgccggtgaattccggtagggctctgcttggtgcca acctccccccaaatgcgcgctgcgggagtgctcttccccaactcatcttagtaacctctcatgtgtgtg cttggtcagcatatctgaggcgacgttccgctgtcccagaccagtccagcaatggacgggccagt gtgcgtagtcgctttccggttttccggcgcatgtttggcgaaacgctgaggtaaggttggtgtgccc aacgcccgtaatttggtgatacctcaagaccacccaggaatgccagggaggtaccccacttcggt gggatctgaccctgggctaattgtctacggtggttcttcttgcttccacttctcttttttctggcatg
18	살리바이러스 NG-J1	tatggcaggcgggcttgtggacggcttcggcccacccacagcaagaatgccatcatctgtcctca ccccccaattttccctttttcttcccctgcaaccattacgcttactcgcatgtgcattgagtggtgcatgtgt tgaacaaacagctacactcacatgggggcgggttttcccgccctacggcctctcgcgaggcccac ccctttcccctcccctataactacagtgctttggtaggtaagcatcctgatcccccgcggaagctgctc acgtggcaactgtggggacccagacaggttatcaaaggcacccggtctttccgccttcaggagtat ccctactagtgaattctagcggggctctgcttggtgccaacctcccccaaatgcgcgctgcgggag tgctcttcccaactcaccctagtatcctctcatgtgtgtgcttggtcagcatatctgagacgatgttcc gctgtcccagaccagtccagtaatggacgggccagtgcgtgtagtcgtcttccggcttgtccggg gcatgtttggtgaaccggtggggtaaggttggtgtgcccaacgcccgtactttggtgacacctcaa gaccacccaggaatgccagggaggtaccccacctcacggtgggatctgaccctgggctaattgt ctacggtggttcttcttgcttccacttctttcttctgttcacg
19	인간 파레코바이러스 1	tttgaaaggggtctcctagagagcttggccgtcgggccttatacccgacttgctgagtttctctagg agagccctttccagccctgaggcggctggtcaataaaagcctcaaacgtaactaacacctaaga agatcatgtaaaccctatgcctggtctccactattcgaaggcaacttgcaataagaagagtgggatc aagacgcttaaagcatagagacagttttctttttctaacccacatttgtgtggggtggcagatggcgtg ccataactctaatagtgagataccacgcttgtggaccttatgctcacacagccatcctctagtaagttt gtgagacgtctggtgacgtgtgggaacttattggaaacaacattttgctgcaaagcatcctactgcc agcggaaaaacacctggtaacaggtgcctctggggccaaaagccaaggtttaacagacccttag gattggttctaaacctgagatgttgtggaagatatttagtacctgctgatctggtagttatgcaaacact agttgtaaggcccatgaaggatgcccagaaggtacccgtaggtaacaagtgacactatggatctg atttggggccagatacctctatcttggtgatctggttaaaaaacatctaatgggccaaacccggggg ggatccccggtttcctcttattctatcaatgccact
20	크로히바이러스 B	gtataagagacaggtgtttgccttgtcttcggactggcatcttgggaccaaccccccttttccccagc catgggttaaatggcaataaaggacgtaacaacttttgtaaccattaagctttgtaattttgtaaccact aagctttgtgcacataatgtaaccatcaagcttgttagtcccagcaggaggtttgcatgcttgtagcc gaaatggggctcgaccccccatagtaggatacttgattttgcattccattgtggacctgcaaactcta cacatagaggctttgtcttgcatctaaacacctgagtacagtgtgtacctagaccctatagtacggga ggaccgtttgtttcctcaataaccctacataataggctaggtgggcatgcccaatttgcaagatccca gactgggggtcggtctgggcagggttagatccctgttagctactgcctgatagggtggtgctcaac catgtgtagtttaaattgagctgttcatatacc
21	Yc-3	actgaagatcctacagtaactactgccccaatgaacgccacagatgggtctgctgatgactacctat cttagtgctagttgaggtttgaagtgagccggttttagaagaaccagtttctgaacattatcatcccc agcatctattctatacgcacaagatagatagtcatcagcagacacatctgtgctactgcttgatagag ttgcggctggtcaacttagattggtataaccagttgagtggcaa
22	로사바이러스 M-7	tatgcatcactggacggcctaacctcggtcgtggcttcttgccgatttcagcgctaccaggctttctg gtctcgccaggcgttgattagtaggtgcactgtctaagtgaagacagcagtgctctctgtgaaaagt tgatgacactcttcaggtttgtagcgatcactcaaggctagcggatttccccgtgtggtaacacacg cctctaggcccagaaggcacggtgttgacagcacccctgagtggctggtcttccccaccagcac

[1697]

		ctgatttgtggattcttcctagtaacggacaagcatggctgctcttaagcattcagtgcgtccggggc tgaaggatgcccagaaggtacccgcaggtaacgataagctcactgtggatctgatctggggctgc gggctgggtgtctttccacccagccaaaacccgtaaaacggtagtcgcagttaaaaaacgtctag gccccacccccccagggatgggggttcccttaaaccctcacaagttcaac
23	산바바이러스 A	tgaaaaggggcgcagggtggtggtggttactaaatacccaccatcgccctgcacttccctttcc cctgtggctcagggtcacttagccccctctttgggttaccagtagttttctacccctgggcacagggt taactatgcaagacggaacaacaatctcttagtcccctcgccgatagtgggctcgaccccatgt gtaggagtggataagggacggagtgagccgatacggggaagagtgtgcggtcacaccttaattc catgagcgctgcgaagaaggaagctgtgaacaatggcgacctgaaccgtacacatggagctcca caggcatggtactcgttagactacgcagcctggttgggagtgggtatacctgggtgagccgcca gtgaatgggagttcactggttaacacacactgcctgatagggtcagggcctcctgtccccgccgta atgaggtagaccatatgcc
24	파시바이러스 A	gcggctggatattctggccgtgcaactgcttttgaccagtggctctgggtaacttagccaaagtgtc cttctccctttccctattatatgttttatggctttgtctggtcttgtttagtttatatataagatcctttccgcc gatatagacctcgacagtctagtgtaggaggattggtgatattaatttgccccagaagagtgaccgt gacacatagaaaccatgagtacatgtgtatccgtggaggatcgcccgggactggattccatatccc attgccatcccaacaagcggagggtatacccactatgtgcacgtctgcagtgggagtctgcagatt tagtcatactgcctgatagggtgtgggcctgcactctggggtactcaggctgtttatataat
25	파시바이러스 A 2	gctggactttctggctgcgcaactgcttttaaccagtggctctgggttacttagccaaaaccccctttc cccgtaccctagtttgtgtgtgtattattattttgttgttgttttgtaaatttttatataagatcctttccgccg atatagacctcgacagtctagtgtaggaggattggtgatattaatatgccccagaagagtgaccgtg acacatagaaaccatgagtacatgtgtatccgtggaggatcgcccgggactggattccatatccca ttgccatcccaacaaacggagggtatacccgctatgtgcgcgtctacagtgggaatctgtagattta gtcatactgcctgatagggtgtgggcctgcactctggggtactcaggctgtttatataat
26	에코바이러스 E14	ttaaaacagcctgtgggttgttcccatccacagggcccactgggcgccagcactctggtattgcgg taccttagtgcgcctgttttatatacccgtccccaaacgtaacttagacgcatgtcaacgaagacca atagtaagcgcagcacaccagctgtgttccggtcaagcacttctgttaccccggaccgagtatcaa taagctactcacgtggctgaaggagaaaacgttcgttacccgaccaattacttcaagaaacctagta acaccatgaaggttgcgcagtgtttcgctccgcacaaccccagtgtagatcaggtcgatgagtcac cgcattcccacgggtgaccgtggcggtggctgcgctggcggcctgcccatggggaaacccat gggacgcttcaatactgacatggtgcgaagagtctattgagctaattggtagtcctccggccccga atgcggctaatcctaactgcggagcagatacccacacaccagtgggcagtctgtcgtaacgggca actctgcagcggaaccgactactttgggtgtccgtgtttctctttatccttatactggctgcttatggtg acaattgagagattgttaccatatagctattggattggccatccggtgacaaatagagcaattgtgtat ttgtttgttggtttcgtgccattaaattacaaggttctaaacacccttaatcttattatagcattcaacaca acaaa
27	인간 파레코바이러스 5	gtacattagatgcgtcatctgcaactttagtcaataaattacctccaatgtcattaccaacattccctac cttttcactaacacctaagacaacaagtacctatgcctggtctccactattcgaaggcaacttgcaat aagaagagtggaattaagacgcttaaagcatagagctagttatctttctaacccacaaagttttgtg gggtggcagatggcgtgccataactctattagtgagataccatgcttgtggatcttatgctcacaca gccatcctctagtaagttgataaggtgtctggtgatatgtgggaactcacatgaaccattaatttaccg taaggtatcctatagccagcggaatcacatctggtgacagatgcctctggggccgaaagccaagg tttaacagaccctataggattggtttcaaaacctgaattgatgtggattgtgtatagtacctgttgatct ggtaacagtgtcaacactagttgtaaggcccacgaaggatgcccagaaggtacccgtaggtaaca agtgacactatggatctgatctggggccagctacctctatcatggtgagttggttaaaaaacgtctag tgggccaaacccaggggggatccctggtttccttttacctaatcaaagccact
28	아이치	tttgaaaaggggtgggggggcctcggccccctcaccctctttccggtggtctggtcccggacc

[1698]

	바이러스	accgttactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatactcccccc acccccctttgtaactaagtatgtgtgctcgtgatcttgactcccacggaacggaccgatccgttgg tgaacaaacagctaggtccacatcctcccttcccctgggagggcccccgccctcccacatcctcc ccccagcctgacgtatcacaggctgtgtgaagcccccgcgaaagctgctcacgtggcaattgtgg gtccccccttcatcaagacaccaggtctttcctcctaaggctagccccggcgtgtgaattcacgttg ggcaactagtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaac ccctggcccttcactatgtgcctggcaagcatatctgagaaggtgttccgctgtggctgccaacctg gtgacaggtgccccagtgtgcgtaaccttcttccgtctccggacggtagtgattggttaagatttggt gtaaggttcatgtgccaacgccctgtgcgggatgaaacctctactgccctaggaatgccaggcag gtaccccacctccgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaatccttt tatgtcggagtc
29	A형 간염 바이러스 HA16	ttcaagaggggtctccggagttttccggaaccctcttggaagtccatggtgaggggacttgatac ctcaccgccgtttgcctaggctataggctaaatttccctttccctgtccttccccctatttccttttgttttgt ttgtaaatattaattcctgcaggttcagggttctttaatctgtttctctataagaacactcaatttttcacgc tttctgtctccttttcttccagggctctcccctgccctaggctctggccgttgcgcccggcggggtca actccatgattagcatggagctgtaggagtctaaattggggacgcagatgtttgggacgtcgccttg cagtgttaacttggctttcatgaacctctttgatcttccacaaggggtaggctacgggtgaaacctctt aggctaatacttcaatgaagagatgccttggatagggtaacagcggcggatattggtgagttgttaa gacaaaaaccattcaacgccggaggactggctctcatccagtggatgcattgagggaattgattgt cagggctgtctctaggtttaatctcagacctctctgtgcttagggcaaacactatttggccttaaatgg gatcctgtgagaggggtccctccattgacagctggactgttctttggggccttatgtggtgtttgcc tctgaggtactcaggggcatttaggtttttcctcattcttaaataata
30	포피바이러스	gggagtaaacctcaccaccgtttgccgtggtttacggctacctatttttggatgtaaatattaattcctg caggttcaggtctcttgaattatgtccacgctagtggcactctcttacccataagtgacgccttagcg gaacctttctacacttgatgtggttaggggttacattatttccctgggccttctttggccctttttccccctg cactatcattctttcttccgggctctcagcatgccaatgttccgaccggtgcgcccgccggggttaa ctccatggttagcatggagctgtaggccctaaaagtgctgacactggaactggactattgaagcat acactgttaactgaaacatgtaactccaatcgatcttctacaaggggtaggctacgggtgaaacccc ttaggttaatactcatattgagagatacttctgataggttaaggttgctggataatggtgagtttaacga caaaaaccattcaacagctgtgggccaacctcatcaggtagatgcttttggagccaagtgcgtagg ggtgtgtgtggaaatgcttcagtggaaggtgccctcccgaaaggtcgtaggggtaatcaggggca gttaggtttccacaattacaatttgaa
31	CVA10	gctcttccgatctgggttgttcccacccacagggcccactgggcgccagcactctgattccacgga atctttgtgcgcctgttttacaacccttcccaatttgtaacgtagaagcaatacacactactgatcaata gtaggcatggcgcgccagtcatgtcatgatcaagcacttctgttcccccggactgagtatcaataga ctgctcacgcggttgaaggagaaaacgttcgttacccggctaactacttcgagaaacctagtagca ccatggaagctgcggagtgtttcgctcagcactttccccgtgtagatcaggtcgatgagtcactgca atccccacgggcgaccgtggcagtggctgcgttggcggcctgcctatggggcaacccataggac gctctaatgtggacatggtgcgaagagtctattgagctagttagtagtcctccggcccctgaatgcg gctaatcctaactgcggagcacatgccttcaacccaggaggtggtgtgtcgtaacgggtaactctg cagcggaaccgactactttgggtgtccgtgtttccttttatccttatattggctgcttatggtgacaatc acggaattgttgccatatagctattggattggccatccggtgtctaacagagctattgtatacctatttg ttggatttactcccctatacataaatctctgaacactttgtgctttatactgaacttaaacacacgaaa
32	엔테로바이러스 C	ttaaaacagctctgggggttgttcccacccccagaggcccacgtggcggccagtacaccggtaccac ggtacccttgtacgcctgttttatactccccctccccgtaaactagaagcacgaaacacaagttcaata gaaggggtacagaccagtaccaccacgaacaagcacttctgttcccccggtgaggtcacatag actgtccccacggtcaaaagtgactgatccgttatccgctcacgtacttcggaaagcctagtaccac

[1699]

		cttggaatctacgatgcgttgcgctcagcactcgaccccggagtgtagcttaggctgatgagtctg gacgttccccactggtgacagtggtccaggctgcgttggcggcctacctgtggtccaaaaccaca ggacgctagtagtgaacaaggtgtgaagagcccactgagctacctgagaatcctccggcccctg aatgcggctaatcccaaccacggagcaggtaatcgcaaaccagcggtcagcctgtcgtaacgcg taagtctgtggcggaaccgactactttgggtgtccgtgtttccttttatttttatggtggctgcttatggt gacaatcatagattgttatcataaagcaaattggattggccatccggagtgagctaaactatctatttc tctgagtgttggattcgtttcacccacattctgaacaatcagcctcattagtgttaccctgttaataaga cgatatcatcacg
33	엔테로바이러스 D	ttaaaacagctctggggttgttcccaccccagaggcccacgtggcggctagtactccggtaccccc ggtacccttgtacgcctgttttatactccctttcccaagtaactttagaagaaataaactaatgttcaac aggaggggtacaaaccagtaccaccacgaacacacacttctgtttccccggtgaagttgcatag actgtacccacggttgaaagcgatgaatccgttacccgcttaggtacttcgagaagcctagtatcat cttggaatcttcgatgcgttgcgatcagcactctacccccgagtgtagcttgggtcgatgagtctgga caccccacaccggcgacgtggtccaggctgcgttggcggcctacccatggctagcaccatggga cgctagttgtgaacaaggtgcgaagagcctattgagctacctgagagtcctccggcccctgaatgc ggctaatcccaaccacggagcaaatgctcacaatccagtgagtggtttgtcgtaatgcgcaagtct gtggcggaaccgactactttgggtgtccgtgtttccttttatttttattatggctgcttatggtgacaatct gagattgttatcatatagctattggattagccatccggtgatatcttgaaattttgccataacttttcaca aatcctacaacattacactacactttctcttgaataattgagacaactcata
34	엔테로바이러스 J	ttaaaatagcctcagggttgttcccaccctgagggcccacgtggtgtagtactctggtattacggtac ctttgtacgcctattttataccccttccccaagtaatttagaagcaagcacaaaccagttcagtagta agcagtacaatccagtactgtaatgaacaagtacttctgttaccccggaagggtctatcggtaagct gtacccacggctgaagaatgacctaccgttaaccggctacctacttcgagaagcctagtaatgccg ttgaagttttattgacgttacgctcagcacactaccccgtgtgtagttttggctgatgagtcacggcac tccccacgggcgaccgtggccgtggctgcgttggcggccaaccaaggagtgcaagctccttgga cgtcatattacagacatggtgtgaagagcctattgagctaggtggtagtcctccggcccctgaatgc ggctaatcctaactccggagcatatcggtgcgaaccagcacttggtgtgttgtaatacgtaagtctg gagcggaaccgactactttgggtgtccgtgtttcctgttttaacttttatggctgcttatggtgacaattt aacattgttaccatatagctgttgggttggccatccggattttgttataaaaccatttcctcgtgccttga cctttaacacatttgtgaacttctttaaatccctttattagtccttaaatactaaga
35	인간 페기바이러스 2	aactgttgttgtagcaatgcgcatattgctacttcggtacgcctaattggtaggcgcccggccgacc ggcccgcaagggcctagtaggacgtgtgacaatgccatgagggatcatgacactggggtgag cggaggcagcaccgaagtcgggtgaactcgactcccagtgcgaccacctggcttggtcgttcatg gagggcatgcccacgggaacgctgatcgtgcaaagggatgggtccctgcactggtgccatgcg cggcaccactccgtacagcctgatagggtggcggcgggccccccccagtgtgacgtccgtggag cgcaac
36	GBV-C GT110	tgacgtgggggggttgatTTTcccccccggcactgggtgcaagccccagaaaccgacgcct atctaagtagacgcaatgactcggcgccgactcggcgaccggccaaaaggtggtggatgggtga tgacagggttggtaggtcgtaaatcccggtcatcctggtagccactataggtgggtcttaagagaa ggtcaagattcctcttacgcctgcggcgagaccgcgcacggtccacaggtgttggccctaccggt gtgaataagggcccgacatcaggc
37	GBV-C K1737	gacgtgggggggttgatccccccccTTTggcactgggtgcaagccccagaaaccgacgccta tttaaacagacgttaagaaccggcgccgacccggcgaccggccaaaaggtggtggatgggtgat gccagggttggtaggtcgtaaatcccggtcatcttggtagccactataggtgggtcttaagggttgg ttaaggtccctctggcgcttgtggcgagaaagcgcacggtccacaggtgttggccctaccggtgt gaataagggcccgacgtcaggctcgtcgttaaaccgagcccactacccacctgggcaaacaacg cccacgtacggtccacgtcgcccttcaatgtctctcttgaccaataggcttagccggcgagttgaca

[1700]

		aggaccagtgggggctgggcggtaggggaaggacccctgccgctgcccttcccggtggagtg ggaaatgc
38	GBV-C 아이오와	tgacgtgggggggttgatccGccccccccggcactCggtgcaagccccataaaccgacgccta tctaagtagacgcaatgactcggcgccgactcggcgaccggccaaaaggtggtggatgggtggt gacagggttggtaggtcgtaaatcccggtcatcctggtagccactataggtgggtcttaagagaag gtcaagactcctcttgtgcctgcggcgagaccgcgcacggtccacaggtgctggccctaccggtg tgaataagggcccgacgtcaggctcgtcgttaaaccgagcccgtcacccacctgggcaaacgac gcccacgtacggtccacgtcgcccttca
39	페기바이러스 A 1220	tgtagcaatgcgcatattgctacttcggtacgcctaattggtaggcgcccggccgaccggccccgc aagggcctagtaggacgtgtgacaatgccatgcgggatcatgacactggggtgagcggaggca gcaccgaagtcgggtgaactcgactcccagtgcgaccacctggcttggtcgttcatggagggcat gcccacgggaacgctgatcgtgcaaagggatgggtccctgcactggtgccatgcgcggcacca ctccgtacagcctgataggtggcggcgggccccccccagtgtgacgtccgtggagcgcaac
40	파시바이러스 A 3	attttctggccgtgtagctgcttttgaccagtggctctgggttacttagccaaatccccctttccttcacc cttttaaatttgatggtctgtgttgtttgttttgtcttgtctaaataatatataagatccttcccgccgataca gacctcgacagtctggtgtaggagggttggtgttattaatttgccccagaagagtgaccgtgacac atagaaaccatgagtacatgtgtatccgtggaggatcgcccgggactggattccatatcccattgcc atcccaacaagcggagggtatacccactatgtgcgcgtttgcagtgggaatctgcaaatttagtcat actgcctgatagggtgtgggcctgcactctggggtactcaggctgttcatataat
41	사펠로바이러스	ccctctcaccctaaggtggttgtatcccacataccccaccctccctccaaagtggacggacaact ggattttgactaacggcaagtctgaatggtatgatttggatacgtttaaacggcagtagcgtggcga gctatggaaaaatcgcaattgtcgatagccatgttagtgacgcgcttcggcgtgctcctttggtgatt cggcgactggttacaggagagtaggcagtgagctatgggcaaacctctacagtattacttagagg gaatgtgcaattgagacttgacgagcgtctctttgagatgtggcgcatgctcttggcattaccatagt gagcttccaggttgggaaacctggactgggcctatactacctgatagggtcgcggctggccgcct gtaactagtatagtcagttgaaacccccccc
42	로사바이러스 B	gtctctttagtgtctatgcttcagagagcggtgaactgacaccgttgcttcttgcacagcccttcgtgc cggtctttccggttctcgacagcgttgggcatcatggctagttaggctaagatagtggatgatctagt gaacagttttggattgtttggagttttgtagcgatgctagtagtgtgtgtggacctccccacgtggtaa cacgtgccccacaggccaaaagccaaggtgttgaaagcacccctactagtcccagactcacccat ctgggaactcctctcatgaaaaatcttagtaacttttgattcggctattcatcaacctctctagtcaagg gctgaaggatgcccggaaggtacccgcaggtaacgataagctcactgtggatctgatccggggc tttggtgcgaccgtctgtccggcgtagccagagttaaaaaacgtctaggcccttccacccccaaggg attgggggtttccccaatcatttgaaagttcact
43	바룬사 바이러스	ttttgaacgccacctcggagcgatatccggggacccccctcccctttttccttcctaccttcttcccaaa tttccctcttcccttgttattttggtttggatttcctggacatgactcggacggatctatctcatttgctttgt gtctgctccaccagtggcatggtcgaaagatcatcaacactggacgtgtactgtaatggccaaacg tgcccacaggggaaaccatgccggtcgctgtagcggcgggtggacgtggtggacccctctccct gctcataaacttgggtaggtgaagggttcaagcgacgcttgccgtgagggcgcatccggatggt gggaaccaacaaactaggctgtaatggccgacctcaggtggatgagctagggctgctgcaccaa aagggactcgattcgatatcccggcctggtagcctagtgcagtggactcgtagttgggaatctacg actggcctagtacagggtgatagccccgtttcccacgcccacctgttgtagggacacccccccc
44	트레모바이러스 A	tttgaaagaggcctccggagtgtccggaggctctctttcgacccaacccatactgggggtgtgtg ggaccgtacctggagtgcacggtatatatgcattcccgcatggcaagggcgtgctaccttgcccct tgacgcatggtatgcgtcatcatttgccttggttaagccccatagaaacgaggcgtcacgtgccga aaatccctttgcgtttcacagaaccatcctaaccatgggtgtagtatgggaatcgtgtatggggatga ttaggatctctcgtagagggataggtgtgccattcaaatccagggagtactctggctctgacattgg

[1701]

		gacatttgatgtaaccggacctggttcagtatccgggttgtcctgtattgttacggtgtatccgtcttgg cacactgaaagggtatttttgggtaatcctttcctactgcctgatagggtggcgtgcccggccacga gagattaagggtagcaatttaaac
45	돼지 파시바이러스 1	gcttttgaccagtggctctgggttacttagccaagtccctttctcttattttcactagtttatgttgtgtgtt gtctgttttgttttgtttaaattgtatacaagatccttcccgccgacacagacctcgacagtctggtgta ggagggttggtgatattaatttgccccaaaagagtgaccgtgatacgtggaaaccatgagtacatgt gtatccgtggaggatcgcccgggactggattccatatcccattgccatcccaacaaacggagggt atacccaccacgtgcgcgtttgcagtgggaatctgcaaatttagtcatactgcctgatagggtgtgg gcctgcactttggggtactcaggctgttcatataat
46	PLV-CHN	acatggggtatgttgtctgtcctgttttgttgaaacaatatataagatcctttccgccgatatagacctc gacagtctagtgtaggaggattggtgatagtaacttgccccagaagagtgaccgtgacacatagа aaccatgagtacatgtgtatccgtggaggatcgcccgggactggattccatatcccattgccatccc aacaaacggagggtataccccactatgtgcgcgtttgcagtgggagcctgcaaatttagtcatactg cctgatagggtgtgggcctgcactctggggtactcaggctgtttatataat
47	파시바이러스 A (더 긴것)	tgaaaaagtggttgtgcagctggattttccggctgtgcaactgcttttgaccagtggctctgggttact tagccaaattcctttcccttatccctattggtttgtgttgtgttgtttgttttgttttgtcttaactatataca agatccttcccgccgatacagacctcgacagtctggtgtaggagggttggtgttattaatttgccccа aaagagtgaccgtgacacgtggaaaccatgagtacatgtgtatccgtggaggatcgcccgggact ggattccatatcccattgccatcccaacaaacggagggtataccccaccacgtgcgcgtttgcagtg ggaatctgcaaatttagtcatactgcctgatagggtgtgggcctgcactttggggtactcaggctgtt tatataat
48	시시니바이러스	gtgtcattaaggtgtgtttggaagttcgaattagctggtttgtggtgattagtagacccccctggaggta cccaattcggatctgaccagggacccgtgactataccgctccggtaattcgggtttaaaacaatga acgtcaccacacaattacttttctcattttattttcatcattgtcttcctatttaccgattacactcgatttcct tggatgttcctggagatttccctggttacctggaccctcattattgttgttgtttcacccagcgagctgt cccaattgcttattatttgcgcttacaacttcgtcctaatatttttctggttgatcgggttgattgagctcc cgggctatcctgccattcaac
49	헤파시바이러스 K	gggaacaatggtccgtccgcggaacgactctagccatgagtctagtacgagtgcgtgccacccat tagcacaaaaaccactgactgagccacacccctcccggaatcctgagtacaggacattcgctcgg acgacgcatgagcctccatgccgagaaaattgggtatacccacgggtaaggggtggccacccag cgggaatctgggggctggtcactgactatggtacagcctgatagggtgctgccgcagcgtcagtg gtatgcggctgttcatggaac
50	헤파시바이러스 A	acctccgtgctaggcacggtgcgttgtcagcgttttgcgcttgcatgcgctacacgcgtcgtccaac gcggagggaacttcacatcaccatgtgtcactcccctatggagggttccaccccgcttacacgga aatgggttaaccatacccaaagtacgggtatgcgggtcctcctagggccccccggcaggtcga gggagctggaattcgtgaattcgtgagtacacgaaaatcgcggcttgaacgtctttgaccttcgga gccgaaatttgggcgtgccccacgaaggaaggcgggggcggtgttgggccgccgccccctttat cccacggtctgataggatgcttgcgagggcacctgccggtctcgtagaccataggac
51	BVDV1	gtatacgagaatttgcctaggacctcgtttacaatatgggcaatctaaaattataattaggcctaagg gacaaatcctcctcagcgaaggccgaaaagaggctagccatgcccttagtaggactagcaaaata agggggtagcaacagtggtgagttcgttggatggctgaagccctgagtacagggtagtcgtcag tggttcgacgcttcggaggacaagcctcgagataccacgtggacgagggcatgcccacagcaca tcttaacctggacgggggtcgttcaggtgaaaacggtttaaccaaccgctacgaatacagcctgat agggtgctgcagaggcccactgtattgctactgaaaatctctgctgtacatggcac
52	보더병 바이러스	gtatacgggagtagctcatgcccgtatacaaaattggatattccaaaactcgattgggttagggagc cctcctagcgacggccgaaccgtgttaaccatacacgtagtaggactagcagacgggaggacta gccatcgtggtgagatccctgagcagtctaaatcctgagtacaggatagtcgtcagtagttcaacg

[1702]

		caggcacggttctgccttgagatgctacgtggacgagggcatgcccaagacttgctttaatctcgg cgggggtcgccgaggtgaaaacacctaacggtgttggggttacagcctgatagggtgctgcaga ggcccacgaataggctagtataaaaatctctgctgtacatggcac
53	BVDV2	gtatacgagattagctaaagtactcgtatatggattggacgtcaacaaatttttaattggcaacgtagg gaaccttcccctcagcgaaggccgaaaagaggctagccatgccctttagtaggactagcaaaagt agggggactagcggtagcagtgagttcgttggatggccgaacccctgagtacaggggagtcgtc aatggttcgacactccattagtcgaggagtctcgagatgccatgtggacgagggcatgcccacgg cacatcttaacccatgcgggggttgcatgggtgaaagcgctaatcgtggcgttatggacacagcct gatagggtgtagcagagacctgctattccgctagtaaaaaactctgctgtacatggcac
54	CSFV-PK15C	gtatacgaggttagttcattctcgtatgcattattggacaaatcaaaatttcaatttggttcagggcctc cctccagcgacggccgaactgggctagccatgcccatagtaggactagcaaacggagggacta gccgtagtggcgagctccctgggtgttctaagtcctgagtacaggacagtcgtcagtagttcgacg tgagcagaagcccacctcgagatgctatgtggacgagggcatgcccaagacgcaccttaaccct agcgggggtcgctagggtgaaatcacaccacgtgatgggagtccgacctgatagggtgctgcag aggctcactattaggctagtataaaaatctctgctgtacatggcac
55	SF573 디시스트로바이러스	aaaaccgaccccagagatcagaaagtcgttgacgcgatcttttattagaggacgttgcgctggcgc gagctttaattagcagacgccaaaaataaacaacaaaatgctgatcgcgagacttaattgtcagac gattggccaaatccgatgtgatctttgctgctcccagattgccgaaataggagtagtag
56	후베이 피코르나-유사 바이러스	ccccaaaaccccccccttaaactcaacactgtagtggattcattttccgttgcaaaacaaaacattac tacccgcatttatgtaggctctgtgttttctatgcgaccgttacattaatctctactctgacccactagttt ataaaaccgaagacctgaatgaaacgattttccttctttttcaacctctaacgaacctctgacggcttga gaaacctgaagttagtaattatgtttaaaagaaaggaaagtcaaacgcgatgactcttacatccctat tccataccgttgctccacaatgtgagcgatgcgaggtcgggactgcagtattaggggaacgagct acatggagagttaattatctctcccctcctacgggagtctcatgtgagctgtagaaagcggttggca cctctcgttacctcgcctgtacatgatcc
57	CRPV	aaaagcaaaaatgtgatcttgcttgtaaatacaatttttgagaggttaataaattacaagtagtgctatttt tgtatttaggttagctatttagctttacgttccaggatgcctagtggcagccccacaatatccaggaag ccctctctgcggtttttcagattaggtagtcgaaaaacctaagaaatttacct
58	살리바이러스 A BN5	tttcctcctttcgaccgccttacggcaggcgggtccgcggacggcttcggcctacccgcgacaag aatgccgtcatctgtccttatcacccatattctttcccttcccccgcaaccatcacgcttactcgcgca cgtgttgagtggcacgtgcgttgtccaaacagttacactcacaccctttggggcgggtttgtcccgcc ctcgggttcctcgcggaaccctccctcttctctctccctttctatccgccttcactttccataactacagt gctttggtaggtaagcatcctgacccccgcggaagctgccaacgtggcaactgtggggatccag gcaggttatcaaggcacccggtctttccgccttcaggagtatccctgccggtgaattccgacagg gctctgcttggtgccaacctcccccaaatgcgcgctgcgggagtgctcttcccccaactcatcttagt aacctctcatgtgtgtgcttggtcagcatatctgaggcgacgttccgctgtcccagaccagtccagc aatggacgggccagtgtgcgtagtcgctttccggtttcccggcgcatgtttggcgaaacgctgagg taaggttggtgtgcccaatgcccgtaatttggtgacacctcaagaccacccaggaatgccaggga ggtaccccacttcggtgggatctgaccctgggctaattgtctacggtggttcttcttgcttccacttctc ttttttctggcatg
59	살리바이러스 A BN2	tatggcaggcgggcttgtggacggcttcggcccacccacagcaagaatgccatcatctgtcctca ccccccatgtttcccctttctttccctgcaaccgttacgcttactcgcaggtgcatttgagtggtgcacgt gttgaataaacagctacactcacatgggggcgggttttcccgccctgcggcctctcgcgaggccc acccctcccctttcctcccataactacagtgctttggtaggtaagcatcctgatcccccgcggaagct gctcacgtggcaactgtggggacccagacaggttatcaaaggcacccggtctttccgccttcagg agtatccctgctagtgaattctagtagggctctgcttggtgccaacctcccccaaatgcgcgctgcg ggagtgctcttcccccaactcaccctagtatcctctcatgtgtgtgcttggtcagcatatctgagacgat

[1703]

		gttccgctgtcccagaccagtccagtaatggacgggccagtgtgcgtagtcgtcttccggcttttcc ggcgcatgtttggtgaaccggtggggtaaggttggtgtgcccaacgcccgtactttggtgatacct caagaccacccaggaatgccagggaggtaccccgcttcacagcgggatctgaccctgggctaat tgtctacggtggttcttcttgcttccacttctttctactgttc
60	살리바이러스 A 02394	tttcgaccgccttatggcaggcgggcttgtggacggcttcggcccacccacagcaagaatgccat catctgtcctcacccccatttctcccctccttccccctgcaaccattacgcttactcgcatgtgcattgag tggtgcacgtgttgaacaaacagctacactcacgtgggggcgggttttcccgcccttcggcctctc gcgaggcccacccttcccctttcctcccataactacagtgctttggtaggtaagcatcctgatccccc gcggaagctgctcgcgtggcaactgtggggacccagacaggttatcaaaggcacccggtctttcc gcctccaggagtatccctgctagtgaattctagtggggctctgcttggtgccaacctcccccaaatg cgcgctgcgggagtgctcttccccaactcaccctagtatcctctcatgtgtgtgcttggtcagcatat ctgagacgatgttccgctgtcccagaccagtccagcaatggacgggccagtgtgcgtagtcgtctt ccggcttgtccggcgcatgtttggtgaaccggtggggtaaggttggtgtgcccaacgcccgtactt tggtgacaactcaagaccacccaggaatgccagggaggtaccccgcctcacggcgggatctga ccctgggctaattgtctacggtggttcttcttgcttccatttctttcttctgttc
61	살리바이러스 A GUT	tatggcaggcgggcttgtggacggtttcggcccacccacagcaagaatgccatcatctgtcctcac ccccaattttcccttttcttcccctgcaatcatcacgcttactcgcatgtgcattgagtggtgcatgtgtt gaacaaacagctacactcacatgggggcgggttttcccgccctacggcctctcgcgaggcccac ccttcccctcccctttataactacagtgctttggcaggtaagcatcctgatcccccgcggaagctgct cacgtggcaactgtggggacccagacaggttatcaaaggcacccggtctttccgccttcaggagc atccccactagtgaattctagtggggctctgcttggtgccaacctcccccaaatgcgcgctgcggg agtgctcttcccaacccatcctagtatcctctcatgtgtgtgcttggtcagcatatctgagacgacgt tccgctgtcccagaccagtccagtaatggacgggccagtgtgcgtagtcgtcttccggcttgtccg gcgcatgtttggtgaaccggtggggtaaggttggtgtgcccaacgcccgtactttggtgacacctc aagaccacccaggaatgccagggaggtaccccgcctcacggcgggatctgaccctgggctaatt gtctacggtggttcttcttgcttccacttctttctt
62	살리바이러스 A CH	ttctcctgcaaccattacgcttaatcgcatgtgcattgagtggtgcatgtgttgaacaaacagctaca atcacatgggggcgggttttcccgccccacggcttctcgcgaggcccatcctccttttctcccat aactacagtgctttggtaggtaagcatcccgatctcccgcggaagctgctcacgtggcaactgtgg ggacccagacaggttatcaaaggcacccggtctttccgccttcaggagtatccctgctagcgaatt ctagtagggctctgcttggtgccaacctctcccaaatgcgcgctgcgggagtgctcttccccaaatc accccagtatcctctcatgtgtgtgcctggtcagcatatctgagacgatgttccgctgtcccagacca gtccagtaatggacgggccagtgtgcgtagtcgtcctccggcttgtccggcgcatgtttggtgaac cggtggggtaaggttggtgtgcccaacgcccgtaatcaggggatacctcaaggcacccaggaat gccagggaggtatcccgcctcacagcgggatctgaccctggggtaaatgtctgcgggggtcct cttggcccaattctcagtaattttcagg
63	살리바이러스 A SZ1	tctgtcctcaccccatcttcccttctttcctgcaccgttacgcttactcgcatgtgcattgagtggtgca cgtgcttgaacaaacagctacactcacatgggggcgggttttcccgccctgcggcctctcgcgag gcccaccctcccctcctcccataactacagtgctttggtaggtaagcatcctgatccccgcggga agctgctcacgtggcaactgtggggacccagacaggttatcaaaggcacccggtctttccgccttc aggagtatccctgctagtgaattctagtagggctctgcttggtgccaacctcccccaaatgcgcgct gcgggagtgctcttccccaactcaccctagtatcctctcatgtgtgtgcttggtcagcatatctgaga cgatgttccgctgtcccagaccagtccagtaatggacgggccagtgtgcgtagtcgtcttccggctt gtccggcgcatgtttggtgaaccggtggggtaaggttggtgtgcccaacgcccgtactttggtgat acctcaagaccacccaggaatgccagggaggtaccccgcttcacagcgggatctgaccctggg ctaattgtctacggtggttcttcttgcttccacttctttctactgttcatg
64	살리바이러스	acatgggggtctgcggacggcttcggcccacccgcgacaagaatgccgtcatctgtcctcatta

[1704]

	FHB	cccgtattccttcccttccccgcaaccaccacgcttactcgcgcacgtgttgagtggcacgtgcgt tgtccaaacagctacacccacacccttcggggcgggtttgtcccgccctcgggttcctcgcggaa ccccccctccctctctctctttctatccgccctcacttcccataactacagtgctttggtaggtgagc accctgaccccccgcggaagctgctaacgtggcaactgtggggatccaggcaggttatcaaagg cacccggtctttccgccttcaggagtatctctgccggtgaattccggtagggctctgcttggtgcca acctccccaaatgcgcgctgcgggagtgctcttccccaactcatcttagtaacctctcatgtgtgtg cttggtcagcatatctgaggcgacgttccgctgtcccagaccagtccagcaatggacgggccagt gtgcgtagtcgctttccggttttccggcgcatgtttggcgaaacgctgaggtaaggttggtgtgccc aacgcccgtaatttggtgatacctcaagaccacccaggaatgccagggaggtaccccacttcggt gggatctgaccctgggctaattgtctacggtggttcttcttgcttccacttctctttttctggcatg
65	CVB3	ttaaaacagcctgtgggttgatcccacccacaggcccattgggcgctagcactctggtatcacggt accttgtgcgcctgttttatacccctcccccaactgtaacttagaagtaacacacaccgatcaaca gtcagcgtggcacaccagccacgttttgatcaagcacttctgttaccccggactgagtatcaataga ctgctcacgcggttgaaggagaaagcgttcgttatccggccaactacttcgaaaaacctagtaaca ccgtggaagttgcagagtgtttcgctcagcactaccccagtgtagatcaggtcgatgagtcaccgc attccccacgggcgaccgtggcggtggctgcgttggcggcctgcccatggggaaacccatggg acgctctaatacagacatggtgcgaagagtctattgagctagttggtagtcctccggcccctgaatg cggctaatcctaactgcggagcacacaccctcaagccagagggcagtgtgtcgtaacgggcaac tctgcagcggaaccgactactttgggtgtccgtgtttcattttattcctatactggctgcttatggtgac aattgagagatcgttaccatatagctattggattggccatccggtgactaatagagctattatatatcc ctttgttgggtttataccacttagcttgaaagaggttaaaacattacaattcattgttaagttgaatacag caaa
66	CVB1	ttaaaacagcctgtgggttgttcccacccacaggcccattgggcgctagcactctggtatcacggta cctttgtgcgcctgttttacatccccctcccaaattgtaatttagaagtttcacacaccgatcattagca agcgtggcacaccagccatgttttgatcaagcacttctgttaccccggactgagtatcaatagaccg ctaacgcggttgaaggagaaaacgttcgttacccggccaactacttcgaaaaacctagtaacacca tggaagttgcggagtgtttcgctcagcactaccccagtgtagatcaggtcgatgagtcaccgcgttc cccacgggcgaccgtggcggtggctgcgttggcggcctgcctacggggaaacccgtaggacg ctctaatacagacatggtgcgaagagtctattgagctagttggtaatcctccggcccctgaatgcgg ctaatcctaactgcggagcacataccctcaaaccagggggcagtgtgtcgtaacgggcaactctg cagcggaaccgactactttgggtgtccgtgtttcattttattcctatactggctgcttatggtgacaatt gacaggttgttaccatatagttattggattggccatccggtgactaacagagcaattatatatctctttg ttgggtttataccacttagcttgaaagaggttaaaacactacatctcatcattaaactaaatacaacaa a
67	에코바이러스 7	ttaaaacagcctgtgggttgttcccacccacagggcccattgggcgtcagcaccctggtatcacgg taccttgtgcgcctgttttatatcccttccccaattgtaacttagaagaaacacacaccgatcaaca gcaagcgtggcacaccagccatgttttggtcaagcacttctgttaccccggactgagtatcaataga ctgctcacgcggttgaaggagaaagcgtccgttatccggccagctacttcgagaaacctagtaac accatggaagttgcggagtgtttcgctcagcactaccccagtgtagatcaggtcgatgagtcaccg ctttccccacgggcgaccgtggcggtggctgcgttggcggcctgcctatgggggaacccatagg acgctctaatacagacatggtgcgaagagtctattgagctagctggtattcctccggcccctgaatg cggctaatcctaactgtggagcacatgcccctaatccaaggggtagtgtgtcgtaatgagcaattcc gcagcggaaccgactactttgggtgtccgtgtttcctcttattcttgtactggctgcttatggtgacaat tgagagattgttaccatatagctattggattggccatccggtgactaatagagctattgtgtatctctttt gttggatttgtaccacttaatttgaaagaaatcaggacactacgctacattttactattgaacaccgca aa
68	CVB5	ttaaaacagcctgtgggttgtacccacccacagggcccactgggcgctagcactctggtatcacg

[1705]

		gtacctttgtgcgcctgttttatgccccctccccaattgaaacttagaagttacacacaccgatcaa cagcgggcgtggcataccagccgcgtcttgatcaagcactcctgtttccccggaccgagtatcaat agactgctcacgcggttgaaggagaaaacgttcgttacccggctaactacttcgagaaacctagta gcatcatgaaagttgcgaagcgtttcgctcagcacatccccagtgtagatcaggtcgatgagtcac cgcattccccacgggcgaccgtggcggtggctgcgttggcggcctgcctacggggcaacccgt aggacgcttcaatacagacatggtgcgaagagtcgattgagctagttagtagtcctccggcccctg aatccggctaatcctaactgcggagcacataccctcaacccagggggcattgtgtcgtaacgggt aactctgcagcggaaccgactactttgggtgtccgtgtttccttttattcttataatggctgcttatggtg acaattgaaagattgttaccatatagctattggattggccatccggtgtctaacagagctattatatac ctctttgttggatttgtaccacttgatctaaaggaagtcaagacactacaattcatcatacaattgaaca cagcaaa
69	EVA71	ttaaaacagcctgtgggttgcacccactcacagggcccactgggcgcaagcactctggcacttcg gtacctttgtgcgcctgttttatatcccctcccccaatgaaatttagaagcagcaaaccccgatcaata gcaggcataacgctccagttatgtcttgatcaagcacttctgtttccccggactgagtatcaatagac tgctcacgcggttgaaggagaaaacgttcgttatccggctaactacttcggaaagcctagtaacac catggaagttgcggagagtttcgttcagcacttccccagtgtagatcaggtcgatgagtcaccgcat tccccacgggcgaccgtggcggtggctgcgttggcggcctgcccatggggtaacccatgggac gctctaatacggacatggtgtgaagagtctactgagctagttagtagtcctccggcccctgaatgcg gctaatcccaactgcggagcacacgcccacaagccagtgggtagtgtgtcgtaacgggcaactct gcagcggaaccgactactttgggtgtccgtgtttccttttattcttatgttggctgcttatggtgacaatt aaagagttgttaccatatagctattggattggccatccggtgtgcaacagagcgatcgtttacctattt attggttttgtaccattgacactgaagtctgtgatcaccctaattttatcttaaccctcaacacagccaa ac
70	CVA3	ttaaaacagcctgtgggttgtacccacccacagggcccactgggcgctagcacactggtattacg gtacctttgtgcgcctgttttatacccccccccaacctcgaaacttagaagtaaagcaaacccgatca atagcaggtgcggcgcaccagtcgcatcttgatcaagcacttctgtaaccccggaccgagtatcaa tagactgctcacgcggttgaaggagaaaacgttcgttacccggctaactacttcgagaaacccagt agcatcatgaaagttgcagagtgtttcgctcagcactacccccgtgtagatcaggccgatgagtca ccgcacttccccacgggcgaccgtggcggtggctgcgttggcggcctgcctatggggcaaccca taggacgctctaatacggacatggtgcgaagagtctattgagctagttagtagtcctccggcccctg aatgcggctaatcctaactgcggagcacatacccttaatccaaagggcagtgtgtcgtaacgggta actctgcagcggaaccgactactttgggtgtccgtgtttccttttaatttttactggctgcttatggtgac aattgaggaattgttgccatatagctattggattggccatccggtgactaacagagctattgtgttcca atttgttggatttaccccgctcacactcacagtcgtaagaacccttcattacgtgttatttctcaactcaa gaaa
71	CVA12	ttaaaacagcctgtgggttgtacccacccacagggcccactgggcgctagcactctggtactacg gtacctttgtgtgcctgttttaagccccctaccccccactcgtaacttagaaggcttctcacactcgatc aatagtaggtgtggcacgccagtcacaccgtgatcaagcacttctgttaccccggtctgagtacca ataagctgctaacgcggctgaaggggaaaacgatcgttatccggctaactacttcgagaaaccca gtaccaccatgaacgttgcagggtgtttcgctcggcacaaccccagtgtagatcaggtcgatgagt caccgtattccccacgggcgaccgtggcggtggctgcgttggcggcctgcccatggggtgaccc atgggacgctctaatactgacatggtgcgaagagtctattgagctagttagtagtcctccggcccct gaatgcggctaatcctaactgcggagcacataccctaatccaaagggcagtgtgtcgtaacggg caactctgcagcggaaccgactactttgggtgtccgtgtttccttttattcttacattggctgcttatggt gacaattgaaaagttgttaccatatagctattggattggccatccggtgacaaatagagctattgtata tctttttgttggttacgtacccctaattacaaagtggtttcaactttgaaatacatcctaacactaaattg tagaaa

[1706]

72	EV24	ttaaaacagcctgtgggttgcacccacccacagggcccacagggcgctagcactctggtatcacg gtacctttgtgcgcctgttttattacccctcccccaattgaaaattagaagcaatgcacaccgatcaac agcaggcgtggcgcaccagtcacgtctcgatcaagcacttctgtttccccggaccgagtatcaata gactgctcacgcggttgaaggagaaagtgttcgttatccggctaaccacttcgagaaacccagtaa caccatgaaagttgcagggtgtttcgctcagcacttccccagtgtagatcaggtcgatgagtcacc gcgttccccacgggcgaccgtggcggtggctgcgttggcggcctgcctatgggttaacccatag gacgctctaatacagacatggtgcgaagagtttattgagctggttagtatccctccggcccctgaat gcggctaatcctaactgcggagcacgtgcctccaatccaggggttgcatgtcgtaacgggtaac tctgcagcggaaccgactactttgggtgtccgtgtttccttttattcttatactggctgcttatggtgaca atcgaggaattgttaccatatagctattggattggccatccggtgtctaacagagcgattatatacctc tttgttggatttatgcagctcaataccaccaacttaacacattgaaatatatcttaaagttaaacacag caaa
348	AP1.0	attctcgggctacggccctggagccactccggctcctaaagatttagaagtttgagcacacccgcc cactagggccccccatccaggggggcaacgggcaagcacttctgtttccccggtatgatctgata ggctgtaaccacggctgaaacagagattatcgttatccgcttcactacttcgagaagcctagtaatg atgggtgaaattgaatccgttgatccggtgtctcccccacaccagaaactcatgatgagggttgcca tcccggctacggcgacgtagcgggcatccctgcgctggcatgaggcctcttaggaggacggatg atatggatcttgtcgtgaagagcctattgagctagtgtcgactcctccgcccccgtgaatgcggcta atcctaaccccggagcaggtgggtccaatccagggcctggcctgtcgtaatgcgtaagtctggga cggaaccgactactttcgggaaggcgtgtttccatttgttcattatttgtgtgtttatggtgacaactctg ggtaaacgttctattgcgtttattgagagattcccaacaattgaacaaacgagaactacctgttttatta aatttacacagagaagaattaca
349	CK1.0	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttccccctttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttctttttcacacaactctactgctgacaactcactgactatccact tgctctgTcacG
350	PV1.0	aacaaaaggctacaccacttgggctacggcccgcgccaccttgtggcgcaaagacattagaaga atagcatacgcccactagggccctgcagccagcagggtaacgggcaagcacttctgtctcccc ggtagaacggtataggctgtacccacggccgaaaactgaactatcgttacccgactccgtacttcg caaagcttagtaggaaactggaaagttcgagttattgacccggagtgttcccccactccagaaac gcgtgatgagggttgccaccccgaccatggcgacatggtgggcatccctgcgctggcacgcgg cctctaagaggataactcgctcctactggtaaccgaagagccccgtgagctacggtttattcctccg cctccctgaatgcggctaatcctaacccatgagcagttgccatagatccatatggtggactgtcgta acgcgtaagttgtgggcggaaccgactactttgggatggcgtgtttccttgttttctccatttgttgttgt atggtgacaagttatagatctcgatctatagcgtttcttgagagatttccaaacatttattcaagtcgta caattcttgtgtttaagcagtacagtgtaacc
351	SV1.0	tctgtcctcaccccatcttcccttctttcctgcaccgttacgcttactcgcatgtgcattgagtggtgca cgtgcttgaacaaacagctacactcacatgggggcgggttttcccgccctgcggcctctcgcgag gcccaccctcccctcctcccataactacagtgctttggtaggtaagcatcctgatccccgcgga agctgctcacgtggcaactgtggggacccagacaggttatcaaaggcacccggtctttccgccttc aggagtatccctgctagtgaattctagtagggctctgcttggtgccaacctccccaaatgcgcgct gcgggagtgctcttccccaactcaccctagtatcctctcatgtgtgtgcttggtcagcatatctgaga

[1707]

		cgatgttccgctgtcccagaccagtccagtaatggacgggccagtgtgcgtagtcgtcttccggctt gtccggcgcatgtttggtgaaccggtggggtaaggttggtgtgcccaacgcccgtactttggtgat acctcaagaccacccaggaatgccagggaggtaccccgcttcacagcgggatctgaccctggg ctaattgtctacggtggttcttcttgcttccacttctttctactgttcgccacc
352	염소 코브바이러스 5Δ40	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttctttcacacaactctactgctgacaactcactgactatccact tgctctcttgtgcctttctgctctggttcaagttccttgattgtttttgactgcttttcactgcttttcttctca caatccttgctcagttcaaagtc
353	염소 코브바이러스 5Δ40/3Δ122	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttctttcacacaactaaagtc
354	염소 코브바이러스 5Δ40/3Δ86_원위	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttctttcacacaactctactgctgacaactcactgactatccact tgctctaaagtc
355	염소 코브바이러스 5Δ40/3Δ 122_코작	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttctttcacacaactgccacc
356	염소 코브바이러스 5Δ40/3Δ86 근위	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttctttcacacaacttcactgcttttcttctcacaatccttgctca

[1708]

		gttcaaagtc
357	파라보바이러스	tgaaccgttacgcaccactcagttggtgtttggtggcaccaatgatggaacaaaaggctacaccac ttgggctacggcccgcgccaccttgtggcgcaaagacattagaagaatagcataccgcccactag ggccctgcagccagcagggtaacgggcaagcacttctgtctccccggtagaacggtataggctgt acccacggccgaaaactgaactatcgttacccgactccgtacttcgcaaagcttagtaggaaactg gaaagttcgagttattgacccggagtgttccccccactccagaaacgcgtgatgagggttgccacc ccgaccatggcgacatggtgggcatccctgcgctggcacgcggcctctaagaggataactcgct cctactggtaaccgaagagccccgtgagctacggtttattcctccgcctccctgaatgcggctaatc ctaacccatgagcagttgccatagatccatatggtggactgtcgtaacgcgtaagttgtgggcgga accgactactttgggatggcgtgtttccttgttttctccatttgttgttgtatggtgacaagttatagatct cgatctatagcgtttcttgagagatttccaaacatttattcaagtcgtacaattcttgtgtttaagcagta cagtgtaagg
358	파라보바이러스 5Δ48	aacaaaaggctacaccacttgggctacggcccgcgccaccttgtggcgcaaagacattagaaga atagcataccgcccactagggccctgcagccagcagggtaacgggcaagcacttctgtctcccc ggtagaacggtataggctgtacccacggccgaaaactgaactatcgttacccgactccgtacttcg caaagcttagtaggaaactggaaagttcgagttattgacccggagtgttcccccactccagaaac gcgtgatgagggttgccaccccgaccatggcgacatggtgggcatccctgcgctggcacgcgg cctctaagaggataactcgctcctactggtaaccgaagagccccgtgagctacggtttattcctccg cctccctgaatgcggctaatcctaacccatgagcagttgccatagatccatatggtggactgtcgta acgcgtaagttgtgggcggaaccgactactttgggatggcgtgtttccttgttttctccatttgttgttgt atggtgacaagttatagatctcgatctatagcgtttcttgagagatttccaaacatttattcaagtcgta caattcttgtgtttaagcagtacagtgtaagg
359	파라보바이러스 5Δ67	tgggctacggcccgcgccaccttgtggcgcaaagacattagaagaatagcataccgcccactag ggccctgcagccagcagggtaacgggcaagcacttctgtctccccggtagaacggtataggctgt acccacggccgaaaactgaactatcgttacccgactccgtacttcgcaaagcttagtaggaaactg gaaagttcgagttattgacccggagtgttcccccactccagaaacgcgtgatgagggttgccacc ccgaccatggcgacatggtgggcatccctgcgctggcacgcggcctctaagaggataactcgct cctactggtaaccgaagagccccgtgagctacggtttattcctccgcctccctgaatgcggctaatc ctaacccatgagcagttgccatagatccatatggtggactgtcgtaacgcgtaagttgtgggcgga accgactactttgggatggcgtgtttccttgttttctccatttgttgttgtatggtgacaagttatagatct cgatctatagcgtttcttgagagatttccaaacatttattcaagtcgtacaattcttgtgtttaagcagta cagtgtaagg
360	파라보바이러스 3Δ60	tgaaccgttacgcaccactcagttggtgtttggtggcaccaatgatggaacaaaaggctacaccac ttgggctacggcccgcgccaccttgtggcgcaaagacattagaagaatagcataccgcccactag ggccctgcagccagcagggtaacgggcaagcacttctgtctccccggtagaacggtataggctgt acccacggccgaaaactgaactatcgttacccgactccgtacttcgcaaagcttagtaggaaactg gaaagttcgagttattgacccggagtgttcccccactccagaaacgcgtgatgagggttgccacc ccgaccatggcgacatggtgggcatccctgcgctggcacgcggcctctaagaggataactcgct cctactggtaaccgaagagccccgtgagctacggtttattcctccgcctccctgaatgcggctaatc ctaacccatgagcagttgccatagatccatatggtggactgtcgtaacgcgtaagttgtgggcgga accgactactttgggatggcgtgtttccttgttttctccatttgttgttgtatggtgacaagttatagatct cgatctatagcgtttgtaagg
361	등출귀 피코르나바이러 스	tttgaaaggggtgcggatatcatggcgtttctcgccatgatatccgcacattgcaaacccatattgca tacccactgggtatgcattatggggaggccccttcaccccctccccccccaattacctttttccccctct agtaaccatacgctttactcagcgtaactactccgggttacgtgatgaagaagaggctacggagatt

[1709]

		ctcgggctacggccctggagccactccggctcctaaagatttagaagtttgagcacacccgccca ctagggcccccccatccaggggggcaacgggcaagcacttctgtttcccccggtatgatctgatagg ctgtaaccacggctgaaacagagattatcgttatccgcttcactacttcgagaagcctagtaatgatg ggtgaaattgaatccgttgatccggtgtctcccccacaccagaaactcatgatgagggttgccatcc cggctacggcgacgtagcgggcatccctgcgctggcatgaggcctcttaggaggacggatgata tggatcttgtcgtgaagagcctattgagctagtgtcgactcctccgcccccgtgaatgcggctaatc ctaaccccggagcaggtgggtccaatccagggcctggcctgtcgtaatgcgtaagtctgggacg gaaccgactactttcgggaaggcgtgtttccatttgttcattatttgtgtgtttatggtgacaactctgg gtaaacgttctattgcgtttattgagagattcccaacaattgaacaaacgagaactacctgttttattaa atttacacagagaagaattaca
362	등출귀 피코르나바이러스 5Δ105	cccctccccccccaattacctttttcccccctctagtaaccatacgctttactcagcgtaactactccggg ttacgtgatgaagaagaggctacggagattctcgggctacggccctggagccactccggctccta aagatttagaagtttgagcacacccgcccactagggcccccccatccaggggggcaacgggcaa gcacttctgtttccccggtatgatctgataggctgtaaccacggctgaaacagagattatcgttatcc gcttcactacttcgagaagcctagtaatgatgggtgaaattgaatccgttgatccggtgtctccccca caccagaaactcatgatgagggttgccatcccggctacggcgacgtagcgggcatccctgcgct ggcatgaggcctcttaggaggacggatgatatggatcttgtcgtgaagagcctattgagctagtgtc gactcctccgcccccgtgaatgcggctaatcctaaccccggagcaggtgggtccaatccagggc ctggcctgtcgtaatgcgtaagtctgggacggaaccgactactttcgggaaggcgtgtttccatttgt tcattatttgtgtgtttatggtgacaactctgggtaaacgttctattgcgtttattgagagattcccaaca attgaacaaacgagaactacctgttttattaaatttacacagagaagaattaca
363	등출귀 피코르나바이러스 5Δ201	attctcgggctacggccctggagccactccggctcctaaagatttagaagtttgagcacacccgcc cactagggcccccatccaggggggcaacgggcaagcacttctgtttccccggtatgatctgata ggctgtaaccacggctgaaacagagattatcgttatccgcttcactacttcgagaagcctagtaatg atgggtgaaattgaatccgttgatccggtgtctccccacaccagaaactcatgatgagggttgcca tcccggctacggcgacgtagcgggcatccctgcgctggcatgaggcctcttaggaggacggatg atatggatcttgtcgtgaagagcctattgagctagtgtcgactcctccgcccccgtgaatgcggcta atcctaaccccggagcaggtgggtccaatccagggcctggcctgtcgtaatgcgtaagtctggga cggaaccgactactttcgggaaggcgtgtttccatttgttcattatttgtgtgtttatggtgacaactctg ggtaaacgttctattgcgtttattgagagattcccaacaattgaacaaacgagaactacctgttttatta aatttacacagagaagaattaca
364	코브바이러스 SZAL6	tttcacaccctcttttccggtggtccggacccagaccaccgttactccattcagctacttcggtacctg ttcggaggaattaaacgggcaccctacccaagggttacatgggaccatattcctcctccccctgtaac tttaagttttgtgcccgtattcttgactccaggcggatgttgtgtcgcccgtcctgtgaacaaacagct agacactttcctccccctccctctgggctgctccggcagtccactccctccccccagcgtaacatgcc ccgctggagtgatgcacctggaagtcgtggacgtgggttagtaacttcggtgaaaacccactataa tgacaactggttgacccccacactcaaaggactcgagtctttctcccttaaggctagcccggccac atgaatttgcagctggcaactagtgagtccaccatgtcccgcaacctcggctgcggagtgctgttc cccaagcgtatgccttccttctgtaagagtgcgcctggcaagcacatctgagaagtcgttccgctgc gtcgtgccaacctggcgacaggtgacccagtgtgcgtagacttcttccggattcgtccggctcttct ctaggaaacatgcgtgtaaggttcatgtgccaaagccctgcgcgcggtgttcttctactgccctagg aatgtgccgcaggtacccctacttcggtagggatctgagcggtagctaattgtctacgggtagtttc atttccatcttctcttcaggtcgacatc
365	코브바이러스 SZAL6 5Δ158	ttgactccaggcggatgttgtgtcgcccgtcctgtgaacaaacagctagacactttcctccccctccct ctgggctgctccggcagtccactcctcccccagcgtaacatgccccgctggagtgatgcacct ggaagtcgtggacgtgggttagtaacttcggtgaaaacccactataatgacaactggttgacccccc acactcaaaggactcgagtctttctcccttaaggctagcccggccacatgaatttgcagctggcaac

[1710]

		tagtgagtccaccatgtcccgcaacctcggctgcggagtgctgttccccaagcgtatgccttccttc tgtaagagtgcgcctggcaagcacatctgagaagtcgttccgctgcgtcgtgccaacctggcgac aggtgacccagtgtgcgtagacttcttccggattcgtccggctcttctctaggaaacatgcgtgtaa ggttcatgtgccaaagccctgcgcgcggtgttcttctactgccctaggaatgtgccgcaggtaccc ctacttcggtagggatctgagcggtagctaattgtctacgggtagtttcatttccatcttctcttcaggt cgacatc
366	코브바이러스 SZAL6 5Δ76	gaattaaacgggcaccctacccaagggttacatgggaccatattcctcctcccctgtaactttaagtt ttgtgcccgtattcttgactccaggcggatgttgtgtcgcccgtcctgtgaacaaacagctagacact ttcctcccctccctctgggctgctccggcagtccactccctccccccagcgtaacatgccccgctg gagtgatgcacctggaagtcgtggacgtgggttagtaacttcggtgaaaacccactataatgacaa ctggttgacccccacactcaaaggactcgagtctttctcccttaaggctagcccggccacatgaatt tgcagctggcaactagtgagtccaccatgtcccgcaacctcggctgcggagtgctgttccccaag cgtatgccttccttctgtaagagtgcgcctggcaagcacatctgagaagtcgttccgctgcgtcgtg ccaacctggcgacaggtgacccagtgtgcgtagacttcttccggattcgtccggctcttctctagga aacatgcgtgtaaggttcatgtgccaaagccctgcgcgcggtgttcttctactgccctaggaatgtg ccgcaggtaccctacttcggtagggatctgagcggtagctaattgtctacgggtagtttcatttcca tcttctcttcaggtcgacatc
367	코브바이러스 SZAL6 3Δ37	tttcacaccctcttttccggtggtccggacccagaccaccgttactccattcagctacttcggtacctg ttcggaggaattaaacgggcaccctacccaagggttacatgggaccatattcctcctcccctgtaac tttaagttttgtgcccgtattcttgactccaggcggatgttgtgtcgcccgtcctgtgaacaaacagct agacactttcctcccctccctctgggctgctccggcagtccactccctccccccagcgtaacatgcc ccgctggagtgatgcacctggaagtcgtggacgtgggttagtaacttcggtgaaaacccactataa tgacaactggttgacccccacactcaaaggactcgagtctttctcccttaaggctagcccggccac atgaatttgcagctggcaactagtgagtccaccatgtcccgcaacctcggctgcggagtgctgttc cccaagcgtatgccttccttctgtaagagtgcgcctggcaagcacatctgagaagtcgttccgctgc gtcgtgccaacctggcgacaggtgacccagtgtgcgtagacttcttccggattcgtccggctcttct ctaggaaacatgcgtgtaaggttcatgtgccaaagccctgcgcgcggtgttcttctactgccctagg aatgtgccgcaggtaccctacttcggtagggatctgagcggtagctaattggacatc
368	살리바이러스 SZ1	tctgtcctcaccccatcttcccttctttcctgcaccgttacgcttactcgcatgtgcattgagtggtgca cgtgcttgaacaaacagctacactcacatgggggcgggttttcccgccctgcggcctctcgcgag gcccaccccctcccctttcctcccataactacagtgctttggtaggtaagcatcctgatcccccgcgga agctgctcacgtggcaactgtggggacccagacaggttatcaaaggcacccggtctttccgccttc aggagtatccctgctagtgaattctagtagggctctgcttggtgccaacctcccccaaatgcgcgct gcgggagtgctcttccccaactcaccctagtatcctctcatgtgtgtgcttggtcagcatatctgaga cgatgttccgctgtcccagaccagtccagtaatggacgggccagtgtgcgtagtcgtcttccggctt gtccggcgcatgtttggtgaaccggtggggtaaggttggtgtgcccaacgcccgtactttggtgat acctcaagaccacccaggaatgccagggaggtaccccgcttcacagcgggatctgaccctggg ctaattgtctacggtggttcttcttgcttccacttctttctactgttcatg
369	크로히바이러스 B	gtataagagacaggtgtttgccttgtcttcggactggcatcttgggaccaaccccccttttccccagc catgggttaaatggcaataaaggacgtaacaacttttgtaaccattaagctttgtaattttgtaaccact aagctttgtgcacataatgtaaccatcaagcttgttagtcccagcaggaggtttgcatgcttgtagcc gaaatggggctcgaccccccatagtaggatacttgattttgcattccattgtggacctgcaaactcta cacatagaggctttgtcttgcatctaaacacctgagtacagtgtgtacctagaccctatagtacggga ggaccgtttgtttcctcaataaccctacataataggctaggtgggcatgcccaatttgcaagatccca gactgggggtcggtctgggcagggttagatccctgttagctactgcctgatagggtggtgctcaac catgtgtagtttaaattgagctgttcatatacc

[1711]

370	크로히바이러스 B 5Δ51	ccccccttttccccagccatgggttaaatggcaataaaggacgtaacaactttgtaaccattaagctt tgtaattttgtaaccactaagctttgtgcacataatgtaaccatcaagcttgttagtcccagcaggagg tttgcatgcttgtagccgaaatggggctcgaccccccatagtaggatacttgattttgcattccattgt ggacctgcaaactctacacatagaggctttgtcttgcatctaaacacctgagtacagtgtgtacctag accctatagtacgggaggaccgtttgtttcctcaataaccctacataatagctaggtgggcatgcc caatttgcaagatcccagactgggggtcggtctgggcagggttagatccctgttagctactgcctg atagggtggtgctcaaccatgtgtagtttaaattgagctgttcatatacc
371	CVB3	ttaaaacagcctgtgggttgatcccacccacagggcccattgggcgctagcactctggtatcacgg taccttgtgcgcctgttttataccccctccccccaactgtaacttagaagtaacacacaccgatcaac agtcagcgtggcacaccagccacgttttgatcaagcacttctgttaccccggactgagtatcaatag actgctcacgcggttgaaggagaaagcgttcgttatccggccaactacttcgaaaaacctagtaac accgtggaagttgcagagtgtttcgctcagcactaccccagtgtagatcaggtcgatgagtcaccg cattcccacgggcgaccgtggcggtggctgcgttggcggcctgcccatggggaaacccatgg gacgctctaatacagacatggtgcgaagagtctattgagctagttggtagtcctccggcccctgaat gcggctaatcctaactgcggagcacacaccctcaagccagagggcagtgtgtcgtaacgggcaa ctctgcagcggaaccgactactttgggtgtccgtgtttcattttattcctatactggctgcttatggtga caattgagagattgttaccatatagctattggattggccatccggtgaccaatagagctattatatatct ctttgttgggtttataccacttagcttgaaagaggttaaaacattacaattcattgttaagttgaatacag caaa
372	CVB3 3Δ91	ttaaaacagcctgtgggttgatcccacccacagggcccattgggcgctagcactctggtatcacgg taccttgtgcgcctgttttataccccctccccaactgtaacttagaagtaacacacaccgatcaac agtcagcgtggcacaccagccacgttttgatcaagcacttctgttaccccggactgagtatcaatag actgctcacgcggttgaaggagaaagcgttcgttatccggccaactacttcgaaaaacctagtaac accgtggaagttgcagagtgtttcgctcagcactaccccagtgtagatcaggtcgatgagtcaccg cattcccacgggcgaccgtggcggtggctgcgttggcggcctgcccatggggaaacccatgg gacgctctaatacagacatggtgcgaagagtctattgagctagttggtagtcctccggcccctgaat gcggctaatcctaactgcggagcacacaccctcaagccagagggcagtgtgtcgtaacgggcaa ctctgcagcggaaccgactactttgggtgtccgtgtttcattttattcctatactggctgcttatggtga caattgagagattgttaccatatagctattggattggccatccggtgaagcaaa
373	SAFV	cacttatttaattcggccttttgtgacaagcccctcggtgaaagaacctctctcttttcgacgtggttgg aattgccatcatttccgacgaaagtgctatcatgcctcccgattatgtgatgttttctgccctgctgg gcggagcattctcgggttgagaaaccttgaatctttttctttggaaccttggttcccccggtctaagcc gcttggaatatgacagggttattttcttgatcttatttctactttgcgggttctatccgtaaaaagggtac gtgctgccccttccttctctggagaattcacacggcggtctttccgtctctcaacaagtgtgaatgca gcatgccggaaacggtgaagaaaacagttttctgtggaaatttagagtgcacatcgaaacagctgt agcgacctcacagtagcagcggactcccctcttggcgacaagagcctctgcggccaaaagcccc gtggataagatccactgctgtgagcggtgcaaccccagcaccctggttcgatgatcattctctatgg aaccagaaaatggttttctcaagcctccggtagagaagccaagaatgtcctgaaggtaccccgc gtgcgggatctgatcaggagaccaattggcggtgctttacactgtcactttggtttaaaaattgtcac agcttctccaaaccaagtggtcttggttttccaattttgttga
374	SAFV 5Δ46	cctctctcttttcgacgtggttggaattgccatcatttccgacgaaagtgctatcatgcctccccgatta tgtgatgttttctgccctgctgggcggagcattctcgggttgagaaaccttgaatctttttctttggaac cttggttcccccggtctaagccgcttggaatatgacagggttattttcttgatcttatttctacttttgcgg gttctatccgtaaaaagggtacgtgctgcccctttccttctctggagaattcacacggcggtctttccgt ctctcaacaagtgtgaatgcagcatgccggaaacggtgaagaaaacagttttctgtggaaatttaga gtgcacatcgaaacagctgtagcgacctcacagtagcagcggactcccctcttggcgacaagag cctctgcggccaaaagccccgtggataagatccactgctgtgagcggtgcaaccccagcaccct

[1712]

		ggttcgatgatcattctctatggaaccagaaaatggttttctcaagccctccggtagagaagccaag aatgtcctgaaggtaccccgcgtgcgggatctgatcaggagaccaattggcggtgctttacactgt cactttggtttaaaaattgtcacagcttctccaaaccaagtggtcttggttttccaattttgttga
375	SAFV 5Δ93	gtgctatcatgcctccccgattatgtgatgttttctgccctgctgggcggagcattctcgggttgaga aaccttgaatcttttttctttggaaccttggttcccccggtctaagccgcttggaatatgacagggttattt tcttgatcttatttctactttttgcgggttctatccgtaaaaagggtacgtgctgcccccttccttctctgga gaattcacacggcggtctttccgtctctcaacaagtgtgaatgcagcatgccggaaacggtgaaga aaacagttttctgtggaaatttagagtgcacatcgaaacagctgtagcgacctcacagtagcagcg gactcccctcttggcgacaagagcctctgcggccaaaagccccgtggataagatccactgctgtg agcggtgcaaccccagcaccctggttcgatgatcattctctatggaaccagaaaatggttttctcaa gccctccggtagagaagccaagaatgtcctgaaggtaccccgcgtgcgggatctgatcaggaga ccaattggcggtgctttacactgtcactttggtttaaaaattgtcacagcttctccaaaccaagtggtct tggttttccaattttgttga
376	SAFV 3Δ47	cacttatttaattcggccttttgtgacaagcccctcggtgaaagaacctctctcttttcgacgtggttgg aattgccatcatttccgacgaaagtgctatcatgcctcccgattatgtgatgttttctgccctgctgg gcggagcattctcgggttgagaaaccttgaatcttttttctttggaaccttggttcccccggtctaagcc gcttggaatatgacagggttattttcttgatcttatttctacttttgcgggttctatccgtaaaaagggtac gtgctgcccccttccttctctggagaattcacacggcggtctttccgtctctcaacaagtgtgaatgca gcatgccggaaacggtgaagaaaacagttttctgtggaaatttagagtgcacatcgaaacagctgt agcgacctcacagtagcagcggactcccctcttggcgacaagagcctctgcggccaaaagcccc gtggataagatccactgctgtgagcggtgcaaccccagcaccctggttcgatgatcattctctatgg aaccagaaaatggttttctcaagccctccggtagagaagccaagaatgtcctgaaggtaccccgc gtgcgggatctgatcaggagaccaattggcggtgctttacactgtcactttggtttaaaaattgtcac
377	SAFV 코작	cacttatttaattcggccttttgtgacaagcccctcggtgaaagaacctctctcttttcgacgtggttgg aattgccatcatttccgacgaaagtgctatcatgcctcccgattatgtgatgttttctgccctgctgg gcggagcattctcgggttgagaaaccttgaatcttttttctttggaaccttggttcccccggtctaagcc gcttggaatatgacagggttattttcttgatcttatttctacttttgcgggttctatccgtaaaaagggtac gtgctgcccccttccttctctggagaattcacacggcggtctttccgtctctcaacaagtgtgaatgca gcatgccggaaacggtgaagaaaacagttttctgtggaaatttagagtgcacatcgaaacagctgt agcgacctcacagtagcagcggactcccctcttggcgacaagagcctctgcggccaaaagcccc gtggataagatccactgctgtgagcggtgcaaccccagcaccctggttcgatgatcattctctatgg aaccagaaaatggttttctcaagccctccggtagagaagccaagaatgtcctgaaggtaccccgc gtgcgggatctgatcaggagaccaattggcggtgctttacactgtcactttggtttaaaaattgtcac agcttctccaaaccaagtggtcttggttttccaattttgttgaccgcc
378	GLuc CK dCTG1	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttttcttttcacacaactctacGTctgacaactcactgactatcca cttgctctaaagtc
379	GLuc CK dCTG1_2	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc

[1713]

		atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttttcttttcacacaactctacGTcGTacaactcactgactatcc acttgctctaaagtc
380	GLuc CK dCTG1_2_3	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttttcttttcacacaactctacGTcGTacaactcacGTactat ccacttgctctaaagtc
381	GLuc CK dAll	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttcccccttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttttcttttcacacaactctacGTcGTacaactcacGTactaC TcactGTctctaaagtc
382	CK SZ1-L1S	gggggtgggggggggcctcggccccctcaccctctttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt ccccctttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcatgttccgctgcggatgccgacc gggtaaccggttccccagtgtgtgtagtgcgatcttccaggtcctcctggttggcgttgtccagaaa ctgcttcaggtaagtggggtgtgcccaatccctacaaaggttgaaccacccaggaatgccaggga ggtaccccgcttcacagcgggatctgaccctgggctaattgtctacggtggttcttcttgcttccactt ctttctactgttcgccacc
383	CK 아이치 스캠 (AV-S)	gggggtgggggggggcctcggccccctcaccctctttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt ccccctttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcatgttccgctgcggatgccgacc gggtaaccggttccccagtgtgtgtagtgcgatcttccaggtcctcctggttggcgttgtccagaaa ctgcttcaggtaagtggggtgtgcccaatccctacaaaggttgattctttcaccaccttaggaatgct ccggaggtaccccagcaacagctgggatctgaccggaggctaattgtctacgggtggtgtttcatt tccaatccttttatgtcggagtc
384	CK 아이치 투프 (AV-L1)	gggggtgggggggggcctcggccccctcaccctctttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt ccccctttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcatgttccgctgcggatgccgacc gggtaaccggttccccagtgtgtgtagtgcgatcttccaggtcctcctggttggcgttgtccagaaa

[1714]

		ctgcttcaggtaagtggggtgtgcccaatccctacaaaggttgaactgccctaggaatgccaggca ggtaccccacctccgggtgggatctgagcctgggctaattgtctacgggtagttttccttttttctttttca cacaactctactgctgacaactcactgactatccacttgctctcttgtgcctttctgctctggttcaagtt ccttgattgtttttgactgctttttcactgcttttcttctcacaatccttgctcagttcaaagtc
385	CK SZ1-L2	gggggtgggggggggcctcggcccccctcaccctctttttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt ccccctttttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcctgatcccccgcggaagctgctcacgtggcaactgtggggacccagaca ggttatcaaaggcacccggtctttccgccttcaggagtatccctgctagtgaattctagtagggctct gcttgcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttgattcttt caccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggctaattgt ctacgggtggtgtttccttttttctttttcacacaactctactgctgacaactcactgactatccacttgctct cttgtgcctttctgctctggttcaagttccttgattgttttttgactgctttttcactgctttttcttctcacaatcc ttgctcagttcaaagtc
386	CK 아이치 트라이투프 (AV-L2)	gggggtgggggggggcctcggcccccctcaccctctttttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt ccccctttttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcatgtgcctggcaagcatatctgagaaggtgttccgctgtggctgccaacct ggtgacaggtgccccagtgtgcgtaaccttcttccgtctccggacggtgcgttgtccagaaactgct tcaggtaagtggggtgtgcccaatccctacaaaggttgattctttcaccaccttaggaatgctccgg aggtaccccagcaacagctgggatctgaccggaggctaattgtctacgggtggtgtttccttttttctttt tcacacaactctactgctgacaactcactgactatccacttgctctcttgtgcctttctgctctggttcaa gttccttgattgtttttgactgctttttcactgcttttcttctcacaatccttgctcagttcaaagtc
387	CK 스캔 결실 (ΔS)	gggggtgggggggggcctcggcccccctcaccctctttttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt ccccctttttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcatgttccgctgcggatgccgacc gggtaaccggttccccagtgtgtgtagtgcgatcttccaggtcctcctggttggcgttgtccagaaa ctgcttcaggtaagtggggtgtgcccaatccctacaaaggttgattctttcaccaccttaggaatgct ccggaggtaccccagcaacagctgggatctgaccggaggctaattgtctacgggtggtg
388	CK 투프 결실 (ΔL1)	gggggtgggggggggcctcggcccccctcaccctctttttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt ccccctttttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcatgttccgctgcggatgccgacc gggtaaccggttccccagtgtgtgtagtgcgatcttccaggtcctcctggttggcgttgtccagaaa ctgcttcaggtaagtggggtgtgcccaatccctacaaaggttgatttccttttttctttttcacacaactct actgctgacaactcactgactatccacttgctctcttgtgcctttctgctctggttcaagttccttgattgt ttttgactgctttttcactgcttttcttctcacaatccttgctcagttcaaagtc
389	CK 트라이투프 결실 (ΔL2)	gggggtgggggggggcctcggcccccctcaccctctttttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt ccccctttttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcgcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaa

[1715]

		aggttgattctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccg gaggctaattgtctacgggtggtgtttccttttttctttttcacacaactctactgctgacaactcactgact atccacttgctctcttgtgcctttctgctctggttcaagttccttgattgtttttgactgcttttcactgctttt cttctcacaatccttgctcagttcaaagtc
413	RhPV	gataaaagaacctataatcccttcgcacaccgcgtcacaccgcgctatatgctgctcattaggaatt acggctccttttttgtggatacaatctcttgtatacgatatacttattgttaatttcattgaccttacgcaa tcctgcgtaaatgctggtatagggtgtacttcggatttccgagcctatattggttttgaaaggaccttta agtccctactatactacattgtactagcgtaggccacgtaggcccgtaagatattataactattttatta tattttattcaccccccacattaatcccagttaaagctttataactataagtaagccgtgccgaaacgtt aatcggtcgctagttgcgtaacaactgttagtttaatttttccaaaatttatttttcacaatttttagttaaga ttttagcttgccttaagcagtctttatatcttctgtatattattttaaagtttatagggagcaaagttcgcttta ctcgcaatagctattttatttattttaggaatattatcacctcgtaattatttaattataacattagctttatct atttata
414	할라스타비 아르바 (1x mut)	cttgattctaaccttgccgtatggtgccctaacgggttcatttaatcatgcgatgagggttgctatacc gcatccattctaaggcgattcaatgcttcatttaggaatttttgttgacgattaaaaggtacccccacaa aaacaaaaccaatctttacttgattttcgttttaactgaccactgcgatcccaaattttcgccttcttatca aagtatgttgtgttctttgggtgtacaacctgagaacttgtctacaactacatattactcgaggaagaa attcggtttaagccgtgccttctcacgtttagtatatctatctGgacacaccttcttcatcttctaatccc catctagtctcctgatcagagacgtcgttattaacaaataacccccctttgttaataagagacaaagta caatcaagctaagttctcttggagttcctgtaggaacttagccattgtgatagagtcataagtctatgt gcatagacagctctagctcaccatttccttcccaacccatctttttcatcagcttaactctatgaatccga tgcaaaaaccattctaacatcttatggtgctttccaagccaaatgagagctcactcttttgagccgcta tttaatggacaataaacgttttatagtgtacatcatattgtaaaaacaaa
415	오스시바이러스 (Oscivirus)	cctcggtccctctttccgtcgccgcccacgacgttaaatgcggtgttgtggtgcttaggtgccacac cactgctatttgggtccccctttccccctatatatgtttgtttgtttatttcaatttcttgaggattggcacctc cttatgccaaatctaaatcgtggaggatcccaggctttctggtctttaacagaactccacgtccaggt catagaaactggttggtaggctgcctgagtagtccatttgctagtagtcccttgtgaacagggtggct cccgtttactgctggtattcccggtgtaggtcgccatggtggtaacaccatcctgcattgtgtgtgaa ccagtaccgcaaggatagcaaggtatgaacacttgtggacgaaatggtaagtgatcaattcactttc atggccggaaggtcacgtggcaatcatgccacccaggtaccctcctctgggaggatctgagggt gggctaagcagaccctgccatgtggctgaacttttcccttattgttttactttgtaacatttatagttgtgt tagtgatttgtgtgttgtgccctgtgagctatatccagtataagttcgcagctagaagttaatccttcg acatcggctgtattggaa
416	카디시바이러스 B	caccaacccttgacctgtaatgtcagtggacagagtgctcctctgttcccggttaccgtgttccagg acacgattgtaatcctgcgcctcaccagcgctgcgtgcacgtctgcataaggaaacgtgccttccc catgtctctatcaattctttggtgagtgaccgccctagttgctcatcctatgggattcttctctcatgggt tctttgtggcatgcgaatgtcaccttaattggaggtctttaattagatatcctttcttcatctttgatatgag tgtcggaatttgattcctagtctctgcaaaacaacccccacttgatgaattcaacttttcaaccgcacaa acataatcaggtttttaaattgaatgtttctaaattctaaatttagtttatttaagtagtttgccatcttgact cgatgtaaaattgtcatacaagtcttctttttctttttctttacactttgaagtttgcacttagcagtcgttctg cacagctttcgagttttgtttgatcgacatcgcaacttccacccacctctctttttctagtgttgaatgcg gctaatcctaacccgagagcaataaacccaggtttattgtcgtaacgcgcaagtcttggacggaac cgactatacacacacctctttaccctttagtacacccttggtacg
417	PSIV (Xba1 에 대해 2x mut)	gcaaaatgggtacgtagttaaccactgcgtatcaggattgcaggccacgaagggtatttgcatatct ttctatgcggtattacggcttaaaacccgttgtatcttgtTgtttgactgcctgtatcactagtggccatt ttatttaggttagagaccccctgatagtaggagagttacaaactctttaaaaattgttgaccccggaaa agatggtgaccccctgtaagtagttgatcAagaagatctatgcgctggcatagtaatccagtgtttcc tgttttaggatgacctctgaaagtagatgaccgtggaaagtcacgtagtgccccaataagcacgttt gggcagcgtgcgctatcacaaggcttgatctccgaggagccccttgttttagctggctggaagcca atgatcttaagtagataagtgctgttgcttgtagttcaacagaaagctttgagtacgtctttcttgcgag aaagaacacatgcattcttatgctctcaattctattatttttattttgggcgaaaggaaagctctcacgc gagtacgaatagccaacccttat
418	PSIV IGR	GCTGACTATGTGATCTTATTAAAATTAGGTTAAATTTCGAG GTTAAAAATAGTTTTAATATTGCTATAGTCTTAGAGGTCTT

[1716]

		GTATATTTATACTTACCACACAAGATGGACCGGAGCAGCCC TCCAATATCTAGTGTACCCTCG
419	PV 마호니	ATGAGTCTGGACATCCCTCACCGGTGACGGTGGTCCAGGCT GCGTTGGCGGCCTACCTATGGCTAACGCCATGGGACGCTAG TTGTGAACAAGGTGTGAAGAGCCTATTGAGCTACATAAGA ATCCTCCGGCCCCTGAATGCGGCTAATCCCAACCTCGGAGC AGGTGGTCACAAACCAGTGATTGGCCTGTCGTAACGCGCA AGTCCGTGGCGGAACCGACTACTTTGGGTGTCCGTGTTTCC TTTTATTTTATTGTGGCTGCTTATGGTGACAATCACAGATTG TTATCATAAAGCGAATTGGATTGGCC
420	REV A	GGGGTCGCCGTCCTACACATTGTTGTGACGTGCGGCCCAGA TTCGAATCTGTAATAAAAGCTTTTTCTTCTATATCCTCAGAT TGGCAGTGAGAGGAGATTTTGTTCGTGGTGTTGGCTGGCCT ACTGGGTGGGGTAGGGATCCGGACTGAATCCGTAGTATTTC GGTACAACATTTGGGGGCTCGTCCGGGATTCCTCCCCATCG GCAGAGGTGCCTACTGTTTCTTCGAACTCCGGCGCCGGTAA GTAAGTACTTGATTTTGGTACCTCGCGAGGGTTTGGGAGGA TCGGAGTGGCGGGACGCTGCCGGGAAGCTCCACCTCCGCT CAGCAGGGGACGCCCTGGTCTGAGCTCTGTGGTATCTGATT GTTGTTGAACCGTCTCTAAGACGGTGATACTATAAGTCGTG GTTTGTGTGTTTGTTTGTTACCTTGTGTTTGTTCGTCACTTGT CGACAGCGCCCTGCGAATTGGTGTACCCACACCGCGCGGCT TGCGAATAATACTTTGGAGAGTCTTTTGCCTCCAGTGTCTTC CGTTTGTACTCGTCCTCCTCTCCCTCTCCGGCCGGGATGGG
421	트로피바이러스 A	tgtcgcatgttgccaacatcaaaattctgggagagtcgcgaactccttaacactgccttgcctcgac ggagccgttgttatagtgtcgacgggatacaaacattaaactaaacccacttgcctcgacggaacc ccttacctttttattttattttatagtatgaaagtgaatcttgtatgaatgttcatagaaaactgcaaatgagt accacgtctaacatgagagaatgatactggagaaatccaagtttagaagtcactacgaatcccagc ggaaacaagggaattctgagcttctaataggcgtttaagactatttgcaaaattctggtgcgtaagtg atattttcattgcgtagaacgctggtaaccactccggctagtataagcattgttagtcacttattatgaa actccacactatcctttctggagaagcacacaaacttacatggtaaagctagaccattatcttaagcg gtgagtacactgcaaccttgtaacaatgcttgtatgactacttttgtatatcttgagcaatattgttgag gtggacatgtccaaaggtaatgttgttgggaatggaggggtccattttcccgtgcacgtagtgtact agtattgggtgatagccttgcggcggatcaaccatgtattttaatccgttgactttcac
422	시마피바이러스 A	ttgggaaatccccaatgcttctttcaacaccgcctgactatgcggtggcgcttcggctcaaacaact agtcacttcccctcttaactactacccaagacttctaactaccctacctacttatttgtctaaatttcaa actttattctcacgcgtcttataaacatctttctatttgttatggtatgttttgtgatttgtgtggtgtatttc atttaatgggatctagtggaccgtgccccggttgggtatccgctccctttaaatgtttgcaagcactct tgacattataacctatcatttagtttacttgtttgtatgatcgtatttctgaatcgtaacatttatgcaattctt tctcgccgagacttgtctaggagataaagttcctgcatatttagtgttacggttgtataatggagactta gatagcttcacactgaggacgctttttcgctatcctttgacctgattcaggccagtgtggagttaatg attgtatggatgggccctacaatttgtctaagacttggtgatagcctcgcggccgctcgccatttata caactgaatagcggttgaaactctct
423	사코브바이러스 A FFUP1 (1x mut)	tcacgcgcttttccggtggtcacccaccgttagggagcgccagcgttcgcgcttccgctaccaggt gacacactccttttccccctccccattcccgttcccatcctctggactggtttctcctcacgattgacca gcagctgggagctgttaccagacgttggacagtaagtcccggatgcactatagggctggtggcta gtgcttggtaagcactcaacgccatacctaatgtgtacctcggcttgccctcctggtcgtggtgacc ggctgtttctcttcccttggctcCagacgggctggtgtcctaccaccaccgttgcatgcagacctcc ccctgcgcactcgaacgccctgtcccagcagggttagtatgtgctgtgcagatctgcatgtgacac cccatccactggtagagcaggaagttgccctagctaacgcggcaagtattactttccgctacacgt ccttgagattcctcggacctctggaactagggtgactgtgggcttgggaaaacccaccttggtcctg tactgcctgatagggtcgcggctggccgaccagtggatgtagccagttgttttgggat
424	로사바이러스 C NFSM6F	cagggagatctccatgaataatcttttccaccctctttagcgtctatgctattgaggacgggttggag ccccgttgacccagcgtcagagtgtgtcggtagcaggctttctgctctcgccccatgccggccaca

[1717]

		cctcccattagtgatgtgaaggttgtaagttacatgtgaaaaggtttctaataattgagctgaatgtag cgattacctaaggtgagcggattccccccacgtggtaacacgtgcctctcaggccaaaagccaagg tgttaaaagcacccctttaggtaggccactaccccgtggcctcagttctcttagaagattcacttagta gtgtgtgcactggcaactcttaagcagagctagtgagtgggctaaggatgccctgaaggtacccg caggtaacgttaagacactgtggatctgatcaggggctcgagtgctgaagctttacagaggtagct cgagttaaaaaacgtctatgcccctccccccacgggagtgggggttcccccacaccaattttagatt gcact
425	로사바이러스 2 GA7403	ccaggcatggcgttaaacatgcattcccttcccctagtaacctcccttcgcccctttcccacgttgta ccccctccgagatggctgctaaggcgcttgctgctacagcagtctcgtgtttcgggtgttataagtg ctttcttttccactccactccctgcctatggggagcggaacggccttgtctcggtcgttgcttcttgca gatcttcacccctccaggctttctggactcgccaggggtggagtagtaggcgcactgtctaagtga aggtagcagtgttgttggcgaagagttgtggacctactttgagtttgtagcgatcatccagagctag cggatctccccacgcggtaacgcgtgcctctaggcccaaaaggcacggtgttcacagcacccttt ggatggcgggggtgccccccctccgcacttaaagtagaaaaacagcttagtagtcaaataacatgg ctttcctcaagcattcagtgctacatgggactgaaggatgcccagaaggtacccgcaggcaacga taagctcactgtggatctgatctggggccctgggccaggtgctatacacctggttaaaaccaaatct ggtagtcagggttaaaaaacgtctaagtcccaccccccgggggacgggggttcccttaaaccct caactgacacc
426	리마바이러스 A	cgaattccggacatctcctttcgggggcgagcgtcaccgtgcccctcatggaggcaactgtgcctc taatcggtgacccactgagaaaattttctttctacgtggctaaacaatgcaacttttataataacacaaa tttaatgcttaatcttaacaccaaagatttgaacatatgtttggaaagtggcacacttcaaacattgcat agttgctaggggtgaagtccctttaaggggttgcagaggatctttcctctttatgagcggctaggagt atcttcttgatattatgtggtcgtgcaactcacttcccagatgtatgacggtgtactaagcgattggaa ctagtcataacctctttgaatttggtattgcgagtctagcagggggatatttaccgctaaagggtgac acactcgtgagggtggccctttggtgtgtgtatatttattccgcccatcttgcatggggtgctaaaattct aatgctgtgaaataaccatttctgaatacattctctacatttggagtcaaatatgaggaatgccactca ggtacccttgacatgatcttggatctgagagtgggctaattatctaattatttggcgactttctaaaatct tctgtttttagtggtgacaatttatggttataaa
427	라피바이러스 LPXYC222841	gtgtccgggaagcgactcaagcttttgactgagtctctacaccttcatccgtaacatctttaagtttatg tgcctatggacctctagtgcactgccatcaccgggggtgtattggactggtttttccacaatccattca tcctgaggaattttggctttgttactaggatggtcccaccacacgcttatctgtgcctattgtgtcaacc atgttcttaagtagttgtgcccgtgggtgagtagataaccacaacaatccgataaagcatctcgcaa ggatgtgagtaatggagtgtatgtgctacagagacccacaacctgaaccaagagagacacagtg aggattgtaaagggggaactctttgaaagggcatgtcccgcaattcctactgactgacaccgggg gttggtgtcggtggattttagcaaatcctgttactgggtgatagccttgtgcacttcacttggttcttgta taagtgctgta
428	라피바이러스 WHWGGF74766	tgcgaatttattcgcacagtctcttttccccccatcttgtgtgtgtgatggggtaagccgcagagtaata cctactctgctgcaaacacactcactcttttctatctactttatatcatgtaataataagtagggaacata ttcaattcatattgttcatctcactgaacccgcatgaaggactgcattgcatatcctggacgaagtgac gtggaatatttggacatttatggattggacaccattacgctttgtgcctctacggagatgtaaccataa tcttaagtagtagtaccccagcacaagaggataaagtggcatacacgacaacgggtgttgctcgc accttagtaatgtggatgttcacccttggagcgtgctgaaactctgtgggtaaagacacacattagta caaatgtgggggaactcactgaaagggcatgtcccgtgtactggtgtgccggaaagtgggggtc gctttctggagaacttagtagttcttgttattgggtgatagccttgcggcggatcaactcacagttttaa tccgttgttttgcat
429	포에시바이러스 BCCH-449	actacacaatcgcaacacgcgcaagtttgtagtttgattggcgtgcaaatgtcaaatcaagcatata acacaatttggtggctgttggtgtttgttataggaattttggttgtgttgaaattgtggatgtgtaggaaa tatgcacaattacgtcagcgtcaggagttttataacctggcgcaacaccaaaatggtcttcgcgcttt aacatcaccagcgaggtgtaaacaaattgaagttgaattagatcgtgtataggccagggaaccatc cctcccaacgccacatcttgtggggaagttgggataatggtgggtctatatgaattggtctgtagac ccacagtgaagagtgaatagtatgcttgcggttccatttgttaatggtctagcatgggtgggggcgg caaccccgtgaggggttccccactggccaaaagcccaggggttagtcatttcaaccaaggaagct ggtaacctggtgacctgaacttgagtggtgagacccccttgctagagtgtgtaaaccgattgtaagc attttgtttgcttagtatctgtggtataagcagtcaattttgtataggctcaaggctgtggtagttagtag

[1718]

		atgcccggaaggttattactgatccggggaccgtgactatacattaggtaaaccggtttaaaaacc
430	메기리바이러스 ALY	ttcgggacactggatgggcgacttggtggggctgccactctatcttgacctttcgttactgactttcg gatctctgactcctccttgtctcttgcgtttggtccacggacggactaattggaatgtttactggctaag cctcgttctgaaatacctagccaatgggttgtagtaggatcctggtgtttccattaaacctcttccga ccatagtagctagagttatggctgtgtaggatgtgggtaagaccgctttttgcgtatctcccacaaga caccggattatggatgtgtccgctggataaggctcgaaacctcccaactgaaggtggtgctgaaat attgcaagcctaggttgtgtagaggcaagtagatgcctgccgcgacattcgtcttccgccctttgg gttagtagtgtacctacatggacgtggggctgggaatccccaccttgcataacactggttgatagac ctgcggctggtcaagttactatggtataaccagttgaaatggct
431	메기리바이러스 E	gcttggcaacctcatatcgttactctgccgaccagtctgggtcgtgtggccacacaatgggattcgtt ctgttgtgtagagtcacatggcattactgggctgatcggtggggatccgttgccacccctaaaccct tacatttactggactgcttttcttggccccggaatgattcgctcacccgcgatgaggactgttgttctta ttatggcaggattacgcgtctggtccgcgtaaggactaattcctatgtttatacgttactacttgttct gaacggtgggcgccacccccgcctagtaggatcctggcttatcgtgtagacctctagggaccacatt agctagagtgtaggctgctatggatggagtagtgacccctttttgggtatcactctctaagactccgg aatgtgtcatagtacgctggaaatccttacttgtttttccatgaggggggaggtggtgctgaaatattgc aagccacccctcggttaaaacagtttggtgccgcttatgccatattaccgccccttgtagttgggctg ttttgcagctccgggttagtagagtaccatagtggacgcggtgttgggaatcaccgccttggctgc acactgcttgatagagctgcggctggtcaagctaattgtggtataaccagttgatttggcat
432	메기리바이러스 C	ttcccgaccggtctggcaaaccggacggttatcctggttagatgtctgatggttgctggaacgtggt ggctactgctgccaccttctggcttcctttaatgggcatctagctgggttctttgccacaatccatctta ctctcttacccatttttctattacccagacttgttgttaactggtaaagttgacctactggcttcgttttgag actattctggtgttggtggacactctttccacaagtagattgtatggagttcatgctcgttttgaaccgg gaatggcacaacccgtagtaggatcttgcctctgccatactaatctgcgcctgttgcttttagactatg ggctgctaaggatgacattggaacccctttttggatattccatgtcaagtcaactgtttcatctggtgta cgctggaaatccttgttccgaggtcttgtctggaggtggtgctgaaatattgcaagccacaggcagt tccttggacttggtgccgctatcagatgctacaccctctatgggcaaatgttgaaccttagtggacgc gtgagatgggaatccacgccggccatagactggctgataagctcgcggctgatcgagttgcaaca gtaatcagttgatttgccact
433	투도끼바이러스	tagaccccacctagccctttttccccgtcagtgggggggcttactcactgggcatctgttaatctggc ctaactagattgacaccactcccttggaacgtaactccacgctaactcactggctctacgcacagac acacggtctttctgctatccccggggaagataccagatggcgaccggctgtcccagcggcctagt agctactcgggttgagtacccaccacggttttgacgcctgctaaaattcaagagacagaggtaggg gtgcttagtgtgtgggggaagttcccacaagcgaggcaaagcattgctccctcgcgtcaccgggt gcaaggtaaattggctggacttccgctctacccttgctactcgccctcttcggagggttcgaagtga cactaggtatacgcatggttgggaaaccatgcctggcctactactgggtgatagcctggcggcgg gtccgtctcttggcttatacccgttgatttgggat
434	리부피 바이러스	tatctacatggggatccaggctgtatggaatgtctgtcttaacaagcactataccagaaagatccac ccaaagtggtgggactgggactgtgaggtgagaaatcccgaaaccagccttctcaagcgtcgga cgatctttctgttttagtgaacaccttgccttttaaatggatgacaacacccctcagcaaatcgcaatc tgaaatcccaaaagactgtttagccgaactctggtaatcactccggagaagtaggatacgcagccc ctgtggactcttgatttcaggactcaaggtagctagagctggaacttcatggaatgacaaaggaata tatgcacattgtgcgctttcctggccttgtagcccgtcgtgaggatatgtcgttgggaatcgacatctt agtccagtactgcttgatagagtgtcggctggcacagttacctgagaataagtcagttgtacttaaca tgaacaaaaaaaataactaccacaactaccacaatctaccaatacttgaattatgctgaatctcgtac agtaaaaacgttccgtggaaggacaagtattgaagtgcggttacatcatccgatacgcgctggatc cctca
435	아이치바이러스 A FSS693	cacccatacacccccaccccctttctgtaactcaagtatgtgtgctcgtaatcttgactcccacgga atggatcgatccgctggagaacaaactgctagatccacatcctccctcccctgggaggacctcgg tcctcccacatcctccctccagcctgacgtatcacaggctgtgtgaagcccccgcgaaagctgctc acgtggcaattgtgggtccccccttcatcaagacaccaggtctttcctccttaaggctagccccgat gtgtgaattcacattgggcaactagtggtgtcactgtgcgctcccaatctcggccgcggagtgctgt tccccaagccaaaccccctggcccttcactatgtgcctggcaagcatatctgagaaggtgttccgct gtggctgccagcctggtaacaggtgccccagtgtgcgtaaccttcttccgtctccggacggtagtg

[1719]

		attggttaagatttggtgtaaggttcatgtgccaacgccctgtgcgggatgaaacctctactgcccta ggaatgccaggcaggtaccccaccttcgggtgggatctgagcctgggctaattgtctacgggtag tttcatttccaattcttttatgctggagtc
436	아이치바이러스 KVGH	tactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatactccccccaccc ccctttgtaactaagtatgtgtgctcgtgaccttgactcccacggaacggaccgatccgttggtgaa caaacagctaggtccacatcctcctttccctgggagggtccccgccctcccacatccccccccca gcctgacgtgtcacaggctgtgtgaagcccccgcgaaagctgctcacgtggcaattgtgggtccc ccctcatcaagacaccaggtctttcctccttaaggctagccccggcgtgtgaactcacgttgggca actagtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaaccctg gcccttcactatgtgcctggcaagcacacctgagaaggtgttccgctgtggctgccagcctggtaa caggtgccccagtgtgcgtaaccttcttccgtcttcggacggtggtgattggttaagatttggtgtaa ggttcatgtgccaacgccctgtgcgggatgaaacctctactgccctaggaatgccaggcaggtac cccaccttcgggtgggatctgagcctgggctaattgtctacgggtggtttcatttccaattctttcatgt cggagtc
437	아이치바이러스 DV	tactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatacacccccacccc ttttctgcaacttaagtatgtgtgctcgtaatcttgactcccacggaacggatcgatccgctggagaa caaactgctagatccacatcctcccttccccgggaggaccccggtcctcccacatcctcccccca gcctgacgtaacacaggctgtgtgaagtccccgcgaaagctgctcacgtggcaattgtgggtccc ccctcaccaagacaccaggtctttcctccttaaggctagccccgatgtgtgaattcacattgggca actagtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaaccctg gcccttcactatgtgcctggcaagcatatctgagaaggtgttccgctgtggctgccagcctggtaac aggtgccccagtgtgcgtaaccttcttccgtctccggacggtagtgattggttaagatttggtgtaag gttcatgtgccaacgccctgtgcgggatgaaacctctactgccctaggaatgccaggcaggtacc ccaccttcgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaattctttatgtc ggagtc
438	뮤린 코브바이러스 1	gtaacttcaagtgtgtgtgctcgtaatcttgactcctgccggaatgccgcccggttcagtgaacaaa cagctaggcaagtccctccctttccccctgtggtcggttctcaccggccaccatccctccccagcct gacgtgttacaggctgtgcaaagcccccgcgaaagctgctcacgtggcaattgtgggtccccccctt tgtcaagacaccgagtctttctcccttaaggctagcccggtcccacgaacgtggaactggcaacta gtggtgtcactacacgcctccgacctcggacgcggagtgctgttccccaagctgtaaccctgacc caagactgtgctgcctggcaagcaccgtctgggaagatgttccgctgtggctgccaaacctggta acaggtgccccagtgtgtgtagtcttcctccagtctccggactggcagtcttgtgtaaagatgcagt gtaaggttcaagtgccaaatccctggaaggagtgaccctctactgccctaggaatgctgtgcaggt acccccaacttcggttggggatctgagcacaggctaattgtctacgggtagtttcatttcccatcctct ctttttttggcatc
439	돼지 코브바이러스 K-30	tttgaaaaggggtgggggggcctcggccccctcaccctctttccggtggccacccgcccgggg ccaccgttactccactccactccttcgggactggtttggaggaacataacagggcttcccatccctg tttaccctactccactcaccccctcccctttgaccaaccctatccacaccccactgactgactccttttgg atcttgacctcggaatgcctacttgacctcccacttgcctctccctttttcggattgccggtggtgcctg gcggaaaaagcacaagtgtgttgttggctaccaaactcctacccgacaaaggtgcgtgtccgcgt gctgagtaatgggataggagatgccaataacaggctcgcccatgagtagagcatggactgcggt gcatgtgacttcggtcaccaggggcatagcattgctcacccctgaatcaagtcatcgagatttctct gacctctgaagtgcactgtggttgcgtggctgggaatccacgcttgaccatgtactgcttgatagag tcgcggctggccgactcatgggttaaagtcagttgacaagacac
440	돼지 코브바이러스 XX	ccaccgttacttcactccactccctcgggactggtttggaggagcataacagggcttcccatccctg ttcaccctcaataccacccaccctttccctcaaccatccctatccacacccccactgactgattcccttg gattttgacctcagaacgcctacttgacctcccacttgcctttcccttctcggattgccggtggtgcct ggcggaaaaagcacaagtgtgttgcaggctaccaaactcctacccgacaaaggtacgtgtccgc gtgctgagtaatgggataggagatgcctacaacaggctcgcccatgagtagagcatggactgcg gtgcatgtgacttcggtcaccacgggcatagcattgctcacccgtgaatcaagtcattgagattcct ctgacctctgaagtgcactgtggttgcgtggctgggaatccacgcttgaccatgtactgcttgatag agtcgcggctggccgactcatgggttaaagtcagttgataagacac
441	염소 코브바이러스	gggggtgggggggcctcggcccctcaccctcttttccggtggccacgcccgggccaccgat acttcccttcactccttcgggactgttggggaggaacacaacagggctcccctgttttcccattcctt

[1720]

	12Q108	ccccctttcccaaccccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcatgttccgctgcggatgccgacc gggtaaccggttccccagtgtgtgtagtgcgatcttccaggtcctcctggttggcgttgtccagaaa ctgcttcaggtaagtggggtgtgcccaatccctacaaaggttgattctttcaccaccttaggaatgct ccggaggtaccccagcaacagctgggatctgaccggaggctaattgtctacgggtggtgtttcctt tttcttttcacacaactctactgctgacaactcactgactatccacttgctctcttgtgcctttctgctctg gttcaagttccttgattgtttttgactgctttcactgctttcttctcacaatccttgctcagttcaaagtc
442	토끼 코브바이러스	gggctataaatatgggcattcctcttccccctttcccctttgaagatgagtgcgcatattcttgactccg cctggattggccgcccaaggcgtgaacaagcagctaggccaccatgacactgcggtggtgtccg aacccgcgggtgccttcacgggcacctgtggtatgtaggactcccaccgtggtcttccctttccccc tcaatctttcccccctggttcgactaacgggaccagtgctggaacctgtccggtgaacggtatagca ggcccccccggcagaaacacccggtgcttacccctaaggctagccccctccatgaatttggttg gggcaactagtgggtgtacagttggcgtgaaccctccggtctaggagtgctcttgcccaatcctct gtgtgtgccttgcagtagggactggcaatccttcgcgtaggtgatccgctgtgccatgccatcctgg cgacaggaggcccagtgtgcgcaacctacgtcccttctgggtgctgcattgcattacctttggagta agcttggtgtgccgaaacccccagggtttacgtaccactcgtggtgtgaggaatgtgccgcaggtac cccatccttgaggtgggatctgagcggtagctaattgtctagcaccactttcttccttttttctttgctgg tcacg
443	아알리바이러스	ttgaaaggggtgctcagggtagctccctgagctcttccctccaccctctttcaacgtctggcccac gatacgggccacctttcaatcttaactaactatcccttttaatctatttggattttctggtttagaataatttg gaacacataattggattatcttttaggattgtggataggatttgttcgggatatcactcccttcctgtgct aacacatattctaattccctccttttgtctattatctcttggaggtggtgctgaaatattgcaagccacttg agtgtatagatgaagtaggctcaagatgaatgttgttgttactcaaggcaagtgtagctatcactaaga tattggtaacgtgaaacggattaccggtagtagcgtgatcttccgtcttagtgctctagtgactagag gacaacgacatggcatcacatatcttaaccctccagttttggcatccgggacagaatgggctggat atccgctttctttctggggtatgtgatgggtggtattggggtaaccaccttgaccatgacgctcgata agagtgaccgcctgatcattgaaacctctagtataaaattcaggctgaaatc
444	그루소피바이러스A	tgcctgagtaggattgtgaatttaggtatgagagggttagccaacccattctgaaccataatagatac gtcaatctgaatccatctaaatctatctcttaggcagtggtgctgaaatattgcaagctactagggata gacgtgatctgattcaagaacctatctaatgtggtgatgagaaggctaggtttatccatagtaatccct tgttctgaacaggcaatgcacatgctctagtaggatctcgggctctgcgattggctctaaaccgacc aatccaggtagaggcactaagtgtaggacttgccaaaatgtattacatgctggtaccgactcactag tctggaaactccacactgaaagtgactgggggggcccccatcacatttgtgctactgcttgatagaa gttgcggctggtcaacttggattggtataaccagttgaa
445	그루소피바이러스B	gccatccgtaggttctggtaaggttccatcaactgttggggcgctagttgctatgaccgcattcacg gacggatgatttatagtatcacccaatccgggcacaacttctttagccacttcttccacattactaagg gctctcttgccgagtttcaacgtctagtccacgacacggaccttcctacttttctatcttcttattttctct actaaattggtatctggtactgaagatatgcggattgtgattttgtgcctgtctaaactaaccctattcta gggttaggtgggtaccatatactaatggtgaacaggattacctatgtatccattagtccctatggatct ggcgacccacaaactcatgttcatagagaggctaagctgagtgctcgccgaataagcattgcttca ggtgccgactattgtctggaaaccactcagtgatagctatagggggggggccccgtagcatctgcct tactgcctgatagggtggcggctggtccatgaacatgcagtaaccagttgacttgac
446	얀청 오스벡스 그레나디어 앤초비 피코르나바이러스(Yancheng osbecks grenadier anchovy 피코르나바이러스)	ccttagggtctggaatgcgtcctttctgggcacttccacaatcctaaggtaattttcaacgccagcga tggagcgatatccaaagaaccttatgttttagttcgtcttgtatgtttataaaatataaaattgggatta gcaacccacaaaaacatttttgttattctcaccacatgagagggtggttaaacctcttcgtaacctatc ttgcttgattggctacttgggctagatttaggacaacctctttagcaagcctaaattactcctcgtactt gcacctggtaacaggcgtacctggaggttacggtggcgctaacttggacttctcgttaattcgtgca agtttaaatgatgcctattttgaatacaagaaagtatgatagtaacttagggcgtgaagttccgcttaa cataaggcagtataagtactaagataaggtgtaagacctaccttaataactgttgtcttttctcatggtc tttccgtgggagcccttgctaggggtaaagttaagtattctcaaataatttttcattcaaactctttctctc tgtttt
447	쉴면조	ccactcgcacttcctggatagtgcgttagatatgcccgacatatcgcttccaggaccaaagccccc

[1721]

	갈리바이러스 M176	cttttctctttcccaaccagcttcgccactcaagctgtaatttccatgtccggtctttccggccttagtatc atggaaatgtggtcgtgctcaaatgaaattgagttgacattgatcaatgaaagttgcactgaacttttg ctaaactggctagcgccacctggtgtgtgccgttggtctcctcacatggtaacatgtgccaacgggg cccgaaaggctagtgggcaattaccgctccaagggaggggtacccaccccgacctgaacagcg gtaatgaagctcacctcccaggctctgaccccgagaagtttagttatttagtaggtgtaattagtactt gtgattggtcaatttgatagtagtttgaaacgttatggatgaatgagtagacccccctgaaggtacccc attacatgggatctgatcagggccacattctgcgtgtctccccgcacttgtggttaaaaccatgaaa gttcatcccaaacaatcttttcctcttctttttctttttagtggtgacaacctactggattggtgattaccaat ctgtactagtgttgtattaagacttgttgtgtggagaaaatggactctttcaagaagatttttg
448	팔코바이러스 A1	taaaagggacgcggtgtggcagctttggctgtcatgccgtgttctccctttaccccaaggactagc cttgggggttttccaaattccttttccctgtaggctttacttctctttatctatctttttctgtaactaagttttgc ctattctaaaaatattttagaatgtgtttggatgtaactaagtttgtgcctgcccctaaaaatattttagggc ttgtttggataacctcgtccctttgtgttcagtgccgcacaatttgctaggcactgttcacttcctttgtttg tccattatgtatgctaaggtatgaattccatcatatgcttagcctctacatgcataatcttattccctccct ggtgcaaactacgcccccaacatatgtgaatctttttaagcatattcctgaccccacacatatatatgtg ttctcgtgaattcccccaccgtgaggtggtcacttggacgtggtgtgtgtcacacagcatatatatga tgcaggatgttgttttttaagataagcatatgtccttagtgctttgcatcatttcctccacaccccgtgaat gcggctaatcttaaccctgttgggtccgtgggtaaaccaacccattaaccacaggacggaaccga ctactttcgggagtgtgtgtttctttttcttctttgtcact
449	트레모바이러스 B	ttcaaatggcccctgggttgatacccagtggtcatttggacactttggtaaggaggtgtaattatcctt cccatgtggaacctagtgcttaggtttacttatatgttctttgtttgtccttgtactttctatcgggcaat cttgttgttcaatacaatatgtatttgaactgcctaagataaattcagttttcaaccaaccctctcttgg ggttgtgtctttctttctttctatatcctcttaagctgacttacttgctaatccgactcctcgtcaacggg agggtaaagcagtatcactagggtattgtgatgtaggagaaaaagtaagtagagatagtgcatgta acgaaagtgacttggtactttaaactctcttaatcccaaagtgtggtattggtcatgttggagtaggct acgggtgaaactccttcacatttagtaatgtgttcacacgctaacgctacggtagatgacagactag gtcttattctcaacgtaggggacgggtgtatgttcatgattagccacatattaaggttttgaggggct gagtcatataagtatgtgcattaatttctggtactggtccctggggactggccctttctaggttgatttt agtttccccaatttttaaaaactaatgagatttacgac
450	디델피스 아우리타 HAV	tctttggtctggggaactaaaatactcagacccgcgtttgcctagcgatataggctttaattgttgtttgt cattgtgcgtttgatatgtgttttaatgtaaataataattctagcaggttctagttcttgatcatgtcctcttt aaggcactcatttcaacttgctatctttctttttcttccttggttctccctacaccaaatgcactggccgct gcgcccggcggggtcaaccacatgattagcatgtggctgtaggtgttgaaggctgggacatgaac atcaatggaatagtgcgcatgcttactggggtccattgaagtagtgggatctttctattggggtaggc tacgggtgaaaccccttaggttaatactcatattgagagatacctggataggttaactgtgctggat atggttgagtttaacgacaaaaagccatcaacagctgtggacagaacctcatccttagattgctcac tatggatatgtgctctgggcgtgtttcttgcatgatggccattggtcaattcatgcctgggccaatgta ggattagccttaaattacttttaaaagtagcctcatttagctggactaatggtggggcgtatgatcctg catttggcctctggggtaatcaggggcatttaggtttccacataatagcaaat
451	헤파토바이러스 G1	gcaaggggtggttttaaccttgcacgcgtttaccgtgcgttaacggttttccatgtttgtatgtcttgttt gtattatgtgttttgtaaatattaattcctgcaggttcagggttctttaatcatgttgggctgtacccacac tcaacttttggccataagtgagtttcttaacgaacctttaacacaggatgttattagggcccaatatttt ccctgaggccttctttggcctctatttttttccccttttctatctccttgtattccgggctcacgtgatgcca atggactgacccatgcgcccgtgggggttaactactggagtagccagtagctgtaggtgctaaaa gtcacgtacgtgtaagactggacgagacctctcagctataactgaaagtagtaagtatgtctgaact tcttgaaggggtaggctacgggtgaaacccctttaggttaatactcatattgagagatacctctgatag gtgaaggtttccggtagaggtgagtttaacgacaaagcctctcaacggatgtgggcccacctcatc agcaagatgctttcataccaatacgtaggggctgggttgttgagttcagtcccaagcgtccctcc cgcaaggttgtaggggtactcaggggcatttaggtttccacaattaaacaaataca
452	헤파토바이러스 D	ctttggatgcccatagtgcgggggtataaataccgcactccctttagctgttccgagggtatcggaa cctatatgtttgttttctgtctgtctgtcagctttatgtgtgctcgtcccctttagggcactcatttcagctt gctttcattctttttcttccccggttctcaccttaccggaggcactggccgttgcgcccggcggggtca acctagtgattagcactaggctgtaggtgtctaaagtggtgacattaagacttggtaactgatttcag cactgttaactgatgttggggatgacttgattgatcttctggaaggggtaggctacgggtgaaaccc

[1722]

		cttatcttaataccactatgtagagatagattcagtaggttaagggcagtggataaggttgagttcattt tggacaataaaccttcaacactggtggacccaatctcactgaccagatgctttcttgactgatccttc agaggggtgattcttctgaataggttgccttgacactgatgcctgagacccattgggtcgggccttа aatcatggaactccactggactttcatggcctagcttctgccttagacagactctggggccccacga ccctctgggcccttcggggtactcaggggcatttaggtttttccacaattaaaagagtta
453	헤파토바이러스 H2	gtcatgtttctctttaagaacactcaattttggccataagtgagactcttgtcgaacctttcatgtcagg accatgttagggccattatccttttccctggggcattcttcttgcccctgtttcatctttctatcatctttctt ccgggctctcacaatgccaatggagcgaccgatgcgcacgtcggggttaacccatggattagcc atgggctgtagctgctaaaagttgtgactcctgaagcatactatcaatggtagtagatgtaactgaaa cactgaagcttctctgatcttgaaagaagggtaggctacgggtgaaaccccttcaggttaatactcat attgagagatacctttggtaggttaacgttggcggataatgttgagtttaacgacaataaacattcaac gcctgtgggcgaacctcaccaatttcatgctttgaagtgaatgtgcgtagggtctctatcggagatg ctatgtggatggtgccctcccctggaaacaggttgtaggggtactcaggtgcacttaggtttccacatt ttaaagatttttc
454	헤파토바이러스 I	ggctgcctgtgtctcaggggtaagtactggggccgcgttgaccgtgcggtacggttatgcttttaga ttaggatgtccgtctgtccggcactctcttttgcttaaaatggccttaaatccatgggaggcgtaacca tgggccctttgttacctagacatgattgcattgggggccgtccttggggcttaggccccagccatttc tcttgactcgtctaagagtttacttcatcctttctttacttttatttccaggctctcagcatgccgacggct ctgaccactgcgcccggtggggttaactgcatgattagcatgcagctgtaggagttaaaagtgctg acaggccaattctgacgtaagtccactctatattaacttgatcaagtaaggttgattgatctttgtgaga gggtaggctacgggtgaaaccctctaggttaatactcatattgagagatacctccagaaggtgaag gttggcggatattggtgagttcttttaggacaaaaacctttcaacgcctgtgggcccacctcactggc acaatgctttcatccccaattgtgatgggtagtttggactgaaatcaggagtaacctgccctacgagt ttaggggtagttcaggggtatttaggcttccacatttgatagagtttatgagagtgagcc
455	헤파토바이러스 C	ttcaaaagccccagcggggtttcattaccccgctgtggcttttggacttccctaggatggggaagta aattaccatcctcgcgtttgccgtgcgttaacggctacttttcttctagctgtagaagtaaaattcagca tgttttatgtttgtttgtcttgtttgttatatacatttttacactcctacaaatgcacatgaagaacagtttgt agagattaacaaacgcttagctgaacctaggtggtgaatctagtagtaagataagtagaggaagct atacttaagttggttgggccctcgtgtttgctctataaacaaaaccaagtgagtagagtggatgaac agtactaaatccctgagtacagggaacctcacaggtgtgatacacttatgtctatgtgacctggttgg aggttgggcgtgccctatgatactggagtgggagatcttttggggaacccacgttttcacactgcct gatagggtcttgccgagagactcacttgtttcggctgtacttgtaac
456	피피바이러스 A	tgcgggtaaactcccgcatgtgtgaatgaggcgatgtcccaggaactaactgccgatcctggtttta actacgatccgtatttgttactaatgcgatatccccccattgtttgcctccatgttgttttcaacgcttttg gccttgagtgttatcaagtgttttagcgacatagtgggaagctacggctgcgtccccatttttgagtg gcgacccagttttagtggccactctgtccctgaactgcgctataatgtgaatttatgttcacaaaaac ggactgatgtaactgttaatgactaaggaatagtacctcactgaagtatcaagaccccgttcgagcg gtgtacatatatggatggaaaccttgtctgagtcatctcgaatactaatcaatgagggatgtcgagta agcatatcatgaaccacatagaatagtggggtttcggggttagaggctctctgcagcaatgtatctct aacaccatggccgaaatgagagatagagaccacgatgtttgtgtgtaagtaatgatgtgtggaaag aaaattctgaatgttggtatgatatcagtctaaggggagtggctcacctaagagctacccaaacattt cacagcagacaacataacgtactgagagtagttggaaggttccagaaatcagt
457	피피바이러스 C	cgcggttaaacccgcgcaacctctttcagccgcgtctgagtagcgcggttagtcctgatacacagt ttcctgttgggtactgtgtcttcgggtgaatgctcttgtgtgaatgttttaggctgttttaagggaagcgtt tccccgtgcgctgtgagggtttctcacgctctttcggggtgcagtctcttctgttgttcattaagatgta tggatgcactgttgtgaaggatttgtgaactggggatcgacaccccgtgaggggtgccccagtgtc ataggagtttgctggagaggtgtgttgctgtagtgactatccgtgacctggcattctaaggtgttga ccccaacctgtgagggtctggatcgcagtgttgaagtgctttggagggttcaatggggtttctgtagt ggatattatgtgcttgacgactactggtacgagtgtattgggggtctacatgtgtga
458	피피바이러스 E	ctcttccgatcttgggggttcgcccccatgtctcatttcaactagccgtgtgtctagttaacgcaccgc ctcaccctggtcgttatcgggtcggttcttgcgaccgttagatcgtgagcgtttctgaggatcagttc gtataagttctccggtgtggcgaccgtaaaatcgtcacgtcccatgcaatagatgacgttaaactcg tttgccagttacataaaggaatgttgttacttttaaattgtctgttacatttaacatcttgccagtatgatg ctactgtacactacgggtgtaggaaccttgtagtgtgacgtatcactcatatgtggatgggtgctcca

[1723]

		gacctttatggaagctctcagttagtagtgatccttgacttcattgagccctggtaacagtggaagtc aagatgtatatgttgctcaacacacttcggtgctacgaagctgtttgtggaagtactggcgaggttca ttctgaatcatatgtttgtcacatagtcagggagtgccgtcgcttacgacggaccctttttctttataatt acaaatctgtgtctcaagtgttgttggctggttttcttcttctgttttcattgttcatatatatacgtcagagt gaaagactcggtatatacaaaactgatccaga
459	아루아마바이러스	ttcaaaggtggcgggagagttggcctcacgctgtttagcgtgagagctggctctcctgccccttcc cctgagccggggatcttggctcattcccctctttctatcctcccctcattggactttacggatgacccg gcataaacttgacaaccgatgttggatttcccttgtggctgtgatggaggacataccctcgggtgta gttgtgtgcgtgtcgctctgcgactcgagcttcaaagtggtgctgaaatattgcaagcgtcgttgctc gattaacggagtggtacaatcctatgaacccaagtgcattcatgcgaaagccccggaggggtgag tagcatggactcgaatcagaagagctggagctcgcttggtacggcacgtagcattgctttgcctaa agaccaaggggtatggctataggtgggggcctatagcttgtccagtgctggttgacagactcgtg ctacgcgtctggttcgagtataagtagctgcaactcact
460	아비시바이러스 A	ttcactcgctttccccccctctctataggggcggtcttttaattcttattaatttcctactttactatcaaatt tcttctaagtagggactgaggtcacttagccctcccctctcctgggctttccagggttatagaggttcta aagctaagccatgtgtcttgagctacacttagtacaaaggtttagtaatgattgtacatgccagtaac cttctagtgcccatggattaaagagtggtaacactctccatggggcccgaaaggctagtgggcata gttggcatcaaggaaggggtccccaccccaacctgaattgctggctagaagctcaccttagaaga agtgctgggtgacaacgtgtccaatcgtgaacgactgatggaaacgtgtggagatggatatgtgg gggttcactgagtagatgccctgaaggagaaatctgatcaggggcccgtgactatacgctaggta aaccgggtataaaaaccatgaaaggtggcccaaaatctcttccttttatttttatttctatgttggtgaca gtcaag
461	아비시바이러스 B	caccccctactgccctaacccccaaagttagttatagggtggctccctacccttactccacggggta agccctaacccggttgaatctcaagatcagccttagcgaggactattagtaccgctcaaaccctttg cctgtagtgcccaggggtcacagaggggtgaccctctccctggggcccaaaaggctaggtggca agacagggtccaagtgaggggctactctaagtagccccaagctgaacatcctgtctgaagccacc cttgcagggccaggtttgattggggaaactagacaccagctttgtcctgggattgggggatatcg agttagtccaggaggtgcgagtagatgcccccgaaggtaccccaggcacatctgggatctgatcg ggggcccgtgactatacaataggtaaaccgggttaaaaaacatgaaagcgcctctctctttcctact tcttttattgactggtgacaaaaatagcagt
462	크로히바이러스 A	gttgaagtccatttcttgcttgcccccgatgaatcctgttaaggcctcacggccctaagggtgaaact cggttatcccctcctgtacttcgagaagattagtacaacactatgaaatctacatcttgtgatccggga taaccccaatcccagaaacctgtgatgggcgtcaccacccctcttatggtaacataagggtgtcgc cgcgttggcacaggacccttttgggctggatgttttagtaatggtgtcgaaggtcctattgagctaca ggagtttcctccgccctggtgaatgcggctaatcttatccctgagcctaaggttgcgatccagcaac ttgatggtcgtaatgcgtaagttgggggcggaaccgactactttccagaaggcgtgtttctttgttttg tctgttactatggtgcatgatatagatattgaatatttgatcttttgagctgtttcttatcttattgctacatc ctttcaggtgttggatttacattttggttaataag
463	푼사기시바이러스 B	gattttctggttatccctttggacttggtaggggcccacgtgcccacccacctctgtgtgtgttgattt ctaatcgatgcctggcagtggcggccacctctccttactggtaaacctccggtgagtgaagttgtca agctacaggtaccgtgcaggatgaaatgcgcacatgtgaacaaactaggagtcatacaccgggtc aaactctggaaacggagtccgggactctgaccttggttgggtgagctcgaggcatcacattgatgg acgcgattcgctatccttccctagtaggaccttgtggtgtacccctggttgggaatccagggctggt cgggtgcagggtgacagcctgttctccacctcaaccattgtaggagaaatcaaccect
464	림니꾀바이러스 A	ttctttggatatccatttaacgtgtaccctatacgataattggggtggattctggatgcctagttccagt gattggttaagaactcgtttactacgtatagtatgattagcaaagtgctcgattgatcacgtaatgatct atgtggttaaaaacccagtagtatggtatatactcagtagtgtacactgtgagtacaactcttggcgta gagagaacaattcacccgaatccgtggcgtatccatggaaataagtttacctaattgtatgttacaag gcatatgagacatttatgagatatggtttattttgactaaacgagtgtagaggtggtggagtctatcca acttcaagccatgcaattgttgtgttgattgatatcattgaccatttttgtggattgtgtacacatacaatt tgaaaattaacccccctcaagaataagacatgggaccattcgtggtagataccgtgctcggatgcttg agattagatgggttagactagttttggaatgagattgccgagaaagtcccgctagacatgttttacaa gtcgtggtattccgctagacttttttcgcagacacatggaagggtccatgtgttgtgcaattgcagggt gacagcccaactgcagagttttccttactagaataaaaatctgttgtcaatttt

[1724]

465	림니피바이러스 C	gtttctgagcactggtaagagcttagacaaacgtttttaaaatttattttctctgcaacttttgtttgtgttt attttttatttgttaattttgcgcctaagcatttgttgcgaagtatttgattcattagtaatattacttattgttta tttagatggtattcaaagtggtgggagtatcgaacccaagcgtcgtatgctatctccttgaacaatttttt aatcattgcgaagtgatcattgaaaaggataggtgtttaagaactcaaagagtgttaataatgttggg tgacaggtgtccccatagaatttattaacatgatttggactggttatctagtaagaagaaccatcgaa cgcacgagcgagcattgcttgcggggcagttaccctgcgtcgatgtaagtgtgtaccgggggggtg cacatgttgattctttatggcctgatagggtgcgtcattcgcgcctagataattagtataatgcgaatg gaataaatttac
466	오리바이러스	ggtcccaggccaatattcttcgtaaggcttggttccaattttccaccactcgtgtttgggttctggccta tggtacccagaggggcggtttgggggaattaactccccctccccctgtggtcctataccaccccaca cctctgtgggctttctttactatcttcttgttttccgacttttaaacactaggcaggcgcgcctagtcata caccgcccggctggtctttccagctcttgtgggcggtgcgcgctggtccatcgtgcccagcgacat agcaccttgtggacacctccgaacgccctccctgtatggggtggtgcccaggggtttcagtgtgg tgacacactccctggggcccgaaaggctagtgtgcaacaggtgaggtacagccagctgcccccg tggctggagggaccaagcttgtgaagcacacctcaccttcttgggggtgggctagtaagtggtga aagcatagtgtccgtgtcgctggccaacacttgggtcaagtccagccactcagtgagtagatgcc caggaggtacccctagtggatctgacttggggcctgttacttaatgcaggttaaaaactatgaaagc tgagtagtgtagcccggctggtggcttctcttccttattcattcatttttatggtgacaaacgcaactga agcc
467	HAV FH1	cttgatacctcaccgccgtttgcctaggctataggctaaatttccctttccctgtccttccctatttcctt ttgttttgtttgtaaatattaattcctgcaggttcagggttctttaatctgtttctctataagaacactcaatt ttcacgctttctgtctccttttcttccagggctctcccctttgccctaggctctggccgttgcgcccggcg gggtcaactccatgattagcatggagctgtaggagtctaaattggggacgcagatgtttgggacgt cgccttgcagtgttaacttggctttcatgaacctctttgatcttccacaaggggtaggctacgggtga aacctcttaggctaatacttctatgaagagatgccttggatagggtaacagcggcggatattggtga gttgttaagacaaaaaccattcaacgccgaaggactggctctcatccagtggatgcattgagggaa ttgattgtcagggctgtctctaggtttaatctcagacctctctgtgcttagggcaaacactatttggcct taaatgggatcctgtgagagggggtccctccattgacagctggactgttctttggggccttatgtggt gtttgcctctgaggtactcaggggcatttaggttttcctcattcttaaacaata
468	HAV HM175	cgccgtttgcctaggctataggctaaattttccctttccctttccctttcctattccctttgttttgcttgta aatattgatttgtaaatattgattcctgcaggttcagggttcttaaatctgtttctctataagaacactcatt tcacgctttctgtctcttcttccagggctctcccctttgccctaggctctggccgttgcgcccggcgg ggtcaactccatgattagcatggagctgtaggagtctaaattggggacacagatgtttggaacgtca ccttgcagtgttaacttggctttcatgaatctctttgatcttccacaaggggtaggctacgggtgaaac ctcttaggctaatacttctatgaagagatgccttggatagggtaacagcggcggatattggtgagttg ttaagacaaaaaccattcaacgccggaggactgactctcatccagtggatgcattgagtggattga ctgtcggggctgtctttaggcttaattccagacctctctgtgcttggggcaaacatcatttggccttaa atgggattctgtgagaggggatccctccattgccagctggactgttctttggggccttatgtggtgttt gccgctgaggtactcaggggcatttaggttttcctcattcttaaataata
469	파레코바이러스 F	ggtcggggagatgtgttcatgatcggttaacaccatcatggatcatctctcccgacctcttttgacc cagctatgggttaaatagtacttttctttttctctttgctttctttgtgttttgtttgttttgcaacatataaca agcattttatcagtattagtgtctgcaactgtataacaagcaaggtggagcaatcatgcgagtatatct caattgaattgtgacacacaagtgtgcactatgtggaataaatgccattttggccaaacctggttagc cagaccagtagtaggacaatttggcacccttagtgggcgcgacctagatgctagggatgagcaaa cctatttccccgagtacaggggctctccttcacctctacattttggacctcttttgagtatcctcgata gaaggtgaagtgacggtgtaccggatggttaattgatctcattgctgggtgacagcccgctaggac caggcagcatctttgtatggacctgtacatgtaac
470	파레코바이러스 D	acatggggcaggtgtgctgtgccaagagcaacactacggtggccgagccgatggttcgtcacca cgtagtaggactccgtagtgcttggttacggcggacgtaagtcagttgagtgatgtctaagtggcaa accatgagtacatggtaaccttgtgtggactcgcgggacggaatttcctatcccattgactccttgta gcaaggtgggtatacccaaccacaatggcagcaccctgggtgggaacccagggcctggatta gtatccagtcacacagcctgatagggtggcggctcagccactgaccagcgtctctaaataattgtg agctgttcatgcacc
471	파레코바이러스	cggtcatccccctttccccacagccggtgtgggttctaatcggctcctactaaacacctaagcatca

[1725]

	C	ctgcgcctctatctctcctatccacaggtctaagacgcttggaataagacatgtgggtgcaataggaagattagctagtccaatctctccttccagctacgcttctcccttcgatgagcgtagggggggccccc acctccctcatctctggataggctcttgctacggggctttcccgtctggaccagcaggcccactg gtgcgcttccattcaagtttagtgtgcattactgtctgaaatattgctttgctaggatctagtgtagcga cctgcatattgccagcggactcccccacatggtaacatgtgcctctgggcccaaaaggcatgtctttt gaccgtatgcagtacaaccccagtataggtcctttctatggcagtatggatctcagtgatgagtctat acagaatatggaagtggttcggatatgtcagcccgaaggatgcccagaaggtacccgcagataa ccttaagagactgtggatctgatctggggcccaccaccttcgggtgggtagaagctaaccatgcct tgggttaaaaaacgtctaagggctgaccagacccgggggatccgggttttccctatcttgacctact ctaatc
472	르준간 바이러스 87-012	ctcattgcccacacctggttggttcccaggttcatacaataaccatcaataaacttttaacatctaagat agtattatcccatactagactggacgaagccgcttggaataagtctagtcttatcttgtatgtgtcctg cactgaacttgtttctgtctctggagtgctctacacttcagtaggggctgtacccgggcggtcccact cttcacaggaatctgcacaggtggctttcacctctggacagtgcattccacacccgctccacggta gaagatgatgtgtgtctttgcttgtgaaaagcttgtgaaaatcgtgtgtaggcgtagcggctacttga gtgccagcggattacccctagtggtaacactagcctctgggcccaaaaggcatgtcatttgaccac tcaggtacacaacccccagtgatgcacacgcttagtaatggcttagtaacaaacattgattgatcatttt gaaagctgttaggaggtttaggtatgacgggctgaaggatgccctgaaggtacccataggtaacct taagcgactatggatctgatcaggggcccaccatgtaacacatgggtagaagtcttcggaccttgg gttaaaaaacgtctaggcccgccccccacagggatgtggggtttcccttataacccccaatattgtat a
473	파레코바이러스 A2	gccgtcgggccttacaccccgacttgctgagtttctctaggagagtcccctttcccagccagaggtg gctggtcaaacaatacccaaacgtaactaaacatctaagataacatagccctatgcctggtctccacc agttgaaggcatcttgcaataaaatgggtggattaagacgcttaaagcatggagtcaattatctttttct aactagtgatcttcactgggtggcagatggcgtgccataactctattagtgggataccacgctcgtg gatcttatgcccacacagccatcctctagtaagtttgcaaggtgtctgatgaggcgtgggaacttatt ggaaataattacttgctgcgaagcatcctactgccagcggatcaacacctggtaacaggtgccct ggggccaaaagccacggtttaacagacccttaggattggttaaaacctgagtaattatggaagata cttagtacctaccaacttggtaacagtgcaaacactagttgtaaggcccacgaaggatgcccagaa ggtacccgcaggtaacaagagacactgtggatctgatctggggccacctacctctatcctggtgag gtggttaaaaaacgtctagtgggccaaacccaggggggatccctggtttccttattttagtgtaaatg tcatt
474	파레코바이러스 A3	agagtccttttcccagccagaggtggctggttaaataatacctactgtaacaaaacatctaagatgta acaaccacacacctggtctccactggccgaaggcaactagcaataaggcaggtgggttcagacg cttaaagtgtgttgtacatattctttttctaacctgtgttttacacagggtggcagatggcgtgccataac tctaacagtgagataccacgcttgtggaccttatgctcacacagccatcctctagtaagtttgtaagat gtctgatgacgtgtgggaacctgttggagataacagtttgctgcaaagcatcccactgccagcgga tctacatctggtaacagatgcctctggggccaaaagccaaggtttaacagacccttgggattggtt caaacctgaactgttatggaagacatttagtacctgctgatttggtagtaatgcaaacactagttgtaa ggcccacgaaggatgcccagaaggtacccgtaggtaacaagtggcactatggatctgatctggg gccagctacctctatcttggtgagttggttaaaaaacgtctagtgggccaaacccaggggggatcc tggtttctttttaatttaagtaatcact
475	파레코바이러스 A8	gggccttatacccgacttgctgagtttctctaggagagtcccctttcccagccctgaggcggctgga taataaaggcctcacatgtaacaaacatctaagacaaaataatttgccttgcacctggtcccctactag ttgaaggcatctagcaataagatgagtggaacaaggacgcttaaagtgcaatgatagttatctttttct aacccactatttatagtggggtggtggatggcgcaccataattctaatagtgagataccacgcttgtg gaccttatgctcacacagccatcctctagtaagtttgtgagacgtctggtgacgtgtgggaacttact ggaaacaatgctttgccgtaaggctttcattagccagcggaccaccacctggtaacaggtgcctct ggggccaaaagccaaggtttaatagaccctaatggaatggttcaaacctggagcattgtggaaagt acttagtacctgctgatctggtagtaatgcaaacactagttgtacggcccacgaaggatgcccaga aggtacccgtaggtaacaagtgacactatggatctgatctggggccaactacctctatcttggtgag ttggttaaaaaacgtctagtgggccaaacccaggggggatccctggtttccttttatttacttttgtcaa t
476	파레코바이러스	ctctattagtgagataccacgcttgtggaccttatgctcacacagccatcctctagtaagtttgtaaga

[1726]

	A17	cgtctggtgacgtgtgggaacttgtgggaatcaatattttgctttaaagcatccattagccagcggat aaaacacctggtaacaggtgcctctggggccaaaagccaaggtttaacagaccctagtggattgg tttcaaaacctgaaatattgtggaacacactcagtacctactgatctggtagtaatgcaagcactagtt gtaaggcccacgaaggatgcccagaaggtacctgtagggaacaagagacactatagatctgatct ggggctggctacctctattttggtgagtcagttaaaaaacgtctagtgggccaaacccaggggga ccctggtttccatttatttttacaaaggcact
477	포타미피바이러스 A	cacatggaaagcttttcgcttccatgtttacgcacacactctctttgacaccctgttgtatggtgttaaa ctacaacatttgtctgtctataatcgtttattttgtttaccctatatgtacccaagtatttgattgcttgactc acataagcatcggtaacccatactgttttatgagctactacctctgctgtctacatacattttatatgaat ggtttgagctctgcctcaggatcaaacatggtaacatgttcctttggtcagttagaatcttattgtataat ctaaggtgtctattagtacgtagaaagttgtaacacatatggggcctgatagccgctatctctgatgg atgtaaggtaaccttctttaggtctgatacattctgcacaggatccaattttcggtgccctgtacgagt gcactcttatgcacgaggacgagatatgctacaacccactgcaaatttaaacccaaacttaaca
478	포타미피바이러스 B	tttcaacgtcgtggctgacgttaaaaagccacaattccacttacctttttaccttttatgtttaatgtttgtta gttttgtgatctttaacaaatagatctaaataatttgttggtaaccaatctcggatgtttcggctgcattgt agtttatttatttcattttagttgtaggtggccactacgtcctggaatcatacatggtaacatgtacctcg gcggttatccactattacgctaatctaagaatatttaaatgaaaatgtaagtgttacggctgactttgg gcctgatagttaaatgctcgcactgacagatagtaccctcctttaggatcgattctgttacatgggatc cattttggtgccccactgattcaacctctttgttgaaaaagagttagcatactacaaatttccaaacaa aaacccttttaatgactacaacttatgattttatgaattttactgctcttgaaaaagatattttgacattga tcgctgtactgtttcagacattcattgcatccatttttgttggctactcctcacaaactcaaaactttcca cacgagaaaccttgtttattgaatttgccttttatattttggaacttgttgttggatttattgtttgcttaatta ttgacctcacacctgttttaaacactacaat
479	베이하이 콘거피코르나바이러스	gggacaaccccacagctggtacaaccattgtgggttggtctccaccctttttcaaccgtggcaactt cggttaaagttgcaaatccccccctctccctattccacctccctttacttttcactccccatatatggtccca gattttattctacctctttatatttttatttagtacagtggtggtgaattactcccagcataaactttgctgg atcagtgttcatcaagcatactaattactaatgtactgagctatactattatctggcatctcacctggat aaccggtgtgaccatatttcctaggttgcctccctatgtattttgtagcacctgtgcatctgcacgttgg ggcgacaaattgtaggtttcctggcacgggtaagaattgtggaaagctagtatgcagttaatgcaa gggcgcgtttttcgctaccccgacactgctaaagttttgggaggggtcccttaaacatttctagtatt gagtgatagctttgcggcaggtcaccacaaccttactataaataaacctgttgaatctcac
480	돼지 사펠로바이러스 JD2011	tacgcatgtattccacactcatttccccctccaccctaaggtggttgtatccccataccttaccctcc cttccacaatggacggacaaatggatttgacctcacggcaaacacatatggtatgatttcggatacac ccttaacggcagtagcgtggcgagctatggaaaatcgcaattgtcgatagccatgttagtgacgc gcttcggcgtgctccttttggtgattcggcgactggttacaggagagtaggcagtgagctatgggca aacctctacagtattacttagagggaatgtgcaattgagacttgacgagcgtctcctcggagatgtg gcgcatgctcttggcattaccatagtgagcttccaggttgggaaacctggactgggcctatactacc tgatagggtcgcggctggccgcctgtaactagtatagtcagttgaaaccccccc
481	돼지 사펠로바이러스 A2	ttgaaatgggtgtggggtacatgcgtattacggtacgcatatattccacactcatttccccccctcca cccttaaggtggttgtatccccatacctttaccctcccttctaaaacagatggacaaatggatttgaact tatggcaagtgaatatggtatgacttggatacactttaacggcagtagcgtggcgagctatggaaa aatcgcaattgtcgatagccatgttagtgacgcgcttcggcgtgctcctttggtgattcggcgactgg ttacaggagagtaggcagtgagctatgggcaaacctctacagtattacttagagggaatgtgcaatt gagacttgacgagcgtctcttagagatgtggcgcatgctcttggcattaccatagtgagcttccagg ttgggaaacctggactgggcctatactacctgatagggtcgcggctggccgcctgtaactagtata gtcagttgaaacccccc
482	유인원 사펠로바이러스 1	ccaaggatctgttgcataggcgttgtatcccctaacctttttacctacccatcccaataggactggtatt tcggttttgattgagtaatggatactgattctatacctgttacccattcaggggaaaaatggagtttcttt catggatctgacttgatatgaccaagagtcaacactttgcgtgttggccgtatggaatgctttaaggtt tattctttggattatgacttcagggttggccgcccaggataaaaggcaattgtggtaagtgatgttagt cattggtggttgaaacctgcctaagacgtcctaggtctacgctgtgcgggccgaagtaagcttagg aataacagggagtatgccatttttctgctttcacccaacacgaccgtacacgaaagagctagaggca ctttggggcaaagggaaaagctttgcttagcccgaatgttcatttgagtccttgacgaatgcgtccc gtctgtcccgacggtgaggcgtatggcgcatgctcatggcattacccaatggtgtatctgtgaggg

[1727]

		gggggctcctcacacttagtctagtgctacctgacagggccgcggctggtcgtttgtgtatggtata accagtagtaatcccccatggattgctttaacttcccctcctcccttaccaagacattctctaag
483	유인원 사펠로바이러스 2	ttttaacttgttatgacattcaaggaaaaaatgtctttttcattatgggactgacctgtttatgaacatgag cagcggcactgctccacgggctatccgtgtaagaaatattgattattcttatggatcatgatttcagg gttggccgcccagtctaaaaggcaattgtggtaagctatgtaagtagttggctgttgaaaggagcc aagtacatcctaggtctacgctgtgcgggccgaagtaagacttggaacaactctgagtaggcagtt tttctctttagcccaacacgaccgcatactgaagagctagaggcactttggggcaaaggtaaaagc attgcttagaccgaatgttcaatgagaccttgacgagtgctgtcacagtgtcccctgatggcagtatg gcgcatgctcttggcattacccatatgtgtatctatagggggggggccccctatacttagtctagtgc tacctgacagggccgcggctggtcgtcggtgtgtggtataaccagtagtaatcccccatggattgc tttaactcccccctcctccctcaacaaaacttttctctaag
484	라보바이러스 C	ccgggtataacccggagttttggggcaggtccaagccccacataggaacatacgatccacggatc gtgtgttcttttatgctttctaaccttaccctttgtaaccattacgctttacgccgcatggtgtttggcggc accatgacgtggacaagaggttacgccattacgatatgtaccctccccttttggggagagaccgac caattatggtacagtatccaactgtattgtggtcaagtacttctgtttccccggtgatgcgggataggc tgtacccacggccaaaacctgctgatccgttacccgactcacatctacgaggaggctagtaaaag gcatgaagttcaagagtatgatccaaccagatccccactggtaaactagtgatgagggttcccgttc cgaacatggcaacatgtgggttccctgcgttggcactaggcccctccgaggggtgctctgaagat ggattgttgatgaagaccaatttgtgcatgtgtttatcctccggccctctga
485	라보바이러스 ANYC-B10	ccgacccactggtcgaaggccacttggcaataagactggtggaacaaggtcgcctgtagttgatt ggaaccttctttctaatgacttatgtcagcggtgctactcacaccgtaactctcctaccctatccccac gctttgtggaactaggagggatgagtgattcaagtaagtactgtcagaatggtgaaaataatctgat tctgaaacgctatggatccatcgaaagatggggctacacgcctgcggaacaacacatggtaacat gtgccccagggccgaaagccacggtgataggatcacccgtgtagtttgagatcatatcaatgttc atagtctagtaagatgatttgaaatctaactggtctgatggctaactgcttgtcttattgcggcctaagg atgtcctgcaggtacctttagagaaccttaagagactattgatctgagcaggagccaaggtggtctttt cccagccttggttaaaaagcgtctaagccgcggcaggggggcgggaggccccctttcctcccaaa ctataatatagattgt
486	파라보바이러스 C	gatgtatccccatccccagtgtgtatgccatactgcatagctcgcctatgccctatggattcacaac cctttcatataccctccctacccaaccccgtaaccacatgctttactccgcttggggttttgcggcccc atgttgtgacgaaatggctacgcaatcaatgcggctaatgggcctgccgcttttaagtggccccca gttagaagtttatgcacacccgcccattaggaggccaccagccaggtggtcagagggcaagcac ttctgtttccccggtgaagtttgataagctgtgcccacggctgaagcagacagatccgttacccgcc tcactactacgagacggctagtagtgtgtaatatccgaatttcattgatccgggtgttcccccaccc agaaacgtgtgatgaggagcggcacccctcctatggcaacatagggcctctcctgcgctggcac acgggctctatgagcatgaaatcaggagaaagtcacacgaagaccttattgtctagtgttgattcc tccgcccccctgaatgcggctaatcccaactccggagcgcccgctggcaaacccgccagaaga gcgtcgtaatgcgtaagtctggagcggaaccgactactttgggtgtggcgtgtttccttttatttccttt gtatttgtat
487	파라보바이러스 B	aacccataatccattgtccatcaatgttttatgggggggacccttttctccctcccctccaaatacct tttaccccctctgtaaccaagagtgtgcaaaatctatttactagcccagaattgcggcttctggggagg tttattcctcatgcctaacaagatgttacgcaaactccgggctacggccctgggctttgccctaaag atttagaagtttacactatcgtccaacaggaggacaacaaaccagttgttctaaggacaagcacttct gtttccccggtgagactggatagactgtacccacggttgaaactggttgatccgttacccgactcac tacttcgagaagattagtaggaaactgtgaaactgattccattgatccggatactttccccgtatcca gaaactactgatgagggttgacttcccgactacggcgacgtagtgtcatccctgcgctggcagtag gcctctttgaggatggaagatgtggatcggtaaccgaaggtcctattgagctagtgtttatacctccg gcctcctgaatgcggctaatcctaacccatgatctagtgctcacaaaccagtgagtagctagtcgta acgcgtaagtcgtgggcggaaccgactactttggagtgaccgtgtttcctattttactttgtttg
488	파라보바이러스 A3	accgttacgcaccactcagttggtgtttggtggcaccaatgatggaacaaaaggctacaccacttg ggctacggcccgcgccaccttgtggcgcaaagacattagaagaatagcatacccgcccactaggg ccctgcagccagcagggtaacgggcaagcacttctgtctccccggtagaacggtataggctgtac ccacggccgaaaactgaactatcgttacccgactccgtacttcgcaaagcttagtaggaaactgga aagttcgagttattgacccggagtgttccccccactccagaaacgcgtgatgagggttgccacccc

[1728]

		gaccatggcgacatggtgggcatccctgcgctggcacgcggcctctaagaggataactcgctcct actggtaaccgaagagccccgtgagctacggtttattcctccgcctccctgaatgcggctaatccta acccatgagcagttgccatagatccatatggtggactgtcgtaacgcgtaagttgtgggcggaacc gactactttgggatggcgtgtttccttgttttctccatttgttgttgtatggtgacaagttatagatctcga tctatagcgtttcttgagagatttccaaacatttattcaagtcgtacaattcttgtgtttaagcagtacagt gtaagg
489	펠리피바이러스 127F	gatgtcggatgacggctggccaccggggaaaaacggcaaatgtgcaccacctctgcaacccac gccgaccacgtttaaccatggcgttagtaggagtggaccactgcagtgggctctggtgtgcgaca gtcagtggtagagtagacagtcctgactgggcaatgggaccgcgttgcgtatccctaggtggcat cgagattcctctgctacccaccagcgtggactcctatgggggggccccataggctaggtctatac tgcctgatagggtcgcggctggtcgaccactgactgtataaccagttgtaactcact
490	부세피바이러스 A	ttgaaagacctcggcatatatcgttgtcacaacggtatatgtcgagatctttctccccaccccctcca attccctttccccctcttgcaacttagaagtttgtacacacagggcaataggatacgtgatccagcc aggacacgtgagctcaagcacttctgtttccccgtcccctcacgtactacgggaatgttagtaatttt gtgtgcactttagtaaggttgatccgggattaaccccaaatcccagaaactggtgatgagcgttacc accccgccgggcgaccggaaggtttcgctgcgttggcaccagggcttcggcacccagaaaaag gtaaagcaaatgaaggcgctactgtgctacgagaagtttcctccaggcccctgaatgcggctaatc ctaaccagtgatccaccggtgcaaaaccatgtactaggtggtcgtaacgcgcaagtcgctggcgg aaccgactactttgggtgtcctgtgtttccatattttatttattcaattttatggtgacaagagtaaagag atacagatttgcagcc
491	부세피바이러스 B	ttttctcccctcccccctccaactacctttcccccctcttgtaacgctagaagtttgtgcaaaccgcctgt agggtactgcaatccagcagtgcataggctaagctttttcttgttaccccaccccacattatactgagg aggattgtgaaattgtgttagtatgggttagtagcggtgacccgggtaaccccaacccagaaactc acggatgagatgaacaggaccccacatggtaacgtgtgttcgtctgccccgcaaggtgaggcc gtgagagctttgcacgcgaaaaccttgaaaaccccaaaagtaccttgagctcttcgctattttgtgtttc ctccaggaccctgaatgcggctaaacctaacccgcgatccgcacgtagcaacccagctagagtgt ggtcgtaatgcgcaagttgcgggcggtaccgactactttggtgttcctgtgtttccttttatttttatttga atttttatggtgacaacagctagaaaataagagtgaac
492	파코바이러스 Pf-CHK1	gtgtgtcatttctcccctcccccctcccaaacctttcccccctctaatcggattgattaacccggttaaag atgattaatggtttgtgagttgatatgatggcccggcattgaatccgggaattcttaagtaatggaatt gcatccaatatgaaagtgagtgtggcaagctcacaagtagtacttggctctgcccattatttgagga ccaactcttcttgactacaatgtgtttaaagtaaactggaccacattgtgtatccagacaactccatttg ataatgtacgctggaaacgttttcagtgcatagggtcctaaagtggtgctgaaatattgcaagctcaa tgggatactgaacgctgaaaaccgccgctgttatcatatgggcccctagtgggtaaatgttggcttt aggcatatactgcttgggaatgcagtactggttgtagacagggtgatagcctaccggctggcgtag ttgagttggtatagccagttgattgccat
493	HRVC3 QPM	ttaaagctggatcatggttgttcccaccatgattacccacgcggtgcagtggtcttgtattacggtac atttccatacagttttatacaccccacccgaaactcatagaagtttgtacacaatgaccaataggt ggtggccatccaggtcgctaatggtcaagcacttctgtttccccggcacccttgtatacgcttcaccc gaggcgaaaaatgaggttgtcgttatccgcaaagtgcctacgaaaagcctagtaacactttgaaaa cccatggttggtcgctcagctgtttacccaacagtagacctggcagatgaggctagacattccccca ccagcgatggtggtctagcctgcgtggctgcctgcacaccctgccgggtgtgaagccagaaagt ggacaaggtgtgaagagcctattgtgctcactttgagtcctccggccccgaatgtggctaacccta accccgtagctgttgcatgtaacccaacatgtatgcagtcgtaatgggcaactatgggatgggacc aactactttgggtgtccgtgtttcctgttttacttttcattgcttatggtgacaattgtatctgatacacttg ttacc
494	HRVB27	ttaaaacagcggatgggtatcccaccatccgacccacagggtgtagtgctctggtattttgtaccttt gcacgcctgtttccccattgtacccctccttaaatttcctccccaagtaacgttagaagtttaaggaaa caaatgtacaataggaagcatcacatccagtggtgttatgtacaagcacttctgtttccccggagcg aggtataagtggtacccaccgccgaaagcctttaaccgttatccgccaatcaactacgtaatggcta gtattaccatgtttgtgacttggtgttcgatcaggtggttccccccactagtttggtcgatgaggctag gaactccccacgggtgaccgtgtcctagcctgcgtggcggccaacccagcttttgctgggacgcc ttttacagacatggtgtgaagacctgcatgtgcttgattgtgagtcctccggcccctgaatgcggct aaccttaaccccggagccttgcaacataatccaatgttgttgaggtcgtaatgagtaattctgggatg

[1729]

		ggaccgactactttgggtgtccgtgtttccttttattctttatattgtcttatggtcacagcatatatagcatatatactgtgatc
495	HRVA73	ttaaaactgggtttgggttgttcccacccaaaccacccacgcggtgttgtacactgttattccggtaaccttgtacgccagttttatatcccttccccccccttgtaacttagaagacatgcgaatcgaccaatagcaggcaatcaaccagattgtcaccggtcaagcacttctgtttccccggctctcgttgatatgctccaacagggcaaaaacaattggagtcgttacccgcaagatgcctacgcaaaacctagtagcatcttcgaagatttttggttggtcgctcagttgctacccccagcaatagacctggcagatgaggctagaaataccccactggtgacagtgttctagcctgcgtggctgcctgcacacccacacgggtgtgaagccaaagattggacaaggtgtgaagagtcacgtgtgctcatcttgagtcctccggcccctgaatgcggctaaccttaaccccgtagccattgctcgcaatccagcgagtatatggtcgtaatgagtaattacgggatgggaccgactactttgggtgtccgtgtttcacttttacttatcaatttgcttatggtgacaatatatatagatatatattgacacc
496	EV L	acatgggccagcccaccacacccactgggtgtagtagtctggttctatggaacctttctacgcctcttttgcttccctccccccatttctcttcgattgctccacctgtgatctttgcaacttagaagaaataatgaaccccgcacaatagcgggcgctgagccacagcgtcaatgtgcaagcacttctgtttccccggaatgggcccataggctgtacccacggctgaaagggaccggcccgttacccgccttggtactgcgagaatgttagtaactccctcgatagctttaggcgttacgctcagccctttgagcccgaagggtagttcgggtcgatgaggctcgtcattcccccactggcgacagtgtgacttgcctgcgttggcggcccggggtggggggcaaccccccatccacgcctactgaaggacagggtgtgaaggcgctattgcgctactaaggagtcctccggcccctgaatgcggctaacccgaaccccgagcccacggtggtaaacccgccacaagtgggtcgtaatgagtaatttggggcagggaccgactactttgggtgtccgtgtttcctgtttttccatacgatggctgcttatggtgacaaccataagcaattggattggccatccggtgttcatattgcgaat
497	EV K	tcagcctgacgcaagtgcctccattggagtctctccaagccctccggggcttggagggcgccgaccccctgcctaggggagcccacgacacggctggagtccattggcacaccgcagccacgattcaagccagaattgaaagcgggaagcacttctgtctcccggtgtggatcatacgctgtacccacggcgaaaagtgaagcatcgttacccgactcggtacttcgagaagcccagtacagttgtggatctctgcagggtatacgctcagcgtgacccctacgtagttccttgagatggctgagagaacaccccacgggcgaccgtgtctctcggcgcgtggctcaaggccgggccttcagtggctcggtgccttgcagagtgaagcctccgaacagcctattgagctaccgtttagcctccgccctcttgaatgcggctaatcctaaccatggagcgcccgcccacagtccagtgggtagagcgtcgtaacgcgcaagtccgtggcggaaccgactactttagagtggcgtgtttccaatttatccttttataaagttgcttatggtgacaccacaagagatccacgatttcttgtttcttatcactgagacacaagtcatattcatcaatctttattgcggaattaacttggtgcgtccaaacacatcagc
498	EV J 1631	caccctgagggcccacgtggcgtagtactctggtatcaaggtacctttgtacgcctattttatttccctttccccacagtaacttagaagcttatctcatagttcaacagtagggtcactaaccaagtggctcagcgaacaagcacttctgtttccccggtcctagtacctgtgaagctgtacccacggcggaaggggaaaaaagatcgttatccggccccctacttcggaaagcctagtaacaccattgaagcaatcgagtgttgcgctcagcacagtaacccctgtgtagctttggttgatgagtctgggcactccccactggcgacagcggcccaggctgcgttggcggccaaccgactcgggcaaccgggtcggacgctcgtttgtggacatggtgtgaagagcctactgagctagagggtagtcctccggcccctgaatgcggataatcctaaccccggagcacccacactcaatccagagtgcaggatgtcgtaacgcgtaagtctgggacggaaccgactactttgggtgtccgtgtttcctgttttacttacttggctgcttatggtgacaatctagtgttgttaccatatagctattggattggccatccggtgttttgaattgtgtgtttatactaattcttttacatatcacagacaaccaaat
499	EV J N125	cggtacctttgtacgcctattttacccccttcccctttgtaacttagaagcaaagcaaaccagttcaatagtaagcaacacaacccagtgttgtgacgaacaagtacttctgtttccccgggagggtctgacggtaagctgtacccacggctgaagtatgacctaccgttaaccggctacctacttcgagaagtctagtaatacattgaagttttgttggcgttacgctcaacacactaccccgtgtgtagttttggctgatgagtcacggcattccccacgggcgaccgtggccgtggctgcgttgcggccaaccaagggcgcaagctccttggacgtcacttaacagacatggtgtgaagaacctattgagctaggtagtagtcctccggcccctgaatgcggctaatcctaactccggagcacatcagtgcaacccagcatttggtgtgttgtaatacgcaagtctggagcggaaccgactactttgggtgtccgtgtttcctgttttaccttatttggctgcttatggtgacaatttgatattgttaccatatagctgttggattggccatccggatttttgaaagagacccaaaactttcttctctacttcagattcaagtgcgaagttttccttttcatatattacttactaatttgaagtaccaaag

[1730]

500	EV I	ttagtactttctcacggggatagtggtatccctccctagtaatttagaagacttgaaaaaccgaccaat aggcacctcgcatccagcggggtaaaggtcaagcacttctgtttccccgggtcgagtagcgatag actgtgcccacggtcgaaggtgaaacaacccgttatccgactttgtacttcgggaagcctagtacc accaaagattatgcttggggtttcgctcagcacgaccctggtgtagatcaggccgatggatcaccg cattcctcacaggcgactgtggcggtggtcgcgtggcagcctgccgatggggcaacccatcgga cgccaagcatatgacagggtgtgaagagcctactgagctacaaagtattcctccggcccctgaat gcggctaatcccaaccacggagcatttgctaccaaaccaggtagtggaatgtcgtaacgggtaac tctgtggcggaaccgactactttgggtgtccgtgtttccttttaatttatcattctgtatatggtgacaact atagtgctatctcgatttgcattactattgttgagattaaaactttattacattgttgcatttttacccttgag tgagttttcacctgaacagattaatttactcatcctgtttatatattacaagcagaaatacttgcaaag
501	EV F1 BEV 261	gcaatgctgcaccagtgcactggtacgctagtacctttcacggagtagatggtatcccttaccccg gaacctagaagattgcacacaaaccgaccaatagggcgcaccgcatccagccgtgcagcggtca agcacttctgtctccccggtctgtaaagatcgttatccgcccgacccactacgaaaagcctagtaac tggccaagtgaacgcgaagttgcgctccgccacaaccccagtggtagctctggaagatggggct cgcaccacccccgtggtaacacggttgcctgcccgcgtgtgcttccgggttcggtctcgtgccgtt cacttcaacttcacgcaaccagccaagagcctattgtgctgggacggttttcctccggggccgtga atgctgctaatcccaacctccgagcgtgtgcgcacaatccagtgttgctacgtcgtaacgcgtaagt tggaggcggaacagactactttcggtaccccgtgtttcctctcattttatttaatattttatggtgacaat tgttgagatttgcgctcttgcaacgttgccattgaatattggcttatactatttggttgccttttacaaaac ctctgatatacccagttcttacattgatctgcttgtttttctcaatttgaagtatagactacaaatagcaaa
502	EV D94	cgtggcggccagtactctggtatcacggtacctttgtacgcctgttttatatccccttccccgcaact tagaagaaaacaaatcaagttcactaggaggggggtacaaaccagtaccaccacgaacaagcact tctgtttccccggtgatgtcgtatagactgtaaccacggttgaaaacgattgatccgttatccgctctt gtacttcgaaaagcccagtatcaccttggaatcttcgatgcgttgcgctcagcactcaaccccagag tgtagcttaggtcgatgagtctggacactcctcaccggcgacggtggtccaggctgcgttggcgg cctacctgtggtccaaagccacaggacgctagttgtgaacaaggtgtgaagagcctattgagctac aagagaatcctccggcccctgaatgcggctaatcctaaccacggagcaagggtacacaaaccag tgtatatcttgtcgtaacgcgcaagtctgtggcggaaccgactactttgggtgtccgtgtttccttttgt ttttatcatggctgcttatggtgacaatctaagattgttatcatatagctgttggattggccatccggtaa tttattgagatttgagcatttgcttgtttcttcaacaatttcacctattcattgcatttcagcagtcaaa
503	PV3	tacctttgtacgcctgttttatactccctcccccgcaacttagaagcatacaattcaagctcaatagga gggggtgcaagccagcgcctccgtgggcaagcactactgtttccccggtgaggccgcatagact gttcccacggttgaaagtggccgatccgttatccgctcatgtacttcgagaagcctagtatcgctctg gaatcttcgacgcgttgcgctcagcactcaaccccggagtgtagcttgggccgatgagtctggac agtccccactggcgacagtggtccaggctgcgctggcggcccacctgtggcccaaagccacgg gacgctagttgtgaacagggtgtgaagagcctattgagctacatgagagtcctccggcccctgaat gcggctaatcctaaccatggagcaggcagctgcaacccagcagccagcctgtcgtaacgcgcaa gtccgtggcggaaccgactactttgggtgtccgtgtttccttttattcttgaatggctgcttatggtgac aatcatagattgttatcataaagcgagttggattggccatccagtgtgaatcagattaattactcccttg tttgttggatccactcccgaaacgttttactccttaacttattgaaattgtttgaagacaggatttcagtgt caca
504	EV C102	ctttgtacgcctgttttacatcccctccccacgtaacttagaagcaattcaacaagttcaatagagg gggtacaaaccagtatcaccacgaacaagcacttctgtttccccggtgattttacataagctgtgcc cacggctgaaagtgaatgatccgttacccgctcgagtacttcgaaaagcctagtatcgctttgggat cttcgacgcgttgcgctcagcactctaccccgagtgtagcttaggctgatgagtctgggcattcccc atcggcgacgatggcccaggctgcgttggcggcctacccatggctaacgccatgggacgctagt tgtgaacaaggtgtgaagagcctattgagctactcgagagtcctccggccctgaatgcggctaat cccaaccacggatcaggtgcctccaacccaggaggtggcctgtcgtaacgcgcaagtctgtggc ggaaccgactactttgggtgtccgtgtttccttttatcttttaaatggctgcttatggtgacaatcataga ttgttatcataaagcgaattggattggccatccggtgaaatacaaacacattatttacttgtttgttggat ttactccgctcacacagcttactcctaagataatatttattgtattgctggtaaggagacactattata
505	EV 30	aagcaaggcaaacctgaccaatagtaggtgtggcacaccagccgcattttggtcaagcacttctgt ttccccggaccgagtatcaataagctgctcacgcggctgaaggagaaaccgttcgttacccgacc agctacttcgagaaacctagtaacactatgaacgttgcggagtgtttcgttcagcacttcccccgtgt

[1731]

		agatcaggtcgatgagtcaccgcattcctcacgggtgaccgtggcggtggctgcgttggcggcct gcctacgggttcgcccgtaggacgctctaatacc gacatggtgtgaagagtccattgagctagctg gtagtcctccggcccctgaatgcggctaatcctaactgcggagcaggtgctcacagaccagtgag tagcctgtcgtaacgggcaactctgcagcggaaccgactactttgggtgttttccttttttcttctcttа tattggctgcttatggtgacaattaaagaattgttaccatatagctattggattggccatccggtgacg agcagagccattgtttacctctttgttggatttgtacctttgaaccacaaagtcttgaataccattcatct cattttaaagttcaactcagctaaaagaaa
506	SA5	agtacttggtattccggtacctttgtacacctatttacaaaccctacccctttgtaaccttagaagcaatt atttaaccgctcactaggggtgtgctatccaagcacatcaagagcaagcacttctgtctccccgg gaggggctaatggtacgctgtgcccacggcggaaatgagccctaccgttaaccggcagtctactt cgggaagcccagtaactacattgaaactttgaggcgttacactcagcacataaccccaatgtgtagt tctggtcgatgagccttggcatcccccacaggcgactgtggccaaggctgcgttggcggccagc ctgcggaccaaaagtccgtaggacgcctaattgtggacatggtgtgaagagcctactgagctaga ctgtagtcctccggcccctgaatgcggctaatcctaaccctggagcatccgcgtgcaacccagtac gtagggtgtcgtaatgcgtaagtctgggatggaaccgactactttgggtgtccgtgtttcttgtttttca tactgggtcgcttatggttacaactaattgttgtaatcattggcagtgcgcgctgaccacgcgattatt gatatttccatttgttggatactccaatagtgtcaactcatatacacaactttaccactgatcaagataa aa
507	EV A114	tgtgcgcctgttttgaaaccccctcccccaactcgaaacgtagaagtaatgtacactactgatcagta gcaggcgtggcgcaccagccatgtctcgatcaagcacttctgtttccccggactgagtatcaatag actgctcacgcggttgaaggtgaaaacgtccgttacccggctaactacttcgagaaacctagtagc accatagaaactgcagagtgtttcgctcagcacttcccccgtgtagatcaggtcgatgagtcactgc aatccccacgggtgaccgtggcagtggctgcgttggcggcctgcctatggggcaacccatagga cgctctaaggtggacatggtgtgaagagtctattgagctagttagtagtcctccggcccctgaatgc ggctaatcctaactgtggagcgcatactcccaaaccagggagcagtgcgtcgtaacgggcaactc cgcagcggaaccgactactttgggtgtccgtgtttccttttattcctatactggctgcttatggtgacaa ttgagagattgttaccatatagctattggattggccatccagtgtgtaatagagcaatcatttaccaatt tgttggatttactccattaacccacacgtctctcaacacactacatttcatcttactactgaacactaga aa
508	모보바이러스 A	tattctcccacaaaccttcttgtaactctgttaagccttttacatccatgtaatttaattttctccacctaaa aggatttcccccatggtcctttttggctcgaacaaatgctacataggg tcttgttttctcccccctggctc tcttgccagggttccatacccc aattcctctatttccatgatttttcatcatggtttatttttactgtcttctta ttttctgaggtgaccaactcctaagccgactgggtcgcggaagcccggactcctcgcatcactagg gtgcgtagcgatgtaggcgaaaatattggttgctagatgcatacatatagtgaattgatactacacca aactctgttctttttgaaactagctattttctaagtaaggtaggctacgggtgaaaccttaccattgcag gtacgtgaaccgcaacggacatttggccgaagactggtgtacccacgtcagttataggacctcttc aacgttggtggacggcatgtcactgattagttaggctagtgaatttaagttcaggggtatcttttagc ttaagcgtgtattctagtaggacttgcagagcctccccacctaggaggatctctgtttatagccccttt tccttgttccgttagttttccacacttttacaatatttgatgatttgtt
509	부르펜가리 바이러스	ctccccccccttccccttcccgagtaggagattggcatgtatgctctacatgcccgattctctcttgct cactctcttaaatcctggtggcggtctcggattaaacatttatgtcgtatctgggatcgtcttacttggt ggtaattcctctgttgcctagggacctccggactgccggattaaaggtctcaacagagggcaatgt acaaggaagtcattatacgctaattaagtatttgatgaatgactagtgtgacagggctgaggaactc ccccccgggtaaccggtgcctcagcgtccgaaagacacgtggataggatccaccctgttatаccca gcacgatgtaatagtcaaatacctctgatttgtgtaggatgtataaattgtgcattgtaaattttgggcg tagagatgctccgaaggtaccccgttttacgggatctgatcggaggctaattacccaatgcgcccta aataacttcatataatttctttttctttattcaaa
510	허니바이러스 A1	taacgtttggcaagaaccctcacctgtcaattgggaccaccactttcagtgaccccatgcgaagtag tgagagagaataagctttcttaccctcatttgtgaaccctcagtcgaagccgcttggaataagata ggaggaaaagttcattctaaatggagtgaaacatgtacttcagaatttctagcacgcgctgggctttc ttgcgtgtgacggcactgtcttgccggagctctccacactgacaccccacgcttgtggaccttggtg gcagatgacaacactgcagctggaattgagtgtctggtacactctgtgtaacagtgaaaacaatgt gatcacttcggtgagctagtagcctgtggaccaacaactggtaacagttgcctcaggggccaaaa gccacggtgtttacagcaccctactggtttgattggagcaatccaagatgtcacagagttagtaattg

[1732]

		ccaagcagtccgtactggtatcttgacataccgtgcagttttggatagtgaaggatgccctgacggt acccataggtaacaagtgacactatggatctaagcaggggctcactctacgctgctttacagctgg ctgtgagttaaaaaacgtctagctatccacaacctaggggactaggttttccttttatttagattacaatt at
511	허니바이러스 A2	acagtttttgacaaggaccctcacctgtcaatcgggaccaccactttcagtgaccccgtgcgaagt gttgagagaaagtgagctttcttacccttcatttgtgaacccttcagtcgaagccgcttggaataagat ggaaggaaatgttcattctaaatggagtgaaacatatacttaatttccagtgtttagtggtctttccact agacaacggcactgtcttgccggaactctacacaccaacattccacgcttgtgggactcaaatgttg gatgacacagttgtagctggaactgagtgtttagtgcactctgtgtaacagtgaaaacaatgtgatca cttcggtgggctagtagcctgtggactaacaactggtaacagttgcctcaggggccaaaagccac ggtgttaacagcaccctactagtttgattggagcaatccatgatgttacagagttagtaactgccaaa cagattgtactggtatcttggcataccgtgcaacttaggatagtgaaggatgccctggcggtaccca taggtaacaagtgacactatggatctaaacaggggctcactctacgttgctttacaactagctgtga gttaaaaaacgtctaactatccacaacctaggggactaggttttccttttatttttatacacaacta
512	Ia Io	ggtgtccgggtgtgtgggatagacccagatgtgcagtggatgcgagcatttgagtcagagtagga gcaagcccaggggcaaagggaccacattgtgtatcccgaatgaaggatcgagatttctctcctcat tacccggtgtcttgtcactgttgggggggcccaacagtcttagtcctatactgcctgatagggtcgc ggctggccggactcaagtgctatagtcagttgattttcactc
513	타우라 증후군 바이러스	ctttaaaagtcgtgcgtggcttcaccacgcacgatcagtactatcagttaaccactcttgaatatgctc aatgaccctattcaacactggtgtctcttagtacattattttagcacttaacgtgcatgagttttgcccat ttctttcaaaaaaatgagtattcgaggagacgtcccgctccccgtcttatttcaaccgtagactcgac atctattggtggacatttaattccagtcgccgtaagttgcttctgccccgcgctatatttttcttatacttat ggttctataggtctggtttaaaacgtaaatagacggcccacaaactatagaacgcgtacccggaac gccaatcccggataagtccctggatatatagatgcaccgcaatataagcctgcagactgtctcatat act
514	ABPV	cccgtcaaaataacaacttataacacgatgttacccgaagaaaccattttagtgtaacatttaagatta gaagtagttcatctaatagagataggcactattagaaggaggccttttctaaaggagccgttagtca gcccagacaagcgcagtactttagaagagagaagttccccgatagcgaccgaaaagacgcgtttt ccgtgctaactaatttaaatgtgggaacgaatattattattgaaattatgtgagccacgtagcaatcaa gtcatgtttttgtcactacgtttactcatctaatgtagataatttttgtttaagtacctatttaggtgtcatccc accagagaagaaataatacgtaccggaacccagagtacacccctttatttttaagccttactgggcttc tctgttagttagtaatctggcccacgttttgcgttgagtggggtcccaacagtaggaattcgacggac aagtagcaagcgagtcggtaccaattggtttagccttcgaaattactctgggcaggaagttactaaa cgagaactttctgcttaaatcccaacgcacaaacaaatagagtaaataaataattata
515	BRAV-2	tttgttttgcggctttgccgttgttcgggttttacctgttttcacacagcaaaacaggccttctagtttcgt gcttaaacgagatcatgctcgaactagaactacatagctggtcactggactcataccacaccttgtg gagctttatgggaaaggtggctagtgggctgtggaagtgactctgaccacatgcctctcaagtgtg ggaaatcacggatcggtgtagcgacgacaacaggccttgggacaccctctccagtaatggagac ccaaggggccaaaagccacgcctcgtgccctgttgttcacaaccccagtgcgacccgtgttagta cctatttgcgagaactgtgtctggacagctaaacacaaccctagtgggagactaaggatgcccag gaggtacccggaggtaacaagtgacactctggatctgacctggggagagagggcttgctttacag gcgcctctctttaaaaagcttctatgtctcatcaggcaccggaggccgggccttttccttttaaaatta cactta
516	BRBV-1	ccccccctacttaaagatgtacggttttgctgctttcacagagtaaagcagatagaggttctgaactg gcaaactttacctcgaaacacgcccgtttttctgctgtgtctcacagactgtcctgtcacacttgtggc ggcttgtgacactgtgaacatagtgagaccgaccaagacaacagtttcaagtgatgaacatcgaac gtctaaactggatccgtaactggacatgttagggcaaggacttccccccctggtaacaggagcctgg ctggccaaaagccccgctcattgagcctagcatgttgtcgaccctggactgttcagtttagttagtac atggaattcacttgtcacggttcttctgaactcggtctctagtatgacagcctaaggatgccctccag gtaccccggggtaacaagtgacacccgggatctgaggaggggactactttacgtagtttaaaaaa cgtctaagctgttatggtgaccagaggctggcacctttcacttttaaaattacactactgactacaatt gaagtgataacggttttacaggctttcaaactagttacacaagcactgttttcctgacacacacacttt
517	ERAV-1 U188	aatattggcgcgcgcatttgcgcgccccccccatttcagccccctgtcattgactggtcgaaggc gttcgcaataagactggtcgtcacttggctgttctatcgtttcaggctttagcgcgccctcgcgcggc

[1733]

		gggccgtcaagcccgtgcgctgtatagcgccaggtaaccggacagcggcttgctggattttcccg gtgccattgctctggatggtgtcaccaagctgacaaatgcggactgaacctcacaaagcgacaca cctgtggtagcgctgcccaaaagggagcggaactcccccgccgcgaggcggtcctctctggcc aaaagcccagcgttaatagcgccttttgggatgcaggagccccacctgccaggtgtgaagtggag tgagtggatctccaatttggtctgttctgaactacaccatctactgctgtgaagaatgccctggaggc aagctggttacagccctgaccaggggccctgcccgtgactctcgatcggcgcagggtcaaaaatt gtctaagcagcaggaacgcgggagcgtttctttttcccttttgtatcgac
518	GFTV	atggggaagggtatgacgtgcccctttccttcttcggagaaactcgctctagtggtctttccacttctgg aaaagagtgagtgcacgtgatcaggaccgtcgaagacgacaaatacctggtgctctatctcatag acgtttcacagctgtagcgacccctcagtagcagcggaagccccctcctggtgacaggagcctct gcggccaaaagccacgtggataagatccactgctgagggcggtgcgaccctagcacccctgtgat gcatactagttgtagcgtgccggactattggtctgtcataagacacctgatagagagaccaagaat gtcctggaggtaccccgcgtgcgggatctgaccaggagaccattgcccaatgctttacaacgggt ctatggtttaaaaactgtcgcagtctctccaaaccaagtggtcttggttttcaattactttgaatatttca ct
519	SAFV V13C	ttttcgacgtggttggaattgccatcatttccgacgaaagtgctatcatgcctccccgattatgtgatgt tttctgccctgctgggcggagcattctcgggttgagaaaccttgaatctttttctttggaaccttggttc ccccggtctaagccgcttggaatatgacagggttattttcttgatcttatttctacttttgcgggttctatc cgtaaaaagggtacgtgctgcccctttccttctctggagaattcacacggcggtctttccgtctctcaa caagtgtgaatgcagcatgccggaaacggtgaagaaaacagttttctgtggaaatttagagtgcac atcgaaacagctgtagcgacctcacagtagcagcggactcccctcttggcgacaagagcctctgc ggccaaaagccccgtggataagatccactgctgtgagcggtgcaaccccagcaccctggttcgat gatcattctctatggaaccagaaaatggttttctcaagccctccggtagagaagccaagaatgtcct gaaggtaccccgcgtgcgggatctgatcaggagaccaattggcggtgctttacactgtcactttgg tttaaaaattgtcacagcttctccaaaccaagtggtcttggttttccaattttgttgaatggcaat
520	SAV P-113	ggagatctaagtcaaccgactccgacgaaactaccatcatgcctccccgattatgtgatgctttctg ccctgctgggtggagcatcctcgggttgagaaaaccttcttccttttttccttggaccccggtccccg gtctaagccgcttggaataagacagggttatcttcacctcttccttcttctacttcatagtgttctatact atgaaagggtatgtgtcgccccttccttctttggagaacacgcgcggcggtctttccgtctctcgaaa agcgcgtgtgcgacatgcagagaaccgtgaagaaagcagtttgcggactagctttagtgcccaca agaaaacagctgtagcgaccacacaaaggcagcggaccccccctcctggcaacaggagcctct gcggccaaaagccacgtggataagatccaccttgtgtgcggcacaaccccagtgccctggtttct tggtgacacttcagtgaaaacgcaaatggcgatctgaagcgcctctgtaggaaagccaagaatgt ccaggaggtacccctcccctcgggaagggatctgacctggagacacatcacatgtgctttacacct gtgcttgtgtttaaaaattgtcacagctttcccaaaccaagtggtcttggttttcactctttaaactgatttt cact
521	VHEV	aattccttcttcctttctccttggacctcggtcccccggtctaagccgctcggaatatgacagggttatt ttcacctcttctctcttctacttcatagtgttctatactatgaaagggtatgtgtcgcccctttccttcttgga gaacgtgcgtggcggtctttccgtctctcgaaaaacgtgcgtgcgacatgcagagtaacgcaaag aaagcagttcttggtctagctctggtgcccacaagaaaacagctgtagcgaccacacaaaggcag cggaaaccccctcctggtaacaggagcctctgcggccaaaagccacgtggataagatccacctt gtgtgcggtgcaaccccagcaccctggtttcttggtgacacacttagtgaaccctcgaatggcaatct caagcgcctctgtaggaaagccaagaatgtccaggaggtacccccttcctcatggagggatctgac ctggagacacatcacacgtgctatacacttgtgcttgtgtttaaaaattgtcacagctttcccaaacca agtggtcttggttttcccttaacttcgaaaagtcactatggcctgcaaacatggatacccagacgtgt gccct
522	TRV NGS910	atgcgacgtggttggagattaaaccgactccgacgaaagtgctatcatgcctccccgattatgtgat gttttctgccctgctgggcggagcattctcgggttgatacaccttgaatccttcatccttggacctcag gtcccccggtctaagccgcttggaatacgacagggttattttccaatcttctcctttctacttttcatgag tcctattcatgaaaagggtctgtgctgccccttccttcttggagaatctgcgcggcggtctttccgtct ctcgaaaagcgcagatgcagcatgctggaaccggtgaagaaaacagttctttgtggaaacttaga gcagacatcgaaacagctgtagtgacctcacagtagcagcggaaccccctcctggtaacaggag cctctgcggccaaaagccccgtggataagatccactgctgtgagcggtgcaaccccagcaccct ggttcgatggttgttctctgtggaaccagagaatggtctttctcaagccctccagtagagaagccaa

[1734]

		gaatgtcctgaaggtaccccgcatgcgggatctgatcaggagaccaatcgtcagtgctttacactg gcgctttggtttaaaaactgtcacagcttctccaaaccaagtggtcttggttttcacttttatcaaactgt ttc
523	EMCV2 RD1338	aaatactggtcgaaaccgcttgggataagaccggggtttgttaatgtctcaatgttattctccaccca attgacgtcttttgtcaattggagggcagtgaaaccttgcccttgcttcttgcagaggattcccagtgg tctttccgctctcgacaagggaattcatgatccaccaaaagttgtgaagagagcaggtcccatgga agctttctgacgactgatgatgactgtagcgaccctttgcaggcagcggaccccccacctggtga caggtgcctctgcggccaaaagccacgtgtttaacagacacctgcaaaggcggcacaaccccag tgcctcatcaaaagtctgatgactgtggaaatagtcaaccggcttttcttaagcaaatttggtgtcgg ggctgaaggatgcccggaaggtaccacactggttgtgatctgatccggggccacagtacatgtgc tttacacatgtagctgcggttaaaaaacgtctaggcccccgaaccacggggacgtggttttccttt gaaaaccacgattacaat
524	EMCV1 JZ1203	gtctgctcgatatcgcaggctgggtccgtgactacccactccccctttcaacgtgaaggctacgata gtgccagggcgggtactgccgtaagtgccaccccaaaacaacaacaaaccccccctaacattact ggccgacgccgcttggaataaggccggtgtgcgtttgtctatatgttatttcccaccacattgccgtc ttttggcaatgtgtgggcccggaaacctggccctgtcttcttgacgagcattcctaggggtctttccc ctctcgccaaaggaatgcaaggtctgttgaatgtcgtgaaggaagcagttcctctggaagcttcttg aagacaaacaacgtctgtagcggcccttgcaggcagcggaaccccccacctggcgacaggtg cctctgcggccaaaagccacgtgtataagatacacctgcaaaggcggcacaaccccagtgccac gttgtgagttggatagttgtggaaagagtcaaatggctctcctcaagcgtattcaacaaggggctga aggatgcccagaaggtaccccattgtatgggatctgatctggggcctcggtgcacatgctttacatg tgtttagtcgaggttaaaaaacgtctaggcccccgaaccacggggacgtggttttccttgaaaaa cacgatgataat
525	EMCV1 AnrB-3741	atgtggtcgaagccacttggaataagaccggcgtgcgcttgtctatatgttacttccaccacattgcc gtcttttggcaatgtgagggcccggaacctggccctgtcttcttgacgaacattcctaggggactttc ccctctcgccaaaggaatgtaaggtctgttgaatgtcgtgaaggaagcagttcctctggaagcttctt gaagacaaacagcgtctgtagcgacccttgcaggcagcggaaccccccacctggtaacaggtg cctctgcggccgaaagccacgtgtataagatacacctgcaaaggcggcacaaccccagtgccac gttgtgcgttggatagttgtggaaagagtcaaatggctttccccaagcgtattcaacaaggggctga aggatgcccagaaggtaccccactggttgggatctgatctggggcctcggtgcaggtgctttacac ctgttgagtcgaggttaaaaaacgtctaggcccccgaaccacggggacgtggttttcctagaaaa ccactatgacaat
526	코사바이러스 D1	cgtgctttacacggtttttgaaccccacaccggctgtttggcgcttgcaggacagtaggtatattttctt tcatttctctttctagccgcgtaggttctatctacgcgggcggagtgatactcccgctccttcttggac aggcggcctccacgctctttgtggatcttaaggctgccaagtcactggtgtttgaagtgaagaatgg agagacactagggcgtttcatgtggctttgccagggattgtagcgatgctgtgtgtgtgtgcggattt cccctcgtggcgacacgagcctcacaggccaaaagccctgtccgaaaggacccacacagtggg gttgccccgacccctcccttcaaagctttgtgtaaacaaactttgtttagactttcttaagcttctctca catcaggccccaaagatgtcctgaaggtaccctgtgtatctgaggatgagcaccaccaactaccc ggacttgtgggacgtgtcccacagacgcatgtggtattccagccccctccttttgaggaggggct tttgctcgctcagcacaggatctgatcaggagattcatctctggtgctttacaccagagcatggattta aaaattgcccaaggcctggcaaacaacctaggggactaggttttctctattttaaaagatgtcaat
527	코사바이러스 B1	cgtgctttacacggttttgaaccccacaccggctgtttggcgcttgcaggacagcaggtttatttttctt tttaactctctctttctagccacacacgatctatgtgtgtgggcggagtgatactcccgttccttcttgga caggcggcctccacgcccttttgtggatcttaaggctaccaagtcactggtgttggaaagtgaagag aaaggagttccttgggaactacatgtggcattgacagaggttgtagcgatgctgtgtgtgtgtgcgg attaccccccgtggcgacacggacccccacaggccaaaagccctgtccgaaaggacccacacagt ggagcaaccccagctcccctcttcaatgttttgtgttagcaaccttggtattatttttctctcaagcttcca atacaccgggccccaaagatgtcctgaaggtaccccgtgtatctgaggatgagcaccatcaacta cccggacttgttctttcgagaacagacgcatgtggtaacccagccccgatcctaaggggtcggggg cttttgctcactcagcacaggatctgatcaggagacctccccccccctgctttacaggggcggggg tttaaaaattgcccaaggcctggcaaataacctaggggactaggttttccttttattttaaagttgtca at
528	코사바이러스 A	ccgtgctttacacggtttttgaaccccacaccggctgtttggcgcttgcaggacagcaggtttatttttc

[1735]

	SH1	ttatgctctttatttctagccaacagggttctatcctgttgggcggagtgatactcccgttccttcttggacagattgcctccacgatctttgtggatctcaaggtgatcaagtcactggtaaatagagcgaaggttgaggaaacctgaggaatttccatgtggttttgccaggagttgtagcgatgctgtgtgtgtgtgcggatttcccctcatggcaacatgagcctcacaggccaaaagccctgtccgaaaggacccacacagtggagcaatcccagctcccctcctacaaagctttgtgagaatgaactcacgtttattcttctttattctctgtttacatcaggccccaaagatgtcctgaaggtaccttgtgtatctgggcatgagcaccatcaactacccggacttgcatttcggtgcagacacatgtggttacccagccccctctgctttggcagaggggcttttgctcgctcagcacgagatctgatcaggagccctcccagtgtgctttacacctggcgggggttaaaaattgcccaaggcctggcaaaataacctaggggactaggttttccttttattaacaatgtctgtcatt
529	말라가시바이러스 B	ctttattttcttatgtaactcttcttttaagttttatttgcctacttgtgagcttatgcgggaccactgtcttagacaaccccacatttgtcatgagtaagtacacgcaaccattacgattactttttaaccgtctgacctttgataacaactgaagttaggcgtgaaacatgcatttatataccaaagtagccccgcatttccccactacggtgggggggctaccctactggctttggaactgtagccattatgtgttgcctggctttcaggatctcaaacacaacagttctctcacaatggaatatgggtgagattgcagtgacatgaacaagtatctagtagtacatagactcaagcctagttgcctgcggaacaacatgtggtaacacatgccccagggtccaaaagacaagggttaacagccccttctaggtgtctgtgtgtgaagaatactttagtagtgttgttatgatctcactgttagtacagaatgagtatggcttggtgaaggatgtcctacaggtacccattatatggatctgagtaggagaccactagtggtggctttaccgccaggtgagtggtttaaaaagcgtctagccaagccaacagcactagggatagtgctttctatattttatattttcagtgtat
530	모사바이러스 A2 SZAL6	ccccccctcaaattgcaacgatatagctaatggcgagattgagatgctatatcacctcttctaagttatagacctcatctgattgataaggacgtaatttggtcgaaaccgcttggaataagaccgatgcgcgtagtcatgatgatgatgtaagatctaggaacttatccaatctgcttatgtctatgtaagtagaggggcaggcctcattgccctaattctttctaccgagtatctgctagggtttctagcggcttgaatacttggattgagggatacaagatactactgatcgattgtcgattgggaaacttgtagatacttcaaagctaccagtagcgtggactcacagccagcggactacccctcatggtaacatgagcctctgggcccacaaggcacgtcgcaagacctgtgagacggcaaccccagcctagctttgttgaggaaacaagcgataacatgacatgagagaccggaaggattcttgtattgtgagccgaaggatggcctctaggtacctcatttttatgagatctgaggaggtgctcttgagttggtgctttacactgacaacacagagttaaaaagcgtctaagctcaccccggaaattgggaaatttccgttatttccatttgtttgcaaagtcgttc
531	SVV	ctgggccctcatgcccagtccttcctttccccttccgggggggtaaaccggctgtgtttgctagaggcacagaggagcaacatccaacctgcttttgtggggaacggtgcggctccaattcctgcgtcgccaaaggtgttagcgcacccaaacggcgcatctaccaatgctattggtgtggtctgcgagttctagcctactcgtttctcccctatccactcactcacgcacaaaaagtgtgctgtaattacaagatttagccctcgcacgagatgtgcgataaccgcaagattgactcaagcgcggaaagcgctgtaaccacatgctgttagtcccttcatggctgcgagatggctatccacctcggatcactgaactggagctcgaccctccttagtaaggaaccgagaggccttcttgcaacaagctccgacacagagtccacgtgattgctaccaccatgagtacatggttctcccctctcgacccaggacttcttttgaatatccacggctcgatccagagggtggggcatgatccccctagcatagcgagctacagcgggaactgtagctaggccttagcgtgccttggatactgcctgataggcgacggcctagtcgtgtcggttctataggtagcacatacaaat
532	PTV A	actagtccttggacttttgttgtgtttaaacacagaaatttaattacctggccatgaattcattggattaacccttctgaaagacttgctctggcgcgagctaaagcgcaattgtcaccaggtattgcaccagtggtggcgacagggtacagaagagcaagtactcctgaccgggcaatgggactgcattgcatatcccctaggcacctattgagatttctctggggcccaccggcgtggagttcctgtatgggaatgcaggactggacttgtgctgcctgacagggtcgcggctggccgtctgtactttgtatagtcagttgaaactcacc
533	PTV B	cttccttttaattcgtaactgataagtgatagtccttggaagctaggtatttgttacgctagttttggattatcttgtgcccaacatttgttttcgaacatatgttgtgtttaaacacagaaatctagtttcttggttatgagtttaatggaatatccttttgaaagacttgccttggcgcgggctagagcgcaattgtcaccaggtattgcaccaatggtggcgacagggtacagaagagcaagtactcctgactgggtaatgggactgcattgcatatccctaggcatctattgagatttctctggagcccaccagcatggagttcctgtatgggaatgcaggactggacttgtgctgcctgacagggtcgcggctggccgtctgtactttgtatagtcagttgaaactcatt
534	토트오리바이러스	ccccttacgtaactgcaacttaaagagtaccctactgcattggatgtgtggtaaactttacgcacacatttgtagtagtgttagttatgttctacctaatgagtatgcatgcacccgtcgaaacacgcttgtgataagataggtgagtccatgtgactaatctcattaagataaataagcaccctacaacgcacggcacgctc

[1736]

		gtgtcttccgtgcggggccgggacaacagcggcctaaatcttctaggtgaccaccatgcttttggg actatggcaccactgtggacgtgagtacctggcagtaagtctgtgaaaagatggaaggtgtccca agctatggggcgtatgcatatagcctgcggaacaaacaacggcgacgttgtccccagggcccaa aaggcacgtggataagatccacctatatgtttaccccatagtgtaagtcactggaagtcctagtaatg gatgtctggagtaaggctcacggggtagggcgaaggatgcccagaaggtacccgtaggtaacct taagagactatggatctgatctggggaccggatggcgccatcaccatgacgtggaggccggttta aaaaacgtctaagcccgaccaacaacctaggggactaggttttccttttttattcatgtatgacgtt
535	포사바이러스 1	acatttccttgcgtgcgcacccgaaaatttattgaacttggcttgaatcataagtaatgcttcttatagc ggacactttgagaatataatgcttgtatggattaatgtatactgttttaaataacaaacctagacacgc agttgcgttgatggttgtatcaatcacatataagtgttgaactcgtgttaatcctcgatcgctatatgttt gccgcctacttccaataaaatagattacatgcgcgtcatgcctttgtgggttacctattggcctctgac aaaaacaagtcgtaagatttgtagcttcccggtgtaaaaagctgggcgcggtctggctctcgtagg gtggaaaggtccaccaatggctggttgagtgtaagctccggtgtcctggttgtcgcaattccaggc gtcgtaataacctatattgcatctgactctaactcttgtggctctactgtatctagttcttgttctactaact ctaataatactactggctctaatactgaaaaacttacatatgttaatatagataatatccttgatcctgat atccctcacgtcactgaagttcgccgaaaacgaatttcagatcatatcattgaatctcaaggatgtac ttgctctgaacctactataactcctcatgcgttttcattttctactcttggc
536	A105-675	ccacccacagcaagaatgccatcatctgtcctcacccccatttctcccctccttccccctgcaaccatt acgcttactcgcatgtgcattgagtggtgcacgtgttgaacaaacagctacactcacgtgggggcg ggttttcccgcccttcggcctctcgcgaggcccacccttcccctttcctcccataactacagtgctttg gtaggtaagcatcctgatccccccgcggaagctgctcgcgtggcaactgtggggacccagacagg ttatcaaaggcacccggtctttccgcctccaggagtatccctgctagtgaattctagtggggctctgc ttggtgccaacctcccccaaatgcgcgctgcgggagtgctcttccccaactcaccctagtatcctct catgtgtgtgcttggtcagcatatctgagacgatgttccgctgtcccagaccagtccagcaatggac gggccagtgtgcgtagtcgtcttccggcttgtccggcgcatgtttggtgaaccggtggggtaaggt tggtgtgcccaacgcccgtactttggtgacaactcaagaccacccaggaatgccagggaggtacc ccgcctcacggcgggatctgaccctgggctaattgtctacggtggttcttcttgcttccatttctttctt ctgttc
537	A110-675	accttgtgcgcctgttttataccccctccccaactgtaacttagaagtaacacacaccgatcaaca gtcagcgtggcacaccagccacgttttgatcaagcacttctgttaccccggactgagtatcaatagа ctgctcacgcggttgaaggagaaagcgttcgttatccggccaactacttcgaaaaacctagtaaca ccgtggaagttgcagagtgtttcgctcagcactaccccagtgtagatcaggtcgatgagtcaccgc attccccacgggcgaccgtggcggtggctgcgttggcggcctgcccatggggaaacccatggg acgctctaatacagacatggtgcgaagagtctattgagctagttggtagtcctccggcccctgaatg cggctaatcctaactgcggagcacacaccctcaagccagagggcagtgtgtcgtaacgggcaac tctgcagcggaaccgactactttgggtgtccgtgtttcattttattcctatactggctgcttatggtgac aattgagagatcgttaccatatagctattggattggccatccggtgactaatagagctattatatatcc ctttgttgggtttataccacttagcttgaaagaggttaaaacattacaattcattgttaagttgaatacag caaa
538	18-675	cccacagcaagaatgccatcatctgtcctcacccccaattttccctttcttccccctgcaaccattacg cttactcgcatgtgcattgagtggtgcatgtgttgaacaaacagctacactcacatgggggcgggtt ttcccgccctacggcctctcgcgaggcccacccctttccctcccctttataactacagtgctttggtag gtaagcatcctgatcccccgcggaagctgctcacgtggcaactgtggggacccagacaggttatc aaaggcacccggtctttccgccttcaggagtatccctactagtgaattctagcggggctctgcttggt gccaacctcccccaaatgcgcgctgcgggagtgctcttccccaactcaccctagtatcctctcatgt gtgtgcttggtcagcatatctgagacgatgttccgctgtcccagaccagtccagtaatggacgggc cagtgcgtgtagtcgtcttccggcttgtccggggcatgtttggtgaaccggtggggtaaggttggtg tgcccaacgcccgtactttggtgacacctcaagaccacccaggaatgccagggaggtaccccac ctcacggtgggatctgaccctgggctaattgtctacggtggttcttcttgcttccacttctttctctgttc acg
539	A115-675	acctttgtgcgcctgttttataccccccccaacctcgaaacttagaagtaaagcaaacccgatcaata gcaggtgcggcgcaccagtcgcatcttgatcaagcacttctgtaaccccggaccgagtatcaata gactgctcacgcggttgaaggagaaaacgttcgttacccggctaactacttcgagaaacccagta gcatcatgaaagttgcagagtgtttcgctcagcactacccccgtgtagatcaggccgatgagtcac

[1737]

		cgcacttccccacgggcgaccgtggcggtggctgcgttggcggcctgcctatggggcaacccat aggacgctctaatacggacatggtgcgaagagtctattgagctagttagtagtcctccggcccctg aatgcggctaatcctaactgcggagcacataccctaatccaaagggcagtgtgtcgtaacgggta actctgcagcggaaccgactacttgggtgtccgtgtttccttttaatttttactggctgcttatggtgac aattgaggaattgttgccatatagctattggattggccatccggtgactaacagagctattgtgttcca atttgttggatttaccccgctcacactcacagtcgtaagaaccctcattacgtgttatttctcaactcaa gaaa
540	A73-675	ttactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatactccccccaccc ccctttttgtaactaagtatgtgtgctcgtgatcttgactcccacggaacggaccgatccgttggtgaa caaacagctaggtccacatcctcccttccccctgggagggcccccgcccctcccacatcctcccccc agcctgacgtatcacaggctgtgtgaagcccccgcgaaagctgctcacgtggcaattgtgggtcc ccccttcatcaagacaccaggtctttcctccttaaggctagccccggcgtgtgaattcacgttgggc aactagtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaaccct ggcccttcactatgtgcctggcaagcatatctgagaaggtgttccgctgtggctgccaacctggtga caggtgccccagtgtgcgtaaccttcttccgtctccggacggtagtgattggttaagatttggtgtaa ggttcatgtgccaacgccctgtgcgggatgaaacctctactgccctaggaatgccaggcaggtac cccacctccgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaatccttttatgt cggagtc
541	코브바이러스 16317	ttactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatacaccccatccc ctttctgcaacttaagtatgtgtgctcgtgatcttgactcccacggaatggatcgatccgctggagaa caaactgctagatccacatcctcccttccccctgggaggaccttggtcctcccacatcctcccccag cctgacgtaccacaggctgtgtgaagcccccgcgaaagctgctcacgtggcaattgtgggtcccc ccttcatcaagacaccaggtctttcctccttaaggctagccccgatgtgtgaattcacattgggcaac tagtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaaccctggc ccttcactatgtgcctggcaagcatacctgagaaggtgttccgctgtggctgccagcctggtaaca ggtgccccagtgtgcgtaaccttcttccgtctccggacggtagtgattggttaagatttggtgtaaggt ccatgtgccaacgccctgtgcgggatgaaacctctactgccctaggaatgccaggcaggtacccc accccgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaattcttttatgtcg gagtc
542	아이치바이러스 Chshc7	ttactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatacatccccatcccc tttctgtaacttaagtatgtgtgcttgtaatcttgactcccacggaatggatcgatccgctggagaaca aactgctagatccacatcctcccttccccctgggaggaccttggtcctcccacatcctcccccagcc tgacgtaccacaggctgtgtgaagcccccgcgaaagctgctcacgtggcaattgtgggtcccccc ttcatcaagacaccaggtctttcctccttaaggctagccccgatgtgtgaattcacattgggcaacta gtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaaccctggcc cttcactatgtgcctggcaagcatatctgagaaggtgttccgctgtggctgccagcctggtaacagg tgccccagtgtgcgtaaccttcttccgtctccggacggtagtgattggttaagatttggtgtaaggttc atgtgccaacgccctgtgcgggatgaaatctctactgccctaggaatgccaggcaggtaccccac cctcgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaatccttttatgtcggga gtc
543	아이치바이러스 고이아니아	actccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatacaccccatccccct tttttgcaacttaagtatgtgtgctcgtaatcttgactcccacggaatggatcgatccgctggagaaca aactgctagatccacatcctccctccccctgggaggacctcggtcctcccacatcctcccccagc cctgacgtatcacaggctgtgtgaagcccccgcgaaagctgctcacgtggcaattgtgggtcccc ccttcatcaagacaccaggtctttcctccttaaggctagtcccgatgtgtgaattcacatcgggcaac tagtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaaccctggc ccttcactatgtgcctggcaagcatatctgagaaggcgttccgctgtggctgccagcctggtaaca ggtgccccagtgtgcgtaaccttcttccgtccccggacggtagtgattggttaagacttggcgtaag gttcatgtgccaacgccctgtgcgggatgaaacctctactgccctaggaatgccaggcaggtacc ccaccttcgggtgggatctgagcctgggctaattgtctacgggtagtttcatttctaattctttcatgtc ggagtc
544	아이치바이러스 ETHP4	ttactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatacaccccaccccc cttttgcaacttaagtatgtgtgctcgtgatcttgactcccacggaatggatcgatccgctggagaa caaactgctagatccacatcctcccttccctgggaggacctcggtcctcccacatcctcccccag

[1738]

		cctgacgtaccacaggctgtgtgaagcccccgcgaaagccgctcacgtggcaattgtgggtccc cccttcattaagacaccaggtctttcctccttaaggctagtcccgatgtgtgaattcacattgggcaac tagtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaaccccctggc ccttcactatgtgcctggcaagcatatctgagaaggtgttccgctgtggctgccagcctggtaacag gtgccccagtgtgcgtaaccttcttccgtcttcggacggtagtgattggttaagatttggcgtaaggtt catgtgccaacgccctgtgcgggatgaaacctctactaccctaggaatgccaggcaggtaccccca ccctcgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaattcttctatgtcgga agtc
545	아이치바이러스 DVI2169	tactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccatacacccccacccccc ttttctgcaacttaagtatgtgtgctcgtaatcttgactcccacggaatggatcgatccgctggagaac aaactgctagatccacatcctcccttccctgggaggaccccggtcctcccacatcctcccccccag cctgacgtatcacaggctgtgtgaagtccccgcgaaagctgctcacgtggcaattgtgggtccccc cttcatcaagacaccaggtctttcctccttaaggctagccccgatgtgtgaattcacattgggcaact agtggtgtcactgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaacccctggc ccttcactatgtgcctggcaagcatatctgagaaggtgttccgctgtggctgccagcctggtaacag gtgccccagtgtgcgtaaccttcttccgtctccggacggtagtgattggttaagatttggtgtaaggtt catgtgccaacgccctgtgcgggatgaaacctctactgccctaggaatgccaggcaggtaccccca ccttcgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaattcttttatgtcgga gtc
546	아이치바이러스 DVI2321	gcttcttcggaacctgttcggaggaattaaacgggcacccatacacccccacccccttttctgcaac ttaagtatgtgtgctcgtaatcttgactcccacggaatggatcgatccgctggagaacaaactgcta gatccacatcctcccttcccctgggaggaccccggtcctcccacatcctcccccccagcctgacgta tcacaggctgtgtgaagtccccgcgaaagctgctcacgtggcaattgtgggtcccccccttcatcaa gacaccaggtctttcctccttaaggctagccccgatgtgtgaattcacattgggcaactagtggtgtc actgtgcgctcccaatctcggccgcggagtgctgttccccaagccaaacccctggcccttcactat gtgcctggcaagcatatctgagaaggtgttccgctgtggctgccagcctggtaacaggtgccccca gtgtgcgtaaccttcttccgtctccggacggtagtgattggttaagatttggtgtaaggttcatgtgcc aacgccctgtgcgggatgaaacctctactgccctaggaatgccaggcaggtaccccacctcgggg tgggatctgagcctgggctaattgtctacgggtagtttcatttccaattcttttatgtcggagtc
547	아이치바이러스 rat08	tactccattcagcttcttcggaacctgttcggaggaattaaacgggcacccactttcctgtcctctccc cttttctgtaactccaagtgtgtgctcgtaatcttgactcccgcggattgaccgctccgctggtgaaca aactgctaggtcatctcctccccaccctgggcgtccttccgggcgtccacaccctccccccagcc tgacgtgtcacaggctgtacaaagacccccgcgaaagctgctaacgtggcaattgtgggtccccccc tttgtaaaggaaccgagtctttctccctaaggctagacccctgtgtgaattcacaggtggcaactag tggttccactgcatgctcccgacctcggccgcggagtgctgttccccaagtcgtaacactgaccttc acttatgtgcctggcaagcatatctgagaagatgttccgctgtggctgccaaacctggtaacaggtg ccccagtgtgcgtagtcttcttccgtcttcggacggtaggtgttaggtaaagatgcggcgtaaggtt caagtgccaacgccctggaagggatgacccttctactgccctaggaatgccgcgcaggtacccc aggttcgcctgggatctgagcgcgggctaattgtctacgggtagtttcatttccctcttcttccactgg catc
548	아이치바이러스 Rt386	actccattcagcttcttcggaacctgttcggaggaattaaacgggcacccactttcctgtcctctccc cctttctgcaactcaagtgtgtgctcgtaatcctgactcccacgggttgaccgccccgttggtgaac aaacagctaggtcattccctccctaccccctgggcgccatttcagtggcgttcatatcctccccccag cctgacgtgtcacaggctgtgcaaagtccccgcgaaagctgctcacgtggcaattgtgggtcccc cctttgtgaaggaaccgagtctttctccctaaggctagacccctgtgtgaactcacaggtggcaac tagtggttccactgcatgctcccgacctcggccgcggagtgctgttccccaagtcgtgacactgac ctccacttatgtgcctggcaagcatatctgagaagatgttccgctgtggctgccaaacctggtaaca ggtgccccagtgcgtgtagtcttcttccgtctccggacggtaagtgtgtggtaaagatgcggcgtaa ggttcaagtgccaacgccctggaagggatgacccttctactgccctaggaatgccgcgcaggtac cccaggttcgcctgggatctgagcgcgggctaattgtctacgggtagtttcatttccctctcttttcact ggcatc
549	노르웨이 랫트 페스티바이러스	gtataagggttgggaaccttgtaccaagctacctctgccattcagtatttgggagtagaagtagatgt gtttacaaactcacacgtgtgggggcggggatagactgtgccagcggtcgtgtaccagcacctac gcatacgtgtggactgcgaaccaggagagcacctaggtctgacaagctgtgagaacacagtagt

[1739]

		cgtcagtgagtcagctggtaaggatcacccacctggatactcacgtggacgagggagtttcccag tcagaaacctacaccagaggaggggtcctctggagacatggatggtctgagtaacagactatcta ctggggtgtgctgcctgacagggtctcggctgatagcctggctagcagtataaaaatcagttgaatt ggcatatgagttgtgaacatctagtaaacaatgaaagacaaaaacaaaaaatgagcataataaaaa aattgtacaatccactactcaggtgtggctgcagactt
550	돼지 코브바이러스 GS2	tttgaaaaggggtgggggggcctcggccccctcaccctcttttccggtggccattcgcccgggc caccgttactccactccactccttcgggactggtttggaggaacacaacagggcttcccatccctgt ttaccctttattccatcatccttccccaagtttaccctatccacaccccactgactgactccttggattt tgacctcagaatgcctatttgacctcccactcgcctctccctttteggattgccggtggtgcctggcg gaaaaagcacaagtgtgttgcaggctaccaaactcctacccgacaaaggtgcgtgtccgcgtgct gagtaatgggataggagatgcctacaacaggctcgcccatgagtagagcatggactgcggtgca tgtgacttcggtcaccacgggcatagcattgctcacccgtgaatcaagtcatcgagatttctctgacc tctgaagtgcactgtggttgcgtggctgggaatccacgcttgaccatgtactgcttgatagagtcgc ggctggccgactcatgggttaaagtcagttgacaagacac
551	코브바이러스 SZAL6	cctacccaagggttacatgggaccatattcctcctcccctgtaactttaagttttgtgcccgtattcttg actccaggcggatgttgtgtcgcccgtcctgtgaacaaacagctagacactttcctcccctccctct gggctgctccggcagtccactcctcccccagcgtaacatgccccgctggagtgatgcacctgg aagtcgtggacgtgggttagtaacttcggtgaaaacccactataatgacaactggttgacccccac actcaaaggactcgagtctttctcccttaaggctagcccggccacatgaatttgcagctggcaacta gtgagtccaccatgtcccgcaacctcggctgcggagtgctgttccccaagcgtatgccttccttctg taagagtgcgcctggcaagcacatctgagaagtcgttccgctgcgtcgtgccaacctggcgacag gtgacccagtgtgcgtagacttcttccggattcgtccggctcttctctaggaaacatgcgtgtaaggt tcatgtgccaaagccctgcgcgcggtgttcttctactgccctaggaatgtgccgcaggtacccctac ttcggtagggatctgagcggtagctaattgtctacgggtagtttcatttccatcttctcttcaggtcgac atc
552	코브바이러스 웹 TB3	gaccttctggtacttcttcgcctgggtcacaaaagcgaagaacctgcctctctaacgccagacgag cggcattaaacttgaacttctggcactctccactctccctttccctgtccctttccccactgcgctctc aaggtcgcgcaatcctgggactagcccagttttaaaagttcctggcacccttgcccctctaggccc ttaaggtaggaactgaccttgtgctgtgatctcggtgcgggagtgctaccacgtagtcatcgtaagc ctcgtttctggttctgccctggcaaggctacagagtaccgtgttccgctgtggatgccatccgggta accggaccccagtgtgtgtagcggtatgttcacggtccgccgtgttcaccagattcctgacctgg ctttgctagaaatggtgtgtgcccaatccctgtgaccagtatcaattacatcacctaggaatgctagg aaggtaccccagtcctgagctgggatctgatcctaggctaattgtctacggtgatgctccttttatttttc ttacaactgctattgactgtctgattgctgattctgctcttgtgctcttctgctctggctcattctcaaggg ttctctttgtccaagatcctttggttctctccttgttccacttgccactgccaacgcttgtc
553	프롱호른 안테토프 페스티바이러스	gtatacgcagttagttcatcctgtgtatacagattggagactctaaaaacaacgattcggaataggg gcccgcggcgaagaccgaagacaggctaaccatgccgttagtagggctagcaccaaaacgcg ggaactagacacttaggagagtggtctggctactctaagaggtgagtacaccttaaccgtcaagg gttctactcctcagttgaggactagagatgccctgtggacgggggcatgcccaagagttagcttag ccggggcgggggttgttccggtgaaagtagcaatattgaccacactgcctgatagggcggagca ggccccctaggtagtctagtataaaatgtctgctgtacatggcac
554	돼지 페스티바이러스 단리물 분고와나(Bungo wannah)	tacgcggggtataacgacagtagttcaagtgtcgttatgcatcattggccataacaaattatctaatttt ggaatagggacctgcgacctgtacgaaggccgagcgtcggtagccattccgactagtaggacta gtacaaataggtcaactggttgagcaggtgagtgtgctgcagcggctaagcggtgagtacaccgt attcgtcaacaggtgctactggaaaggatcacccactagcgatgcctgtgtggacgaggacatgtc caagccaatgttatcagtagcgggggtcgttactgagaaagctgcccagaatgggtagttgcacat acagtctgataggatgccggcggatgccctgtattttgaccagtataaatattatccgttgtaaagcat
555	돼지 페스티바이러스 1	gcagatatcggtggtggacctgggggttgggctcaccgtgccccttcatggggtagacctcactg cttgatagagtgccggcggatgcctcaggtaagagtataaaatccgttgttcactaac
556	페스티바이러스 지라페-1	gtatacgagtttagctcaatcctcgtatacaatattgggcgtcaccaaatatagatttggcataggca acaccccgatgcgaaggccgaaaagggctaaccatgcccttagtaggactagcaaaaaatcggg gactagcccaggtggtgagcttcctggatgaccgaagccctgagtacagggcagtcgtcaacagt tcaacacgcagaataggtttgcgtcttgatatgctgtgtggacgagggcatgcccacggtacatctt

[1740]

		aacctatccgggggtcggataggcgaaagtccagtattggactgggagtacagcctgatagggtg ttgcagagacccatctgataggctagtataaaaaaactctgctgtacatggcac
557	돼지 열병 바이러스	gtatacgaggttagttcattctcgtatgcatgattggacaaattaaaatttcaatttggatcagggcctc cctccagcgacggcccgaactgggctagccatgcccacagtaggactagcaaacggagggacta gccgtagtggcgagctcccctgggtggtctaagtcctgagtacaggacagtcgtcagtagttcgac gtgagcagaagcccacctcgatatgctatgtggacgagggcatgcccaagacacacctaaccct agcgggggtcgctagggtgaaatcacaccacgtgatgggagtacgacctgatagggtgctgcag aggcccactattaggctagtataaaaatctctgctgtacatggcac
558	인간 페기바이러스 단리물 JD2B1I	tcagggttggtaggtcgtaaatcccggtcaccttggtagccactataggtgggtcttaagagaaggt taagattcctcttgtgcctgcggcgagaccgcgcacggtccacaggtgttggcccctaccggtggg aataagggcccgacgtcaggctcgtcgttacaaccgagcctcattcccgtcgttaaaccgagcccgtcacccacctgggcaaacgacgc ccacgtacggtccacgtcgcccttca
559	인간 페기바이러스 단리물 GBV-C-ZJ	cccggcactgggtgcaagccccagaaaccgacgcctatttaaacagacgttatgaaccggcgcc gacccggcgaccggccaaaaggtggtggatgggtgatgccagggttggtaggtcgtaaatcccg gtcatcttggtagccactataggtgggtcttaagggttggtcaaggtccctctggcgcttgtggcga gaaagcgcacggtccacaggtgttggccctaccggtgtgaataagggcccgacgtcaggctcgt cgttaaaccgagcccattacccacctgggcaaacaacgcccacgtacggtccacgtcgccctaca atgtctctcttgaccaataggctttgccggcgagttgacaaggaccagtggggctgggcgacgg gggtcgtataggaagaaaaatgccacccgccctcacccgaaggttcttgggctaccccggctgca ggccgccgcggagctggggtagcccaagaaccttcgggtgagggcgggtggcatttttcttccta taccgatc
560	인간 페기바이러스 단리물 JD2B8C	tggtcaccttggtagccactataggtgggtcttaagagaaggttaagattcctcttgtgcctgcggcg agaccgcgcacggtccacaggtgttggccctaccggtgtgaataagggcccgacgtcaggctcg tcgttaaaccgagcccatttcccgcctgggcaaacgacgcccacgtacggtccacgtcgccctttt aatgtctctcttgaccaataggttcatccggcgagttgacaaggaccagtgggggccgggggtca cagggatggaccctgggccctgcccttcccggcggggtggggaaagcatggggccacccagct ccgcggcggcctgcagccggggtagcccaagaaccttcgggtgagggcgggtggcatttttctt cctataccgatc
561	간염 GB 바이러스 A	gccgggtggaaggcccggaaccgccccaccacctcaactaggtggtaagggtacgtctatcggt ccggctggcccgaaaggcggtggatcctgtgtgttagggttcgtaggtggtaaatcccagcacag gtggtaatcgctatagggcaggcttatcccggtgaccgcttccctggatcctggagcgggtcgtgg cggcacggtccacaggagtggggcctccggtgtgaataagccctcgtctggagcatcagacgtt aaactgagacgtcccgaagagatcggaacgacgcccccacgtatggcaacgccgcttaaaaccct tcggggacagctatgcgggttgacaatgccagtgggggccgggcccactattgttgtgggctcc gagttcctctagggatggccgaaaggcagccatggggccacccaggcggcgccgtgctacagg cggcaaggggaaaaatccttcgggtgaccccgggtggcattcccctcccttagcagcatgagtgtg gtggtagctgcaacc
562	유인원 페기바이러스	ggggaatctcaccccccgtccggttccggaagaatcggaaaccgacaccctgaccaatcattctt gatcatagagtggatgttagtgaaagccagacgaaagccggcggatgggtggtgacagggttgg taggtcgtaaatcccggccaccctggtacccggtataagttgggcggaagctgactgaagctccg tgctctttttctgtgcgttcttggtgcacggtccacaggtgacgcctataccggtgtgaataataggcc gactcgagcggagtcgttaaactgagaacctccatacggatggcaacttggcttgcgtacgggga cgccgctaaagtcacagtgggttaagtccggcgggttgacaaccccagcaaggcgaggggtc ctattgttggactctgccagttcccggtggaggtaggcatggggtggcccagctccgcggcgcgc tacagccggggtagcccaaaatccgaaaggtgagggcgggccacatgtccgaaatttagtcaag c
563	페기바이러스 I	agaatggtctaagtggttgccaccgtggtccgaaggggaggaggacctacgctgccagggttgg caggtcgtaaatcccgggtgtaggagatccctccttgttaggactgctggtagctgggggtcggt gacccccgggcaaccgccaaacccggacgaccgggtggcggctccatgttggcacggtccac aggtgtgaaccctaccggtgtgaataagggttggtggttgccggtccacctttaaacgtagtatgcatt gggcttggtaaaacaccgctcgtagtacggaacgccgcctttaaagacacagtaggcgtagccg gcgggttgacaatccatacgggggtggggtgtggtcatggatctgtccacaccaccttcatgcg gccctctaagcaagccatacccgggggaggcgcgcggcaccgcactgccgggcaaggggaa gaaccttcgggtgacccccccccaaccaccgtccgatcaatgctaatgttgcgtttaggcgtgaca

[1741]

		ccggcaca
564	폐기바이러스 K	agaatggtgtgatccgtcgccgctccagcggaaagcgggcgggatctagtggttagggttgttcg cgtaaatcccacactagtggtacgctcgtataacgtgggagcagccggTggggtcgacccccc Acctggcggctgctgagcaccggacgaagcgcgggggtgaacgctaacccgcggcccgggg ctgccaacgttaggcacgtcttggctggaagacgttaaacacagggccccccctcaaccctgatc cgaggccagagaccaaggtacgccgccctttttaaaggcgttactcgtccaataggatctctccgg cgggttgtcaaaccttgctggccctggtgatggttacgggaggggtggggcggggagtagaag ccccgcccggcatggggtaccaagctcggcacgcccagcacgcgtggcgtaggggaaaaat ccttcgggtgacccctggtaccataaagtaattaacatgagcatgccgctagggtgtgctttttcttcc ttccttgggaaggcggtggcacc
565	테일러병-연관 바이러스	tgataccgtgtcccggtacgacctcgcgcgtccccaagctcgccctgaggggggagcgtaaggg cgcgtagtggggtagccccccaaaccgagccaccctagtgagtgactttagaatggttagggaga ctaccgccttcgctgtttggggacctaatgatccgcgtgccagggttcttcgggtaaatcccggcgc ggtgttttgggttcagggcagtaggggcagacgggccagcagtcgctggttcctggtaccaccac cctatccggacgacctccctcacgaaaggtcgccacggtctgtggctcgacgacgcctataattca gtccgaggggcgcagccctcgttaaacttaggcaaggttcctcgccattgatttggccaggggttt aagtgaacgccgccctttttaatgtttaatagggttctttcccggcgggttgacaaacacttccctggg ctcttcgttggcctcggttccttgatgcttcggcacccatgagcgcacaggggggggaccctgcga cagtccgccaagaggaaaatccttcgggtgacctcgtgcgcaacccaatcccttcttcttccacatg gcgtgtctgtggtgcatgctgtg
566	설치류 폐기바이러스	ggacttcggtcccccctgttactctgcgagccaccgcagagccagggttggtacgcccgaggtgtt agaccccggccgaaagctcctaaccatggggttagtaggacgtggtaaatgccactgaggggtt ggagagctggtagagcgagtaagtcggcgtaaggcccgagtacgggcctcccagcccgggtca gcctaaacctggctgtgatacccggtgcatggagggcgtgtcccaacgctcgatcgctgtagggt gggtccctgcagttgggtgtggctaccctgctctgtactgcttgatagagtcccggcggacggacc agctctcgtcagtccgtggagttgcac
567	인간 폐기바이러스 2	aactgttgttgtagcaatgcgcatattgctacttcggtacgcctaattggtaggcgcccggccgacc ggccccgcaagggcctagtaggacgtgtgacaatgccatgagggatcatgacactggggtgag cggaggcagcaccgaagtcgggtgaactcgactcccagtgcgaccacctggcttggtcgttcatg gagggcatgcccacgggaacgctgatcgtgcaaagggatgggtccctgcactggtgccatgcg cggcaccactccgtacagcctgatagggtggcggcgggccccccccagtgtgacgtccgtggag cgcaac
568	GB 바이러스 C/G 형 간염 바이러스	cccccggcactgggtgcaagccccagaaaccgacgcctatctaagtagacgcaatgactcggcg ccgactcggcgaccggccaaaaggtggtggatgggtgatgacagggttggtaggtcgtaaatcc cggtcaccttggtagccactataggtgggtcttaagagaaggttaagattcctcttgtgcctgcggc gagaccgcgcacggtccacaggtgttggccctaccggtgggaataagggcccgacgtcaggct cgtcgttaaaccgagcccgttacccacctgggcaaacgacgcccacgtacggtccacgtcgccct tcaatgtctctcttgaccaataggcgtagccggcgagttgacaaggaccagtgggggccggggg cttggagagggactccaagtcccgcccttcccggtgggccgggaaatgc
569	말 폐기바이러스 1	agaatggggagttaactcctggcactggcccgaagcatgaactgatcgcggtggcagggttcttc gggtaaatcccggccgcgtgttgtgattgtgttagggcaggtgacagtcggcagggtcgacccct tgcttcaggaccactgtcttcctggacgaccgttgctgaaaaagggccgccacggtctgtagctcg ccgacgcttctaattcaggccggaggaccacgctccgtaatcgagcccaagtactcaaaccccag caccccctgggtcacgccctacgccgccctttaacgcttcggctaatagggtctatccggcgggtt gacaaagggcgcagggttacctggtactacgagcttgggtgtccctgggagtaatcccagggtgcc c
570	쿨렉스 테일레리 플라비바이러스	atataaatcccagtttggttaaacctatttcaaggcttaagttgtttattattttatcgccgctcgtgactat aaagttgcctagcggagagagataaagaagaaggagttcaaggctcagggcagggcgcaagtt ccctggtccctaggccgctcgcaggaaggaggagtgaagaagaagaaagagaaggagaggac caccgccgaaagaaggcaggtgcctcacaagagggccaaccagcgtgttggaccagtggcca acgccggacggcgtggtggcctgctgggacgcctggggattggatggagtgccttcctacagga agacatcgttcaagccatc
571	부수루아라 바이러스	agtatttcttctgcgtgagaccattgcgacagttcgtaccggtgagttttgacttaacgcagtgagaa aagttttcgaggaaagacgagaagcgaattctctga

[1742]

572	지카 바이러스	agttgttgatctgtgtgagtcagactgcgacagttcgagtctgaagcgagagctaacaacagtatca acaggtttaatttggatttggaaacgagagtttctggtc
573	요코세 바이러스	agtaaattttgcgtgctagtcgctgagcgtcagaccgcaaagtgagtttttagtgatctaaagtgagg agttattcttactgtcatcaaacactacaaataaacacgttgaaattatttccggaagaacaactgtcc ggaatcaaagacg
574	웨셀스브론 바이러스	agtatattctgcgtgctaatcgttcgacgttagtccgtggagtgagcttctattagagtcgttaacacg tttgaataatttctactgaaaggagtagaagaaaggagattcattccca
575	말 헤파시바이러스	acctccgtgctatgcacggtgcgttgtcagcgttttgcgcttgcatgcgctacacgcgtcgtccaac gcggagggattcttccacattaccatgtgtcactccccctatggagggttccaccccgcccacacg gaaataggttaaccatacctatagtacgggtgagcgggtcctcctagggccccccccggcaggtcg agggagctgaaattcgtgaatccgtgagtacacggaaatcgcggcttgaacgtcatacgtgacctt cggagccgaaatttgggcgtgccccacgaaggaaggcgggggcggtgttgggccgccgcccc ctttatcccacggtctgataggatgcttgcgagggcacctgccggtctcgtagaccataggac
576	헤파시바이러스 B	accacaaacactccagtttgttacactccgctaggaatgctcctggagcacccccccctagcagggc gtgggggatttcccctgcccgtctgcagaagggtggagccaaccaccttagtatgtaggcggcgg gactcatgacgctcgcgtgatgacaagcgccaagcttgacttggatggccctgatgggcgttcatg gttcggtggtggtggcgctttaggcagcctccacgcccaccacctcccagatagagcggcggc actgtagggaagaccggggaccggtcactaccaaggacgcagacctcttttgagtatcacgcct ccggaagtagttgggcaagcccacctatatgtgttgggatggttggggttagccatccataccgta ctgcctgatagggtccttgcgaggggatctgggagtctcgtagaccgtagcac
577	헤파시바이러스 I	cagggtttcgaccctggcccggatacctatcgccttacgccgaaaggtaacgagtaggagtcggg tccccaggcccttaccgccaccaagccaggtggggaggtatgggagccgggggtgcagctg gtagctccatgggggacgccccgtgagcggatgctgcatcgatacccgggttagctctctgggaga gcggcacttgacaccacgaatccgggaaccggacaatcgccggcgtgggacgcgttgcctccg tggccgagcaatttggcatgcccgtggtgaagagtgatggtgggggggggccccccttccagta ccgtactgcctgatagggtccttgcgaggggatctgggagtctcgtagaccgtagcac
578	헤파시바이러스 J	gccgctcccgaaagggagtccggcgcgtcatcccactccgaggagtggggtggcgtccccgtg tgccggggaaccatgaagcctaagggcatccacatttttagaatgaacttgaagcttcgtttcgctgg ccggaaagtcctgggttcccatggccagggttccgcaggtgggtaaatcccggtggggttccatc caggatatacggcaggcgggcgtagtccggcggttccggacgacgtgtgggtcgcctacggtgg attgttcacaggatgggcactccggtgtgaataggccccgtcagggtgcgctgacgttaaactcag gccttgcctggtgttcggggaggattgcagggccacgccgcctctaagggccgtatggcacagta cttcttcgggcgggttgtcaaggccctccaacgcgacaccagtgcctcggcaggcatggggcca cccagctcggcgtcccgcacacagacggcgtaggggaaaatcagcaatgtgaccccgggtggc attttccttctctctacttccatgcatgatcaaccgcaatc
579	헤파시바이러스 K	gggaacaatggtccgtccgcggaacgactctagccatgagtctagtacgagtgcgtgccacccat tagcacaaaaaccactgactgagccacacccctcccggaatcctgagtacaggacattcgctcgg acgacgcatgagcctccatgccgagaaaattgggtatacccacgggtaaggggtggccacccag cgggaatctgggggctggtcactgactatggtacagcctgatagggtgctgccgcagcgtcagtg gtatgcggctgttcatggaac
580	이카바이러스	cgaagtttaagctaagcaccctcgggcgttcccggattatgtgatcacatcaatttgatggctggtca ccacgcaacgcctggagagatactcttactttctcttaagatccccggtcatttgacgcttgtaggat gatagggttatttttccactataaaatactttcatactcttggatgttctatatccaagacgggaggaccta ccccgtacccctagaggtgagatgccaagaacaggccctttctgttctctcgacaatggcatcata ggcaacaagcatcacaccaagattgctaagttttgttaagagttcttcaagctatagggtggctgtag cgaccttctgatgcctgcggatttcccacggagcgatccgtgccacaggggccaaaagccacg gctaacgcccatcaggagcggcacttaccccgtgccccaccctgaaacttgaatgttcacactgg cttctctcggctttctgaactgtctgcttgttggggccccgaaggatgccctggaggtaccccatttta tgggatctgaccaggggacacctcagctctctaagttgctggtgtttaaaaaacgtctaagggccc ccaccccttaggtggagggatccaccttccttattttttaaaactcttttatggtcacaattgttt
581	안타크틱 펭귄 바이러스 A	ctaggagactacgcagtgggataagatgactatgatgtcgtacgggcagaagccagtacagtcga agtcgagaccgacgtcgaggatttgactctgcctgacctagtgccatc
582	포레스트	tattggatccgcctccgggcaaaggttacttttcttgtacctcggcttagccacagggtgacccctttgt acgtaggggccccgacgtagcactggtctgacaacaccttctcggcatttcaccttctgcccgctct

[1743]

	파우치즈 자이언트 래트 아르테리바이러스	tccgggcggtggtgtcaagaagcagcagtgctcttctcttttcttcctgcagttcaccgagccctacg gggggtaggtg
583	아비시바이러스 Pf-CHK1	cattccccttttccccagccatgggttaaatggcccctcaccaggttcggtgctgtctaggcttccagt aaagaagtcaaccgagcattgaacaaaacctcagtgggtatggtagtttaaccccgtccactggac aactttgtcctcttaaaagtggatcaatccaccccaactccccccctagccacctgagccatggtgg atagcagtgacgaaactagggaccccaatacctctagtgccaagagaattcccccctcgcgagag gtgctcttgggcccgaaaggctagttggcagggtgaagtgaaggaagctgctagcgtggcaacc ttaagcgtagcccgaagctgaccttagaggttaaccctagtggaccactggatgaagctgtggag gtggtggataggaaagttggccacttgtgagtagatgcccagaaggcataaggctgatctggggc cagtgactataccgttccggtaaacctggtataaaaaccatgaaagcaagtgggtttaaaatttcttct aattccttcatttcagtagtgataactggcaga
584	조류 파라믹소바이러스 펭귄	accaaacaaggactagataacccacgtgaccgttaactggaaaataagatgttgtaggggcgacc tagttggaattcgaccccggctccgaaacctctaattgtggttattggcagtctagtctacttctaacg
585	뉴캐슬병 바이러스	accaaacagagaatctgtgaggtacgataaaaggcgaagaagcaatcgagatcgtacgggtaga aggtgtgaaccccgagcgcgaggccgaagctcgaacctgagggaaccttctaccgat
586	박쥐 Hp-베타코로나바이러스	ttaagcttcggcttgttgcataggaccggaaaggtactatctaccctaactcttgtagttagactctcta aacgaactttaaaactggttgtgtccttcagtagtctgtatggccattggaggcacaccggtaattatc aaatactaagaagattcatagtacatccttgtctagcttttggttggcagtgagcctacggtttcgtcc gtgtcgctcacaattatccacacagtaggtttcgtccgctgtggttgagttgctagtccgttgctgtttc gtcagccatctacaactcgacacc
587	바쉴라 알바 엔도나바이러스	ggaatatggctaatcggcttattctaatcaaacgcaaaagacttatgacacagaccggacctgaac gaggtgataaaacacctcgttcaggttcaaaacgtagaagattcattcctccgattagaaatacaact acgtctaagcacgatagagatggtatcaagatcggttttagaccacgagaaaatcggcaaatgaaa gtacaagttgggtggtttaaattaccaagaacagtaacattcaagaacaacggcaacccgtttgtta cctcatttcgtaaattgtttagaagtaacaaagataaattatttaatggtgggaagaacctaagtacag taccagccagaagtagtgaaatgacagaaatgtttatgttcatgtccacgctagagggccaattgtc aatccaagatcgagatccaaaaataatcaataagtctatatacatgatagaggta
588	불 피톤 니도바이러스	ccccttcacccataggcactaggagaacaggataaccccctaacggggcatcctgcctgtgaccttt cagattcgctagttagatatcttcacagactctgctaggcttctgacccagtccgttcccaaagtccgt taccgcccgagtagcgcttaggcgcgaaagggacggaagtacctccagtaagcgaaagctgaa gtaagggaaatacggcaagactaacttgttagtcttacagtgtggataacctggtagttatccccga caagagacctgactcgatgtgaaaacatccaactaggttggcttcaaactagctacaggcaggata tcccccaacggggcctccaaggaatcgagaaccaaccctcattccacgtctgtagtaagcaaaaa caggggcgatcttcaccgacacctctcaccacagagcacaccaacctctgtgaagccaatttcctc gtccaaggacaggttattgaggtcaactttcttccgaccagaagaagggatttcctaccaaaaga aaaaccaaatccaccaacaccacaaggtaaaacaacaacttgtgaagccaatacttagtcaaaga ctaactattgagggtcaactttctcttcaatagagaagggatttcctggtaaaacaaataacaacaac taacatcagcaact
589	박쥐 사펠로바이러스	gacaggtgttttggagggcggatgacgatatctggctggccaccagggaataacggcaaatgtct gatcatacggttcacaagtctaccggcgatagtggttcaacaccatgtgtagcagggattcttgcgt atgtgaaggcgacagtgc
590	박쥐 피코르나바이러스 3	taagcggaaagcattcttgtcccccggtcagtaacctataggctgttcccacggctgaaagggtga acatccgttacccgcctcagtacttcgagaaacctagtacgcctgatgattccaaattggtatgatcc ggtcaaccccagaccagaaactgtggatgggggtcaccattcctagtatggcaacatacaggtgt ccccgcgtgtgtcacaggcccttacgggtgccatttcggatgagtctggccgaagagtctattgag ctactgttgatacctccggccccctgaatgcggctaatctcaaccccggagccactgggtggtgaa ccaaccacttggtggtcgtaatgagcaattctgggacggaaccgactactttggggtgtccgtgttt cttttgttcatattaaactgttttatggtcacaacacaacttggtacgatttgtgattattcactgctcactt gtcacagtaaatatacacaatcatc
591	박쥐	gggttttacgaaacccgtatacaccagacctttttctcccctcccctccacctacctttttcccccctctt

[1744]

	피코르나바이러스 2	ggaccgaaacaaggacacgtaagtggaaacgcgattttatatgtggttggccaccacggaataac ggcaattgtctacatgtgggaagtgcaacctccctagccgataacccctgaccgggtgtgtaggat aggaaaggtgcccactgtgggcgacaggttatggtagagtggatacctagccagggcaatggg actgctttgcatatccctaatgaagtattgagatttctctgctcattacccggtgatggttgtgtggggg gggccccatacactagatccatactgcctgatagggtcgcggctggccgaccataacctgtatagt cagttgaattcagccaag
592	박쥐 피코르나바이러스 1	gaaacccgtatacaccggacctttctcccctccctctccacttacctttttccccctcttcggcatgaaa caaggattattcaagtggaaacgcgatttaatatgcggctggccaccgcggaataacggcaattgt gtatctgctggaagccaagcctgcctagccgatagcccttgaccgggtgtgtaggatagcccagg aaccagcaatacgcgacaggttatggtagagtagatacctagccagggcaatgggactgcattg catatccctaatgaaccattgagatttctctggtcattacccggtgatggttactagaggggggcctc tagtactagatctatactgcctgataggtcgcggctggccgaccatgacctgtatagtcagttgatt tgagcaat
593	박쥐 이플라바이러스	acgaatcggtatacgcttcggtacctattgttgcaagttcgttccctattttcgatttgcctgcccgaatt tgactcaaacaattgtgacatactatgtctctgtttgaaagcactacacgagtttgcgcccagcatgta tgttttcaagtcttttgtataagtctgcctctatagtggttttctttgaccttaaagccttgtcaaccatcct atatgctgcatcgagacttgatgtcaatctgcctctactacgcaaatgtctagtaattagttataaggtt ttactattttccctcatttccaattttagtttgtagtgtatgtgagtatcattcttactccgactgttaagaga aaccaatttatagtcgttaaatatgataaatggaatgaatgatggtgtcattttaaaaacactcttctcta taggcgtaagcattctcgctcttagagtcgtaaagaagaaatgccgtgtctatcagtatgttatgcga tttattttctgccacgcgcttctagtgcaatctagttgacatacagacattgcctaccactcgcgaggg tcgaccggtagtgtaaggagtaagtgatgatttccgcttattctgtacccttgcctggtgaggacag atcctgactaattttaaatataaatgaacactagcttccaag
594	박쥐 디시바바이러스	gtatagcaccggaatggtattttactactccaagtatacgtactagagagttaaaccctgtaatttacag gggatttagtgacttttatccgtaaaagtcgattggacgttaatcggtaacgaggccaagtaccgtg aaccaatttaaaaacgtattttctcatgtggtagaaccaacttggaaatagcatggcatatagggttgttt aggg
595	베타코로나바이러스 HKU24	gataaagtgtgaatcgcttccgtagcatcgcaccctcgatctcttgttagatctaatctaatctaaacttt atataaaaacactaggtccctgctagcctatgcctgagggtttaggcgttgcatactagtgtcttagga atttgactgataacacttccctgctaacggcgtgttgcactctcagtctaagcctcccacccataggaa ggtatc
596	베타코로나바이러스 잉글랜드 1	atttaagtgaatagcttggctatctcacttcccctcgttctcttgcagaactttgattttaacgaacttaaa taaaagccctgttgtttagcgtattgttgcacttgtctggtgggattgtggcattaatttgcctgctcatc taggcagtggacatatgctcaacactgggtataattctaattgaatactatttttcagttagagcgtcgt gtctcttgtacgtctcggtcacaatacacggtttcgtccggtgcgtggcaattcggggcacatc
597	분 카디오바이러스 1	tacgatcgctgtacattccactactgccaattagctccccctcccgttgctcccctctataaggaga gccttctcttgcaaaggtgaagccttcacccccggtcgaagccgcttggaataagacagggttattt tctcctctcctcggcgcttgcctcttctaagctgaataggttctatctattcaggcggatggtctggtcc gttccttcttggacagagtgtgtatctgggttttccggatctcgaccacacactcaccagagctcagg agtgattaagtcaaggcccgatctgcggcgaaaaggaaatgaagtatttttgcagctgtagcgacct ctcaaggccagcggatttcccccacctggtgacaggtgcctctgggggccaaaagccacgtgttaat agcacccttgagagcggtggtaccccaccaccctgcaaattatggatttgacttagtaactaaaaga ttgacttggcatacctcaacctgagcggcggctaaggatgccctgaaggtacccgtgttgaaatcg cttcggcgaccatggatctgatcaggggccctgcctggagtggttctatcccacacagcgtagggt taaaaaacgtctaaccgccccacaaagaccccggcagggatgccggtttcctttttaccaattcttg acact
598	브레다 바이러스	atcacctagtacttacaagcgggtcaaaccgccctccggaacggtcataacccccctcccgaacgt gcgcttgacgtgactggtctttcagtctagctttctgagaaatactccgggggttgtaacccaccatttt gacctttggtcagtttggtaacactccaaccaaacagcatctgacccacctccagcttgctgcaggc catttggacCaaacgggttcagatatcttgtggctaaacctctgacccacctccagtttactgcagg cctttttggactaaacgggtAcagactctttgtggttagtttttaactacccactgtttagccgccaacc tgatttttattgttacaaaatttttgtgtttacacattatttttcttacggttggcagtttgtttggttgtttgcaca gttttgctgatacccaatttttactgtgctttttggtgttttcggctaaggctgtttttcacatacttagtttgct tgaagtaaccttcacaacatctgttttgttttttggtttctaaggtaaagagtttcaggaaaaaacatagg

[1745]

		cgcccatcttgtggtgtctagttttaattaatctggcaaacaagtatcaagtcatcgactccctttggag tgagacttacgagtaccaattcgcctattttggccatccatataaaa
599	소 바이러스성 설사 바이러스 3	gtatacgcccagttagttcaggtggacgtgtacgattgggtatcccaaattaataatttggtttaggga ctaaatcccctggcgaaggccgaaacaggttaaccatacctttagtaggacgagcataatgggggg actagtggtggcagtgagctccctggatcaccgaagccccgagtacgggggtagtcgtcaatggtt cgacgcatcaaggaatgcctcgagatgccatgtggacgagggcgtgcccacggtgtatcttaacc caggcgggggccgcttgggtgaaatagggttgttatacaagcctttgggagtacagcctgatagg gtgttgcagagacctgctacaccactagtataaaaactctgctgtacatggcac
600	소 비염 A 바이러스	ttttgcggctctgccgccgttcgggttttacctgttttcacagagcaaaacaggacctctagtttcgtg cttaaacgagatcatgctcgaactagaactataacgctggtcactggacccgtgccgcgccttgcg gatctttgcgggaatggtggctagtgggctgtggaagtgactctaaccacacgcccctcaagtgtg ggaaaacacgaactggtgtagcgacgacgataggccttgggacacccctctccagtgatggagac ccaaggggccaaaagccacgccttgtgccctgtcgttcacaaccccagtgcagttcgtgccagta cctgcttttgggaagtgtgctttggacagctgaaaacagtcctagtgggagactaaggatgcccag gaggtacccggaggtaacaagtgacactctggatctgacttggggagagcgggtctgctttacag acgccactctttaaaaaacttctatgtctcgtcaggcaccggaggccgggccttttccttttaaaacaa tacacttt
601	소 피코르나바이러스 단리물 TCH6	ttggatctgagcaggggccccctttgggttgctttacaactcaactgggggttaaaaaacgtctaac ccgacacgccagagggatctggtttccttttatttctttttactcaccactggatgcagattgacgataa acgttgttgtttgtgactattgacttgatctgcttctacgggtttacttttcactgttatacttcttgctttgttt ggtgttcactgtactttgtctccttctacatttcaca
602	소 니도바이러스 TCH5	caccaatagattagtcaagctgtctataggcataaactaacccccaaccccattaccccggggcca ggtgggccgccgccttcgggcaaacccgtgcgctggtataatcaaggttcacagccagattcact gccggttagctagtggggcggtagcctggcaaaacccgaagaggttggaaagggaacttcagg gtagtttatcctaggctagcgtagctacagttcggtcaagataaccgtcctggtgctagggctagta gagacagtggtaacttggacaagggtccagggccacttttagggaataccctacggaaggctagg tccgtaaggaagacccccgcagttgtccgcggttgagcagagctcctgcgtagacaaaaggcaa aaagtggattacattcgcctgcaggaaaaggcaaacgtcgttggagtcggagctaaagtactgga cgattgataccacgcctgctgcggtagataaa
603	소 헤파시바이러스	ccatcgacactccaggctcacggattaagttaggttccgccgaagcgggctaaccaggcccctag taggaggcgcctatcccgtgagccctttccccacggattgagtggagctggagctgggaaggac cgagtacggtccaatcgagaagaaccctgatgaacattccaggcctcttcggtagatttggatatat ccaccagtgaaggcggggtcgtgggtacaggccccctagtccacacagcctgataggtcctgc cgcaggatccgtgggtgcggctgtacatgtacc
604	보트리티스 시네레아 미토바이러스 4 RdRp	gccccgccgaccttcctatttatttctaaataggaaggtcccgactagtcggataattcggtttaacg aattatggttagatctattaaagttaaaatagattgaatttctttctcatcctccttattcctatacttgggga gtaaatgacaaatgtctatcctcaaaccgaaatggcttaagtgatgaatttgaaagaaaggtaggttt taagaatataaggcatcaatatattatacccttgaatgttaagtgaccacggcgtgacgattagggct atcttaggatagacagccatctaacgcgacagcagtggaaatcagcttagcatctcaagatcatgta taatatatacataaccttacaattataaaaccaaaccaaaacacactataatttttatataaattatagaa gtatcggacctggacggtacctactattaactgatagtagtccaaatgcaggaagctc
605	보트리티스 시네레아 미토바이러스 2 RdRp	ggaacttttcagttccagaagtggctttattaagccttcaaagtttacactttgaacctgcgattaattcc ccatagtgactcttgttactgtgaattaggaatagttgtagttcaacttctaatgaggtgaacaatataa taactcatcttattaaccctatacgtagacaattgtccaaagagacagttggaattctgccaatctgga atgtttggtacgcgtagaagataataagagaccctctattccctgcctcatgactaagtcatggccc gggggtgtaatagagatacttttatatattatacaatc
606	개 피코디시스트로 바이러스 균주 209	cgctcttttatacaaatctgtcaaccctttgtataactctaagccgaacaattatagctaggctttttattat ataacattaattaggcattagcgttgtcgccaatctcttggtaatcctaaggatacctttcctgttgacta agatgaagcgccttcggttaccgatgcccggtgtccacgaagccatcgtggtcggccgcgtcccc cacctctcccaacttggactccatgtttcagtaggtgtaatgattagtattattgattctgctcgttcaat gtgtttatcttcacgatctgggacccaacacatgcttcactcatgtttaaatgttggttccctcattttga agacacccaaaccatagagtgcgagaatgaggatttctacttccattctggtaacagaaatgaattc ctgcgtgtgtctcgtaaatggaatctttaagaacttcagataaatcgaacaatacactaatacaagttg ttttctaccaacatgttcaatgcggctaatctgaccgtggagctgtgaagcgctcaaacccgagtgtt

[1746]

		gtatacagtcgtaatgcgtaagtccatgaggaaccgactactgttacctctgttggtgtgtttctcctttCCtctcttttattattatttgttattgcaaatactacaactttgatcaac
607	개 디스템퍼 바이러스	accagacaaagttggctaaggatagttaaattattgaatattttattaaaaacttagggtcaatgatcctaccttaaagaacaaggctagggttcagacctaccaat
608	개 코브바이러스	tttaagtgttgtgcccaatctcttgactcctgctggaaccaccgaccagtagtgtccaaaatgccaggtggaaaatcctcccttccccctctgggcttcatgcccggcatcctccccccagcctgacgtgccacaggctgtgcaaagaccccgcgaaagctgccaaaagtggcaattgtgggtcccccctttgtcaaggcgtcgagtctttctcccttaaggctagtcctgtcagtgaactctgtcgggcaactagtgacgccactgcatgcctccgacctcggccgcggagtgctgccccccaagtcatgcccctgaccacacaagttgtgctgtctggcaaacattgtctgtgagaatgttccgctgtggctgccaagcctggtaacaggctgccccagtgtgcgtaattctcatccagacttcggtctggcaacttgctgttaagacatggcgtaaggggcgtgtgccaacgccctggaacgagtgtccactctaatacccgaggaatgctacgcaggtacccctggctcgccagggatctgagcgtaggctaattgtctaagggtattttcatttcccaccctcttcttcttgttcata
609	낙타 알파코로나바이러스	cttaagtgtcttatctatctatagatagaaaagtcgcttttagactttgtgtctactcttctcaactaaacgaaatttttgctacggccggcatctctgatgctggagtcgtggcgtaattgaaatttcatttgggttgcaacagtttggaaataagtgctgtgcgtcctagtctaagggttctgtgttctgtcacgggattccattctacaaacgccttactcgaggttctgtctcgtgtttgtgtggaagcaaagttctgtctttgtggaaaccagtaactgttccta
610	크리파바이러스	gcaaaatcggtagtacgttaaacgtacgaccaccgatgagactgaaatgacactagttgagattatttcaatatcctagtgttataaagtcaatatttgttggttgatcgtttcgtcaatcgatggcgctgacagccggaaagacggcaataataaaaaccaagatttagtttttaagttttgattgaattgcaaaagctatcttgaatagacaatcaaaatattaagtaaagcaaaagcttcttaaagaagacaaatatttaattagttagtaaccaaacctcatcgtgccccctaagggttaaccggttacgtaaaagcgtagaggtattaaggtcactgcggagacctaaaatccgcaattttatgttttgtaatgttttagttatagacttagatgtaactataagagtttataaatacttgtttcaagatttatagacaagatctgatcctatggattttagataaccttcatgttagtggatagtgtgtgtacctatctaaacgcataaggctcttatttcatatttaaagtaggactatgtattacggcgcatctaacggtaacgttagtcaagaccggagaatctcggaatgaattttagtaattcccaaatttata
611	인간 콕사키바이러스 A2	accttgtgcgcctgttttatatccccaccccgagtaaacgttagaagttacgcaaccccgatcaatagtaggtgtagcactccagctgcatcgagatcaagcacttctgtctccccggaccgagtatcaatagactgctaacgcggttgaaggagaaaacgttcgttacccggccaattacttcgagaagcccagtagtgccgtgaaagttgcggagtgtttcgctcagcacttcccccgtgtagatcaggctgatgagtcaccgcgatccccacaggtgactgtggcggtggctgcgttggcggcctgcctatggggcaacccatagggacgctctaatacagacatggtgcgaagagcctattgagctaattggtagtcctccggcccctgaatgcggctaatcctaactgcggagcacatgccctcaaaccaggggtggtgtgtcgtaacgggtaactctgcagcggaaccgactactttgggtgtccgtgtttcttttattcttataatggctgcttatggtgacaattaaagaattgttaccatatagctattggattggccatccggtgactaacaaatcgctcatataccagtttgttggttttgttcccttatcacatacagctcataacacccctcttatatttactacaattgaatagcaagaa
612	코로나바이러스 AcCoV-JC34	agaaacaagtagtgttttaaaaaccttcaaattagtgcctgtaacatctttgcaatgaaagtagcgctcactagcctctatgcaaagaatgttaaaagaaatacgaagcatttaaagaatacaatctatctaggatagtacaattctcctccccctcttgacttcggtcaactcaactcaactaaacgaaatccccctttgcatggttccgacccgtgtaaggttgtgtatttcgtgcagtcgttgcccttactagtgtaagcgtaacggcatctaggtttgcacgtcttggaggaaacggtgtgtacgtttctagtgtttacgccgtatcggttccggcccgataggtattgcattagacgtcctgggtggttctgcctgcccttgtgtgattcggctgttccgtcagtttggtcacctcacacgtccttaagac
613	닭 피코르나바이러스 3	gggtatggtggttaaccccgtccactgggcatctttgccctctcagaagtggatcaatccaccccaactccccccctggttacctgagccacagtggactccggtgacgaagctagggaccccaatacctcaagtgccaagagagtccccccctcgcgagaggtgctcttgggcccaaaaggctagttggcagagtgaagtgaaggaagctgctaacgtggtgaccttaagcgtaattcgaagctgacctttgaggttaacctctagtggaccactggaggaatctgtggaggtggtggttaggaaagttggccacttgtgagtagatgcccagaaggcataaggctgatctggggccagtgactataccgttccggtaaacctggtataaaaaccatgaaagcaagtgggtgaaattttctcttttatccttcattcagcagttgatattggcaaa
614	닭 피코르나바이러	gtggccgacttgcagaacttcctaccgaaccaccacctcaccccataactccttccctctacacctttccgctatggtggttaccactgctttattgccttgactgagaatggccacccccctcgacacctgcccc

[1747]

	스 1	ctactgccccaccgcgcaaccttttgcttgtccactcggttggagaggcatgggggccccgttcat atccccagtccagttggtgaccttccccctccccgtccggtagatggtccagagggctttgccgac gccctctatgatgcttgcttgtctttcctccgtcagcgcgagcatgcacttgtcgagcccacggaaa cacagcctagcttttgcttctctcaccctgcgtaccctgggcgccttccgctcgagattcgctttgttc gacaccctggcgtccccccaccgctacgtgattttactcgtggcatacaccgccctggcgttcagt acattccactgcctaatttgggtggcctccctcaatctcccgcaccccccattgcgcacgtcatcac cgccgccgctaacgcgatccggcgcggttctcactggcactgtcccctcgtccgccgggtttcca ctcatggttggctttcacttattctggttactggtgttccatcctacttcatgatggtcgccatgaccat gac
615	닭 오리바이러스 1	aaaccctcacgagtgcttgtggtaggtcccaggccaatattcttcgtaaggcttggttccaattttcca ccactcgtgtttgggttctggcctatggtacccagaggggcggtttgggggaattaactcccctcc cctgtggtcctataccaccccacacctctgtgggctttctttactatcttcttgttttccgacttttaaaca ctaggcaggcgcgcctagtcatacaccgcccggctggtctttccagctcttgtgggcggtgcgcg ctggtccatcgtgcccagcgacatagcaccttgtggacacctccgaacgcccctccctgtatggg gtggtgcccagggggtttcagtgtggtgacacactccctggggcccgaaaggctagtgtgcaacag gtgaggtacagccagctgccccccgtggctggagggaccaagcttgtgaagcacacctcaccttct tgggggtgggctagtaagtggtgaaagcatagtgtccgtgtcgctggccaacactttgggtcaagt ccagccactcagtgagtagatgcccaggaggtacccctagtggatctgacttggggcctgttactt aatgcaggttaaaaactatgaaagctgagtagtgtagcccggctggtggcttctcttccttattcattc tatttt
616	닭 갈리바이러스 1	ggttaacttgtttaaccaaggcttccgtgcagggcttgcacgttaggagttcatgttgattcatgctcc aatgcccaaaactttgtgtgtttgtttatgtctttcccaaagtttcccccaaggtttcggtactcaaacct taattcctagtcccatcttttgggccttagtatctaggaaatgtacccgtgccttgacgaacgtaagaa agctgtcttttattgaacggttctaatgaactaagtataactggctcgcgccacctggtgtgtgccgtt ggaattccccatggtaacatggtccaacgggcccgaaaggctagtgggcaatcggtcctccaa ggaaggggttcccaccccgacctgaacaggatttgatgaagctcacctcccaggctcctaacccc aaggaagtttacttatagtaattagaatttagtatgtaattgctggcaatcttgctagtagtcaggaacg ttatgaccaaatgagtagaccccagaaggtaccccattatatgggatctgatctgggcctcatact gtgtgtctccccacatatgaggttaaaaccatgaaagtttggtccaaaatattcttttccttttatctttttct ttagtggtgacgccattatatcagcagtttgctg
617	닭 칼시바이러스	ggtgcatcatcactgaacaccctcgggcagagatgcaagggtggaagtcactcctgccccctgg caacatgcaggtgcccgatcccaagcttagttctgacttcctctcctcgggtggtgcaacactccaag gttgatgaacaaacctggagggacctctgggcaacctggtctctcgaggatctccggcgcatctcc acagactacctcatgctcccggaacctgagaagaacactgatgcctatgatggctggctgatggtc ggcgagatgttgacctttgccggtctcattggctgggagg
618	잉어 피코르나바이러스 1	agctacaggaaagagagagataatcacagcacataaatacaactacagaagagagccctttccctg agcactatttacagcaaaccacgctgggaaaagtggtagcatgacccacttacgggttacttagtat aggattttaatatcttcgattctttatttttctttttcttaactaagttcgcccggactttatccgtgttttatttta gtttaagcaaaattgagtaactaagttttttaacctgccaaatggtgagaagtaactctgtgaaaatacc atttgtgcatgaaattgtcttgaaaactcttaggcttttgggggtcccactgctgtttggaggactatt gacagactctattgtagttgagtagtgactaatgataacgatttgcgtattacgcaatgggctgtacc cgttagatttagtatgccgggggggaggggtcccactggattgcactatgtaacctgacagggcgtc tgccgacgcactacaatgaggataagatcggctgtttttattt
619	매 피코르나바이러스	cgcttggaataagttgagaggaattatgcatgctagttgtgtttgtttacaactaattgttctaatccaa gtgaagctcttcgcttggggcggcacgacacttgccgtaattcttctaccgtccctccacaccttgtg gatgaagggccggatgtgtggcctctggctaaccccctctctctggggtgatgctactggatgtttta ctcctagaccaaatcacatgaactcctcttgatccacttcggtggggctatgagcctgcggattaata gctggcgacagctaccccaggggccaaaagccacggtgttagcagcaccctcatagtctgatgc ccaagggctgatgttgggagctagtagtgtgtgtctggcctatgtttaggacttctggccaagcgca gaggagtggggctgaaggatgcccagaaggtacccgtaggtaaccttaagagactatggatctg atctggggccccctcacgtggctttaccacgtgttgggggttaaaaaacgtctaggccccaccagc ccacgggagtgggctttcccttaaaaagcccaacaatatttatggtgacaattcactgtttcttctttgc aatttttgtattcttggactccttatttttgtttgcttgattttagtggacattcagattcaaatacaaa
620	말 비염 B	cgacaggcacaggtcgctccgagttctagtagtgtgggaacttgttactactgatgaaacgaggta

[1748]

	바이러스 1	gtgacactcagtacctgcgaacgaggtcggggccctccctccttcttccttcacccaactttcacttttcgt tccactttagcaggggtcttctttctatccccctggcggcattggaactagccgtcgcgtcttaacgc gcagccctgaaggccccacacctttgtggatcttgccgtgggtatgtttctggcatgtgtttctcaagc ctgcaaccgaagccgaacagccacatgaacagtttgagcgtggtagcgctgtgtgagttggcggt ggatccccctcgtggtaacacgagccccgtggccaaaagcccagtgtttacagcacctctcaca tccaggacgaccccatcctggcgctcactcttagtagtatggcttagtacgcattaggtggtaagcc gagctctccctcggccttgttctgaatgcacacatgtctaggggctaaggatgtcctacaggtaccc gcacgtaaccttcagagagtgcggatctgagtaggagaccgtggtgcactgctttacagatgcag cccggtttaaaaagcgtctatgcccctacagggtagcggtgggccgcgccctttccttttaaaacta cttgttct
621	말 비염 A 바이러스	aagggttactgctcgtaatgagagcacatgacattttgccaagatttcctggcaattgtcacgggag agaggagcccgttctcgggcacttttctctcaaacaatgttggcgcgcctcggcgcgccccccctttt ttcagccccctgtcattgactggtcgaaggcgctcgcaataagactggtcgttgcttggctttttctatt gtttcaggctttagcgcgcccttgcgcggcgggccgtcaagcccgtgtgctgtacagcaccaggt aaccggacagcggcttgctggattttcccggtgccattgctctggatggtgtcaccaagctggcag atgcggagtgaacctttacgaagcgacacacctgtggtagcgctgcccagaagggagcggagct ccccgccgcgaggcggtcctctctggccaaaagcccagcgttaatagcgccttctgggatgca ggaaccccacctgccaggtgtgaagtggactaagtggatctccaatttggcctgttctgaactacac catctactgctgtgaagaatgtcctgaaggcaagctggttacagccctgatcaggagcccgctcg tgactctcgatcgacgcggggtcaaaaactgtctaagcagcagcagaaacgcgggagcgtttcttt ttcctcatttgtttc
622	말 아르테리티스 바이러스	gctcgaagtgtgtatggtgccatatacggctcaccaccatatacactgcaagaattactattcttgtg ggcccctctcggtaaatcctagagggctttcctctcgttattgcgagattcgtcgttagataacggca agttccctttcttactatcctattttcatcttgtggcttgacgggtcactgccatcgtcgtcgatctctatc aactacccttgcgact
623	엔테로바이러스 종 단리물 CPML	actctggtatcacggtaccttgcacgcctattttataccccttccccatcgtaacttagaagcaacaa acaaactgcccaatagcagcacaacacccagttgtgttaggggcaagcacttctgtttccccggaa gggtctgacggtatgctgtacccacggcagaagtatgacctaccgttaaccggccatgtacttcga gaagcctagtaccattatgaaggttgattgatgttacgctccccagcaaccccagctggtagttctg gtcgatgagtctcggcattccccacgggcgaccgtggccgaggctgcgttggcggccagcctac accatacggtgtaggacgtcaagatactgacatggtgtgaagagcctattgagctacgtggtagtc ctccggcccctgaatgcggctaatcctaactccggagcatccgccagtaagcccactggaagggt gtcgtaatgcgaaagtctggagcggaaccgactactttgggtgtccgtgtttccttgttttacttattgtt tggctgcttatggtgacaacttatagttatcatcataagctacttggtcttgccaaccggagaattattt ggttatttgttggtttcataaacctacagtcgtatttcctgtcttattaattgttctcaaaattaacaaca
624	엔테로바이러스 AN12	taccgctgcaccagtgagctggtacgctagtaccttcgcacggagtagatggcatccccccacccc gtaacttagaagcaaagtacacatctggccaatagtggcgctgcatccagccgcgcaacggtcaa gcacttctgtttccccggtccgcaagggtcgttatccgcccagtccactacggaaagcctactaacc attgaagctatcgagaggttgcgctcggccacgaccccggtggtagctctgagtgatggggctcg caaacaccccgtggtaacacggatgcttgcccgcgcgtgcactcgggttcagcctattggttgtt cacctcaacatagtgtaaatggccaagagcctactgtgctggattggttttcctccggagccgtgaa tgctgctaatcccaacctccgagcgtgtgcgcacaatccagtgttgctacgtcgtaacgcgtaagtt ggaggcggaacagactactttcggtactccgtgtttcttttgttttattttgaattttatggtgacaattgc tgagatttgcgaattagcgactctaccgctgaacattgccctgtactacctaatcgcatttcacaaaa cctcagagatacccaagctcttacattgatctgcttgttttcctgaatctcaaatataaattggaacaagc aaa
625	돌고래 모빌리바이러스	accagacaaagctggctaggggtagaataacagataatgataaattatcatacttaggattaatgat cctatcaattggcacaggatttggataaaggttcacagtc
626	디안케 바이러스	tgttttcaaccataatactactactacaagtataaaaccccgtccgtctgtcggagacgctaaactctg accaccaatctagccacatcagttgcttaaagaacctcttgagacactctcccacttaacatcttttag gaatcttcgatgctacaacaacttggctagtgaacaataaatccgtacaattcacagttgtaagagg ccataggtccagactttgaaaggtttgtttctattgttacaaatacttagattaacagaggctatttaata gtgctcatcacgttaacagagtaaccttgtgcaatagtatgagcttgttgtaaaacgtcttgatacgac acc

[1749]

627	게레자 헤파시바이러스	cactcaatactacactccgcatttggggagaagcgctggcgttcgcggaaccgcgttaaccatac gcgtagtacgagtgcgacagaccccggtgctactggtggtagcgagacacgagccgaagtctgt gggggggaactccacttagagggcatgcccgggcgtaggcttctgagttgggatgggccccaact tggcccctgagtgggggggtgttacgacctgatagggtgcgggctggcgcctaccactttccagt cgtacatgagtc
628	그라제빈 연관 나나바이러스-1	gcccgggggtgcagtcctgtgaaagggtctgcaccatactatatatgtatatgattacatcccaaa aggcgacttcgttcaggttttaaatctgacgtaggtccagtaaataagcatgtcaaaacatgtaagttt atcctgtaatctactctcataagatgagataagatgatattgcagttcccatgtaaataaatccattatg aattcattcatataaggtagaagtggtaactatggttgaaacattaatataaaacggtcattttgcatga acgtcattaaggaactggcataccaatgtctatttagtgactatgatatttagagtatcccttatattaat taacaattattccttttagcatatcatccgacaacaaattttaaaagaagaaatattactcattaaaa
629	엽소 토토바이러스	gtacttacaagcgggttaaaccgcctccggaacggttacaaccccctcccgaacgtgcgcttga cgtgactggttctcagtctggctttctgagaaatactccagggttgtatcccaccatcttgacctctgg tcattttggtaacaccataaccaaacaaactctacacacctaacccacctccagcttgctgcaggcc ttttggactaaacgggtttaggtgttttgtgaccaactcgtctacccacctccagttttactgcaggcct ttttggactaaacggCtttagacttttgtggttagttttaactacccactgtttagccgccaacctgatt ttcattgttgtaaaatttttgtgtttatacactattttcatacggttggcagtttgtttggttgtttgcacagttt ttgttgatacccaatttttactgtgcttttagtgtattctgctaaggctgtattttacatacttagtttggttga agcaatttttacaacatttatattgtttttgatttctaaggtaaagagtcttaggaaacaccatagacgcc attcttgtggtgtctagttccaactaatctggcaaacaagtaccaagtcattgactcactttggagtga gacttacgagtaccaatttgcctatttcggacatccatataaag
630	구제역 바이러스 O 단리물	acaagcttgacaccgcctgtcccggcgttaaagggaagtaaccacaagcttacaaccgcctaccc cggtgttaatgggatgtaaccacaagatacaccttcacccggaagtaaaacggcaaattcacaca gttttgcccgtttttcatgagaaacgggacgtctgcgcacgaaacgcgccgtcgcttgaggaggac ttgtacaaacacgatctaaacaggtttccccaactgacatacaccgtgcaatttgaaactccgcctg gtctttccaggtctagaggggtaacactttgtactgtgcttgactccacgctcggtccactggcgagt gttagtaacagcactggtgcttcgtagcggagcatggtggccgtgggaactcctccttggtaacaa ggacccacgggccgaaagccacgtcctgacggacccaccatgtgtgcaaccccagcacggc aacttttctgtgaaactcactctaaggtgacactgatactggtattcaagtactggtgacaggctaag gatgcccttcaggtaccccgaggtaacacgcgacactcgggatctgagaaggggactggggctt ctgtaaaagcgcccagtttaaaaagcttctatgcctggataggtgaccggaggccggcgccttttcc attataactactgacttt
631	고양이 전염성 복막염 바이러스	acttttaaagtaaagtgagtgtagcgtggctataactcttcttttacttttaactagccttgtgctagatttg tcttcggacaccaactcgaactaaacgaaatatttgtctctctatgaaaccatagaagacaagcgttg attatttcaccagtttggcaatcactcctaggaacggggttgagagaacggcgcaccagggttccg tccctgtttggtaagtcgtctagtattagctgcggcggttccgcccgtcgtagttgggtagaccgggt tccgtcctgtgatctcccctcgccggccgccaggaga
632	파밍턴 바이러스	acgacgcataagcagagaaacataagagactatgttcatagtcaccctgtattcattattgacttttat gacctattattagacccttcacgggtaaatccttctccttgcagttctcgccaagtacctccaaagtca gaacg
633	조류 전염성 기관지염 바이러스	acttaagatagatattaatatatatctattacactagccttgcgctagatttttaacttaacaaaacggac ttaaatacctacagctggtcctcataggtgttccattgcagtgcactttagtgccctggatggcacctg gccacctgtcaggtttttgttattaaaatcttattgttgctggtatcactgcttgttttgccgtgtctcacttt atacatctgttgcttgggctacctagtgtccagcgtcctacgggcgtcgtggctggttcgagtgcga ggaacctctggttcatctagcggtaggcgggtgtgtggaagtagcacttcagacgtaccggttctgt tgtgtgaaatacggggtcacctccccccacatacctctaagggcttttgagcctagcgttgggctac gttctcgcataaggtcggctatacgacgtttgtaggggtagtgccaaacaacccctgaggtgaca ggttctggtggtgtttagtgagcagacatacaatagacagtgacaac
634	인간 리노바이러스 1	ttaaaactgggtgtgggttgttcccacccacaccacccaatgggtgttgtactctgttattccggtaac tttgtacgccagttttccctccccctcccatcctttacgtaacttagaagttttaaatacaagaccaat agtaggcaactctccaggttgtctaaggtcaagcacttctgtttccccggttgatgttgatatgctcca acagggcaaaaacaacagatacgttatccgcaaagtgcctacacagagcttagtaggattctga aagatctttggttggtcgttcagctgcatacccagcagtagaccttgcagatgaggctggacattcc ccactggtaacagtggtccagcctgcgtggctgcctgcgcacctctcatgaggtgtgaagccaaa

[1750]

		gatcggacagggtgtgaagagccgcgtgtgctcactttgagtcctccggcccctgaatgcggcta accttaaacctgcagccatggctcataagccaatgagtttatggtcgtaacgagtaattgcgggatg ggaccgactactttgggtgtccgtgtttcacttttccctttattaattgcttatggtgacaatatatatattg atatatattggcatc
635	EV22	ccttataacccgacttgctgagcttctataggaaaaaaccctttcccagccttggggtggctggtcaa taaaaaccccccatagtaaccaacacctaagacaatttgatcaaccctatgcctggtccccactattc gaaggcaacttgcaataagaagagtggaacaaggatgcttaaagcatagtgtaaatgatcttttcta acctgtattatgtacagggtggcagatggcgtgccataaatctattagtgggataccacgcttgtgg accttatgcccacacagccatcctctagtaagtttgtaaaatgtctggtgagatgtgggaacttattgg aaacaacaatttgcttaatagcatcctagtgccagcggaacaacatctggtaacagatgcctctggg gccaaaagccaaggtttgacagacccattaggattggtttcaaaacctgaattgttgtggaagatatt cagtacctatcaatctggtagtggtgcaaacactagttgtaaggcccacgaaggatgcccagaag gtacccgcaggtaacaagagacactgtggatctgatctggggccaactacctctatcaggtgagtt agttaaaaaacgtctagtgggccaaacccaggggggatccctggtttccttttattgttaatattgaca tt
636	인간 TMEV- 유사 카디오바이러스	tccgacgtggttggaattaacatcttttccgacgaaagtgctattatgcctccccgattgtgtgatgctt tctgccctgctgggcggagcgtcctcgggttgagaaaccttgaatcttttccttggagccttggctc ccccggtctaagccgcttggaatatgacagggttattttccaaactctttatttctactttcatgggttct atccatgaaaagggtatgtgttgccccttccttctttggagaatctgcgcggcggtctttccgtctctc aacaggcgtggatgcaacatgccggaaacggtgaagaaaacagttttctgtggaaatttagagtg gacatcgaaacagctgtagcgacctcacagtagcagcggattcccctcttggcgacaagagcctc tgcggccaaaagccccgtggataagatccactgctgtgagcggtgcaaccccagcaccctggttc gatggccattctctatggaaccagaaaatggttttctcaagccctccggtagagaagccaagaatgt cctgaaggtaccccgcgcgcgggatctgatcaggagaccaattggcagtgctttacgctgccactt tggtttaaaaactgtcacagcttctccaaaccaagtggtcttggttttccaattttgttgactgacaat
637	인간 코로나바이러스 229E	acttaagtaccttatctatctacagatagaaaagttgctttttagactttgtgtctacttttctcaactaaac gaaatttttgctatggccggcatctttgatgctggagtcgtagtgtaattgaaatttcatttgggttgca acagtttggaagcaagtgctgtgtgtcctagtctaagggtttcgtgttccgtcacgagattccattcta caaacgccttactcgaggttccgtctcgtgtttgtgtggaagcaaagttctgtctttgtggaaaccagt aactgttccta
638	후베이 자오바이러스- 유사 바이러스 1	gtgcaggatggcctttcccatcttaagtggtttgtaggatttcgtgggtccataccccccgatttcttg gtacgtattccatgcacggagaatacgaccaaaactcttatttcaaaaaatattatttttttactcttgtgg gctgagtgcgacccaccagttccagcttagcaacctggaagttgttggagatttatggaaccaaatt acacatgcgtggagtgccgccactccgtatctgttcactcattacgcgattaagttctgcgacgaga cgagcgaa
639	후베이 톰버스- 유사 바이러스 9	ggaccatccaggcaggtgtaggctagtaccctcacctgacctgtcgcgatgtttggctttgtgagg cttgtgggaggatcccttggcccatgcattgctgtcatcgtgaaaaatgttgtatgctcgcacct ggcgtggaggaaacggcatttgacggatgctaaggaggatttgaaggtcctcaactcccatgcctt attacaagtccccttttccttcaagcttcgagcccacgtctgtgacgagcagaggtgaggttggagtc aagaggagtcctattcaagctgttcgcagcaagaaccataaaactttcattgctcaatgggcaaga gcggctaaggctcggttctcgtttgcacgacagtgtgagcggagccctgtaaatgtcgcggccctt catcggtggtttgcatcggagtttaagaaaattggcatgaacttgctccaggtttcgtacgtgatcgat gaggttgttgaacttagtttggaaccaacgtatgagcgtatggtgtccgaaactaagaagcaattcc gtcatcgggcacgtatggaatactacaacgagaaaaaatgccttgaaaagatccactaggaacgc
640	후베이 톰버스- 유사 바이러스 32	ttcgggtttacccgcgtaagcggccacactgactggttgtcggtgttgaatttgtataccagatgagg agacgttaccaccgtctcggcagtgctacgtctgggaaaggactgtgatagtggacagtcctacc gtctttagatacattgcgttgtgtatagcccgggagggattaactaatagcaacgcaatgcacacgg cggttcggatttgcttgactgatggaaagacatctaagttctaattgaacc
641	후베이 소베모- 유사 바이러스 3	gagatgtttgcgtgggccgttgcgctgcgggcggcccacctccctacggggaccgtgagacacc gctggggaaggccccccaccccgggccaaggggatcctgccgagaggcaggagaaagaggcc cagccctctggggcgcctttaggggTgcctgggagggaagtacccgagccgggcggccggtc gggtgcggctgtgcagttcgaggctaaccgtaaggaaggcctgagctgcctcggcttgtcggaa aggaagacgaaggcacttatcaaggctttcaggaagcaggagcgcaattacagcgcgcgccgt gcggcctggattaaccgcatttggccg

[1751]

642	후베이 피코르나-유사 바이러스 2	agcaacttctttctgaaaactagctagagttcgacgatctctctggctaatgacaaataaccaatcaa aaagtcaaatgttcatgtatatatatatttagtagtgaccttatttagaaaaactttagatgatttatcgtc aagttgccctttgtgaagcgatcagctttttatatcgtttcatttttgtatacgtcttaattgacgttgtaag tacgttttgcatacctcacattgaggatagtatcgtttcctgactaagaagttaaactagtctaccaata gcaaccatataggatatagctttgtttttaacaaggtttaatctgatcttatgctcttgcttacggtgtttta tgtatagaaaaGtttttataaaaactacataattgtcaaaagaaaaagcgttacgtactgacgcattta tgttcacagtgtgacacaaaccttctatattttgattgtaaaataggctttgcctgaccatttatcaaata caaactgtttcaaacgcctctccgagccatattggccgacgcgaatcgacataacagggtgagttt acagctgcagttgcagccgaggatccacttatttgagagagaattactcttaacgaaattttgaagtc acttctacacagttaagtctgagtaggcgttatccaaacgtaaa
643	헤파시바이러스 P	acatgggggggggctgacagtgagtacactgtgccaagcaggtgctacgctatgcctaggtgct gctgtaggccaaggacatgtccagtcatcccaggtgagggggggggttcccctcaccgctgcc actgcctgatagggtcctgccggagggtctcggtgtccggctgtac
644	헤리어 피코르나바이러스 1	gatgtgtgacggtgtaattacttccggatcccacttcctattataactctttcatcccaaggttaggg aaagaacctggctcggtaccaccagaccctccgccacgctagtggactctccggagataacggt acccccttgtagtcacctgtgctggtgaagaaccacctagtattgcttgggtgcgtgccgcctagct tccatttcttctggagcactgtgcaatgaggtttccccacttggtaacaagtgcctcaggtcccgcaa ggatactgtggggtggtgtgaccgcagggagctgtctccacggctcctctaatgttacgccgctat ccacaggccagtgcgtgtcatcgatcccggatgacagagctagtattgcgaacccccaagtaaga aaagtggctagtaacctgatagctggtgaagagggtggtcagttgagtagatgccctagaggta cccgaaaggatctgactagggacccgtgactatacattaggtaaaccgggtataaaaaccatgaa aaactgaccactttttcttttaacctcttctttctttttatgtgttaagtgttttgagtttggactgtaccttg cccgcctttcttggattttctctatcgctcttcttacacctactgttatcaaggcactctttagagata
645	쿤사기시바이러스 1	tttcaaatcggactccggtagttatacccggagcccggtttggacgcttgggccgcgttaacagccc ccacccctttcccactgactgatttctcggattggactcatttgcattgctaactctgattctggatttc ccgtttatgtcgtcgcggtcggaagtgcacgtacttcgacgttgatctgatggcgtttgtttccagg ggggaggtggcggcagaaacgcccccgccgtaaacttcggcgggccacgcctgtcaagccact ccctggggccgagcgcctgaggtgatacagagagataagcacactgggcgctgacaacgcccg ggacctcagtgagaagagcagtagggccgtgtttatgggactccattggatatcccccgcttgtcg gaactcacggctactccgggttgggaagcccgcgactggtactgtactgggtgatagcctggtgc cttccctctcactgttgtatgaaggctgaaaacccccct
646	카고시마-2-24-KoV	tatagtttgcctgtttctcgcaccgttaccgctcgttcgggaatgtgaaactggcacccctcctctccc ctaccacccttctccttcgccccattcataatttacaacgccgcacacagcggcggccgccaag ggctagcctggcggttataaaaggaacctgggtctttccctcttcaagccaaaaggtaggttccctg tgtccctgaatgctcggtgaggaatgctgcaccgtaacgctttgtgaagtgtttgcaagttctggccc ggcaagcctacagagtgctgtgatccgctgcggacgccatcctggtaacaggacccccagtgtg cgcaacagtatgttcagacttcggtttgttcacttgctttcatggaccattgcgcgaaagtgcgtgcg ccatatcccctgtacttcaggtgtgcttctctggaccctaggaatgctgcgaaggtaccccgtttcggc gggatctgatcgcaggctaattgtctatgggttcagtttccttttttctttactccacaattgactgcttaa ctgactctggatcttgtgcttccactgctctttctgctctcaaaacggcttcacttaccaactctcaccctt tcgaccaacaccatttacacactaacttttttcgactcttctgactcctggcttggtgaagac
647	캐시미어 비 바이러스	tacgtacaattttgacgcttcgttcatgcaacaatgaactcacatgtggcgctcggtagtaaccagag gggcgtcattccccgtatggagtggagaatataagctaccgactcgagctgtagaataattcagc aacttataacgaacacgaattttagtcgacgaaaccattttagccttaattctatgatttgattataatg ataaacagctcatgtaactgtctaaactacataaatacaactggattacgaacctttaagtaactatctt agatgaagtctagtagtctcccaatatttccgtgaaaagaatgagggacgagatagctctatttaaa gacgtgaggctttaaattctgataaatacattacctgagaattcctccttttaggagttagaatttgaaaa gattagtacctatacttaatagaatttaaatatgttaataatgctgaagacaagtatttcgattattaacct ctatatttctatataaagtatctgtgagtctcagtggacatcacagtaaggtcgcagttaacttgtaatc ttttcattcctgtgtcggagcagtggtaatggagccggacgatttcgccaaaac
648	징먼 피코르나-유사 바이러스	tgtgtttttgtcaagataattgttctgtgattaacagtgattgtggttcgtgtaatgcgacgcagtcaaat gctagttttgatgaagtgtatgagagagtggaaaacttatctcataagaagattgaagagtgtgtaga tcaggctattgatcgagcttctaagcttcgtgattacaagcttaatgttcacaatggctcccgacggg aatcatctgatcctctctttatttcgccccattcgttgttatcgcttggggtatctaagtttgttgcgtttga

[1752]

		gcagcatcacagtttcgcttcagttgagtctctgaagttgcttgctctgtct
649	멈프스 바이러스	accaaggggaaaatgaagatgggatgttggtagaacaaatagtgtaagaaacagtaagcccggaagtggtgttttgcgatttcgaggccgggctcgatcctcaccttcattgtcgataggggacattttgacactacctggaaa
650	마우스 모사바이러스	gaagttgatcatgaacttggttattggtggaacgcacatgaactcccaacaatgatcttgaagacacagcgtggtaacaattaccatgcccttgtggctgcccaagacattgatggctattgggtgatttatgatgac
651	미니옵테루스 슈라이버시이 피코르나바이러스 1	gttgtcgacccgttgcttggtttaagcatgaggtggattccccccgattatgtctacccgttactatggcgggcggtcgtttcagggtgtttctactgaggactgcaccaagtctttatctttattctcagatctccggctgtttgacgcttgtaggacagcaggactattttctcttaatctttttctacccactagggtcctatcctagtggagggagggtgccaccccttctctctttagagagtgcgcctggcggtctttccgtctctggaaaaaggagcacatggcatgctacaattggcacaagaaaacaagctttgcggattctttctttgtactagaggaagctgtagcgaccctgtatggcgagcggactcccctctcggcgacgagagcctctcgggccaaaagccaagtgttaatagcacccatacaggcggcagtaccccactgccctttctcaacatacaatgactgatgaaccactgttggtttttctgacacctttgtagtaggattccaaggaatgtcctgaaggtacctgttagcttacgcgcaggatctgatcaggagtctctttacagtgctgtacactgtgcaagggattaaaaattgtttgaggaatccccgagatagtggtctatctcttcctattttgttttacagacacg
652	린다 바이러스	gtatagcagcagtagctcaaggctgctatacgattggacatacaaattccaattggtgttagggaccacctaggtgaaggccgacgacaggtagccattcctgttagtaggacgaaccgttatggtggactggttgctcaggtgagcaggctgcaatgcgtaagtggtgagtacaccacagccgtcaaaggtgccactggtaaggatcacccactggcgatgccttgtggacgggggcgtgcccaacgcaatgttagcggtggcgggggctgccatcgtgaaagctaggtcttgatggaccttgttgcctgtacagtctgataggatgccggcggatgccctgtgacagccagtataaagaatatccgttgtgattgcac
653	레사바이러스 2	tctttctttatttttcttatgtaactcttctttttaagttttattttgcctacttgtgagcttatgcgggaccactgtcttagacaaccccacatttgtcatgagtaagtacacgcaaccattacgattacttttttaaccgtctgaccttttgataacaactgaagttaggcgtgaaacatgcatttatacccaaagtagccccgcatttccccactacggtgggggggctaccctactggctttggaactgtagccattatgtgttgcctggctttcaggatctcacaacacaacagttctctcacaatggaatatgggtgagattgcagtgacatgaacaagtatctagtagtacatagactcaagcctagttgcctgcggaacaacatgtggtaacacatgccccagggtccaaaagacaagggttaacagccccttctaggtgtctgtgtgtgaagaatactttagtagtgttgttatgatctcacctgttagtacagaatgagtatggcttggtgaaggatgtcctacaggtacccattatatggatctgagtaggagaccactagtggtggctttaccgccaggtgagtggtttaaaaagcgtctagccaagccaacagcactagggatagtgctttctatattttatattttcagtgtatatggtgacaa
654	레사바이러스 1	gtaactaataagcaagttttactgcctgcaaactgcttcaatgggaccaccgcttcggcgacccctttgttgagtttgtatgttttaagtaatctttgcaaccatacgattattttagccgcctctctataatgatcttgttatagtgggacgtgaaacattggtttttctcacacacgtccggtcacccgggcgtgtgttcttccgtaagtcctatccacatacctgtgggtaggccagcatgtttgcacaggctgtgacagtgtgggtgggctttccacctctcaacaacacactgaattgcaatgcactcacggaggaaatgacaatttggttatagttttgaactgtgctagtaatttctttcacattaagccatgttgcctgcggaatcacatgtggtaacacatgcctcagggcccaaaaggcacgggttagcagccccttcatggtgtgttagaagtgaaaacacatagtatgagctataatatctttgttgtcttctctgtagtgtaccccgccaaatgtaaggatgcccagcaggtacccatttatggatctgagctggggattgatagtgtatctataaatgcactgatcaatttaaaaagcgtctaagtttggcacaaacactggggacagtgtttttccttattctttttatttgatta
655	포피바이러스 균주 뉴잉글랜드	gggagtaaacctcaccaccgtttgccgtggtttacggctacctatttttggatgtaaatattaattcctgcaggttcaggtctcttgaattatgtccacgctagtggcactctcttacccataagtgacgccttagcggaacctttctacacttgatgtggttaggggttacattatttccctgggccttctttggccctttttcccctgcactatcattctttcttccgggctctcagcatgccaatgttccgaccggtgcgcccgccggggttaactccatggttagcatggagctgtaggccctaaaagtgctgacactggaactggactattgaagcatacactgttaactgaaacatgtaactccaatcgatcttctacaaggggtaggctacgggtgaaacccctttaggttaatactcatattgagagatacttctgataggttaaggttgctggataatggtgagtttaacgacaaaaaccattcaacagctgtgggccaacctcatcaggtagatgctttggagccaagtgcgtaggggtgtgtgtggaaatgcttcagtggaaggtgccctcccgaaaggtcgtaggggtaatcaggggcagttaggtttccacaattacaatttgaa
656	케스티바이러스	gtatacgagattagctcatactcgtgtacaaattggacgtagcaaatttaaaaattcggatagggtcc

[1753]

	균주 아이덴	ccatccagcgacggccgaacggggttaaccatacctctagtaggactagcagacggatggacta gccacagtggtgagctccctgggaggtctaagttctgagtacagaacagtcgtcagtagttcaacg ctggtaaaccccagccttgagatgctacgtggacgagggcatgcccaagacacacttaacctgg acgggggtcgtccaggtgaaagtacccatctttgggtgctgggagtacagcctgatagggcgctg cagaggcccactgacaggctagtataaaaatctctgctgtacatggaac
657	메추라기 피코르나바이러스 QPV1	tttgcatcagttcgcccctccctcaccatacctttttccccctctttaggactgatacttggttatgatga gcagaggatttcgcaagttatgcttcttgataaaaagtaattcacgaatcatgggattatagcctgga agtgaacactcatgtggcaagtgggttagtagctctccatgttcccatgtgcagtggactgacaaca gtgagttcggggttgtgtagtaaagggaaagtattacttacccgcacctgctatacgtggtgtacgta ggatacgagttagtagtgcttagcaactttaaactggtgctgaaatattgcaaggtcactgaagttgt gaacgcgaacgctccgccactgccatgtatagcgtgcaatgcataaatggtgcactacatgatacg agggaatgggaaaccctccatggccgaatgcagggtgacagcctgccggcggatgcctgttgtt agtataatccgttgtttgccac
658	돼지 사펠로바이러스 1	acactcatttccccctccacccttaaggtggttgtatcccctacaccctaccctccttccacatagg acgaataaacggacttgagattaaggcaagtacataaggtatggttttggatacacttaaatggca gtagcgtggcgagctatggaaaaatcgcaattgtcgatagccatgttagcgacgcgcttcggcgtg ctcctttggtgattcggcgactggttacaggagagtagacagtgagctatgggcaaacccctacag tattacttaggggaatgtgcaattgagacttgacgagcgtctctttgagatgtggcgcatgctcttgg cattaccatagtgagcttccaggttgggaaacctggactgggtctatactgcctgatagggtcgcg gctggccgcctgtaactagtatagtcagttgaaaaccccc
659	돼지 생식기 호흡기 증후군 바이러스 2	atgacgtataggtgttggctctatgccttggcatttgtattgtcaggagctgtgaccattggcacagc ccaaaacttgctgcacagaaacaccctctgtgatagcctccttcaggggagcttagggtttgtccct agcaccttgcttccggagttgcactgctttacggtctctccaccccttttaacc
660	돼지 엔테로바이러스 9	Gaaccttagaagtttacacaaacaaagaccaataggagtccaacacccagttggattgcggtcaa gcacttctgtttccccggacctagtagtgataggctgtacccacggccgaagatgaacccgtccgtt atccggctacctacttcgggaagcctagtaacattctgaagtctctgaggcgtttcgctcagcacga ccccggtgtagatcgggctgatgggtctccgcatacccccacgggcgaccgtggcggaggccgc gttggcggcccgcctatggcgaaagccataggacgcctcttagatgacagggtgtgaagagcct actgagctgggtagtagtcctccggccccctgaatgcggctaatcctaaccacggagcgtccacca gcaatccagctggcagggcgtcgtaacgggcaactctgtggcggaaccgactactttgggtgtcc gtgtttccttttgatcctatttggctgcttatggtgacaacgataagttgttatcataaagctcttgggtt ggccacctggaaaaagttatcagtgtttgatattgttcggctctcacgcctaccaataagacaagcc ctatatttacttgttgcattttactcgtcagaagaaatcacagagtatcatttggatttgttactcacatta aggacaag
661	비둘기 피코르나바이러스 B	tttatttagctgttaagttttatttgtgccgagAccccatagtaggatcttggtgttccacattaagctct Cccgaccacacatccaaacgataggcggtgtaagggctccctggctaagtgtttactcattgctag ggaagtgttgcgacccgttaccagtaggaatacaggaggtcttagttgcctaaccagataaagtgg tgctgaaatattgcaagctcaatgtctggcgaacgacggactaccgttgaactattgttaacgcccg cgtgtgtaggcaacacacgggttagtaggtcacttacattgacatccgtgccgggaaagcggatct gagctatcgattgcctgatagggtgcccggcgggcgcggtacgtgtggtatagtccgctgtcttggg gtatggcgtctactcggttttgttgtcgttttggaatgtcccatgtcggagatgtcgttaccggtgtttg ctttacttgtgcacgaataagaaaacagagtttggagttttggaagttaatggataacatcaagtttcc attgcgttcctcgtgcattttaggccagaaaatcgaccacaaaatcgaga
662	피코르나바이러스 HK21	aaacgggaggatcggctttggcttatcctcttaatagtctacaaaactgggctgactggtggggga gctactagatccgggagaggactagttaccccgcgtaacttccctttttgtcactccctccacctacc ccttccctgtcccttagactcttaagtaggtgtgcaggcctggccccccggaaatgggcaaatcgg acgtttgcgtgtagggtcgttctgaaaggtggggcccactccaccgtagtaggatcttcctgtttcac gtttaaacctctccgggcaggtatcgctagacactaggctgtataggatggcacgcacttaggtttc gcaccttcctagtgcgccaagatcccgcttttgagctcaagtacatggttctttgtaagatgtctaacc agaggaggtggtgctgaaatattgcaagccactctggcgaacgtccactgcattgcctggaacag gctctctagcgccctccactggtagcgcggtggtgggttagtaggatacctatatggacaggggat gcgggaataccctcactagctagtgactggttgatcgactggcggcggatccagtgatacttgca taatccgcagacttgggag
663	피코르나비랄레	agccatgaactagtgtgcgattcccacagtgttggtcgaaagccccgcactgtctactcgcatattt

[1754]

	스토토리-HG1	gactccagtccttccgcagccagcggttaggttctggttaaagtgcattatgtgcacggcgccacg cagaactgccagaaatggtaagctgcgcccaacgccaacggtttgtgtatcccgttagtcacacgt ttacagctgttcgttaacagtagggttttgtcgacccggacccctttaatgcgcgcgaattcaaccgc gcctagtcggcaatggtatcatttaatcccatgcactacgggagaaatttgagaccaaagaattcct gagggccactgcttgctctaagtgcaatgcctcgggagacttctgtcaggagcctagcggctttca accgcgacagctaactcctgcgggatgtttggtgtccatacttactggcgtcctcacaacgctaagt ggatgttgtccacaggtaggcaaacaccgagccccacattcaggagacctgtatgaacgatcctat cagcattagagttggaattgggtgtgctaacgtccgcataagtgcaccccgtggtaacgctgggaa actatccagcgcaacgtactgtcctcaatgtctagggaaggaccgccctaagcgtacaaccgggc catgtgtcgagc
664	설치류 헤파토바이러스	ttcaaagaggggtccgggatttttcctggtccccctctttgggcacccttggctcgggggtgtgaata ccgtgctcgcgtttgccgtgcgttaacggcttcatttatgtttgtttgtctgttttattatgttGgtttgtct gtttgttatgttggtattgttcgtgtttaatgttatgaccacattacactccagccaatgaagaacagatg gtgcggttattgctggcggaattcctaacgtcctggatccgttggtacgcatcacaaaacaatttgca gagagagtggtgaaacggcttgggaatccctgagtacagggaaatcacactgatagctcatcttg gctgttttcagtcatggaccttatgcagtgtaatttgggtgtacccccatagcttaggaggaatgttc tgtcttggcactagagtgggacgctgatgcctccgtgtctaggatggtctaagggacagaatggggg tgcctctgatgccatactacctgatagggtgctctcacggcctctgcatcttagtgagaagttcaatttt
665	린더케스트 바이러스	accaaacaaagttgggtaaggatcggtctatcaatgattatgatttagcacacttaggattcaagatc ctatcgactggagcaggcttaaggtaaaggttctttaaa
666	라보바이러스 A	ctacggatatttgcatgacccgctttctatcgccccaacaatccccttttgtaaccacaagctttactca ggctagcagcccgactagctgtttggaagaaaaggctagggcacacaccaacaacaccgaccc cactggtcgaaggccgcttggcaataagactggtggaacagggtcgcctgtagttgtttggaacat tctttctaatgactttgtcagcggtgctactcacaccgtaactcttctaccctatccccacgcttgtgga actaggaggggatgagtgattcaagtaagtactgtcagaatggtgaaaatgatctgattctgaaacg ctatggatccatcgaaagatggggctacacgcctgcggaacaacacatggtaacatgtgccccag gggccgaaagccacggtgataggatcacccgtgtagtttgagatcatatcaatgttcatagtctagt aagatgatttgaaatctaactgagctgatggctaactgcttgtcttattgcggcctaaggatgtcctgc aggtacctttagataaccttaagagactattgatctgagcaggagccaaagtggtctttcccagcttt ggttaaaaaacgtctaagccgcggcaggggggcgggaggccccctttcctcccaaaacttaatatt gattgt
667	싱글백 니도바이러스 1	ctgtgagtaccgacaggctcgaagtctattatgaggcgtcgaaacagaaaacctgtaacaactccg gtttcatctatcactgccgtcaagaggcagaagaggacgaccacgtgtcaccagatcacttgtatct gtttcagtcaggaagtcaacttttcgacgaagttcgaccattcatcgacccgctgaaaagcgtagaa gtcgatgaagatgagctccaaagagccatcgccgatttcgacaaccaaagtgactgtttccactcct tcgagctcgtgaatttcgagctgaaacaacaaatcaacgagaacgagtggtacggttattataatta cgacaaccaaaactgcaaagttcagttgccagtcacatgtcgaatcgaggacgtaacctgggatc aggtttacgtg
668	세네카 밸리 바이러스	tttgaaatgggggctgggccctgatgcccagtccttcctttccccttccgggggttaaccggctg tgtttgctagaggcacagaggggcaacatccaacctgcttttgcggggaacggtgcggctccgatt cctgcgtcgccaaaggtgttagcgcacccaaacggcgcacctaccaatgttattggtgtggtctgc gagttctagcctactcgtttctcccccgaccattcactcacccacgaaaagtgtgttgtaaccataag atttaacccccgcacgggatgtgcgataaccgtaagactggctcaagcgcggaaagcgctgtaac cacatgctgttagtcccttatggctgcaagatggctacccacctcggatcactgaactggagctcg accctccttagtaagggaaccgagaggccttcgtgcaacaagctccgacacagagtccacgtga ctgctaccaccatgagtacatggttctcccctctcgacccaggacttcttttgaatatccacggctcg atccagagggtggggcatgacccctagcatagcgagctacagcgggaactgtagctaggccttа gcgtgccttggatactgcctgatagggcgacggcctagtcgtgtcggttctataggtagcacatac aaat
669	스클레로티니아 스클레로티오럼 dsRNA 마이코바이러스	ttgaattaatcttttacgtttacgcgcataaaatcaggacacatctcttgtatactttagtatatcaatgat gtttttgttttatgcgattaatcgtaagagaacttctttccatccgcctgtatgggcgggataataagttc accgccttggtcgaggcgcaaacttgtatgtgcaaaggtgagctatatgctcgaaatagtcgtaact aacacacagccactacctgtagagctctattgatccggaatcctttagtgggaatgcagagctcaca

[1755]

	-L	ccggacctgcgggtatcttcggcgttagggacttctgtttcagccttgaatcatttacctttataccttct ctgaggcgcctgggccgggcgcgatattaagtacaagtcaaggacatcgcgggtagtggtctaat cagccgctagtcctgctggagagttccaacttagttgggtgtggtgcatactagctggatagagtag gtatgtattgctaacgtatgccggaggctatccgtcctcggtagaacgtgccgaggagtagtctctg cagacccccgaacgcgtggggtctttacttaaatgtaggcggagggagcgctcgtaggtggaac gactgcctcccagtcgaatgcaagattttgcacgcggaccagtctgcccggcaattcccgggtg
670	야크 엔테로바이러스	ctccggcacagccgcaccagtgcactggtacgctagtacctttcacggggtagtcggtatccccc cccgtaacttagaagcatgtaacaaaccgaccaataggtgcgcggcagccagctgcgttgcggtc aagcacttctgtctccccggtccgcaaggatcgttacccgcccactccactacgaggagcctagta actggccaagtgattgcggagttgcgttcagccacaaccccagtggtagctctggaagatggggc tcgcacatccccgtggtaacacggttgcttgcccgcgtgtgcttccgggttcagtctccgactgttc acttcaacatcacgcaaccagccaagagccgattgtgctggagtggtcttcctccggggccgtga atgctgctaatcctaacctccgagcgtgtgcgcacaatccagtgttgctacgtcgtaacgcgtaagt tggaggcggaacagactactttcggcaccccgtgtttccttttatttttattcttatttttatggtgacaattg cagagatttgtgatattgcgactttaccgttaaacatagcactgcattacctggttgcattccacaaaa cttcagagattcctagttcctacattgacctacttgtttatttgaatcttaaatacaaacttgagcaagtg aa
671	워블리 포섬(Wobbly possum) 질병 바이러스	cggctgtgagtgcttagcatatgctagagtactacagccgggtgttggagtcatatgcactggttgc ctgtataatagtcgggatctgtctgacctacattatctttgggagttgcttatcacgacaattctcgaag tgtctgtcgacagcttacccccgattcgacaaggccccttgtccaccgcagacctatcgattttcaac gagaacactatcagaggtttaaatttaaaactcaccaaca
672	조류 오쏘레오바이러스 분절 S1	gctttttcaatcccttgtgtcgatgttcccgtatgtcatctggttcatgtaacggtgcaacttctatttttg gtaacgttcactgtcaggcagcgcaaaactcggcgggtggtgatcttcaagcgacttcctctcttgtt gcttattggccctacttggctgcgggcggtggtctgctcgttgttattattataattgttggcgctgtttg ttgttgcaaggctaaggttaaagcggacgcagcccgaaacgttttctaccgagagctgttcgcactt aattcgggtaaaagtgatgcaggacctccgatttaccaggtttagtgtacgacgatttgagttttcac ctttcgtcttagaggagtgcactactccatctttcacgactataactaataccgatccggctctctactt taacattgagtttccgtcaagtcatcgtctctccccctttcattccagaactgttgtctcagccttgtacc gttcacgtttcattgattcggagattcgctctctgtgcaaccttatctagtatttgtgaatacgactgtgc gctactgccatccatcaacgctattacgacgatccctacaccaggtgcgtcatcatctctgattgttc attggg
673	염소 코브바이러스 d10	ggggcctcggccccctcaccctcttttccggtggccacgcccgggccaccgatacttcccttcact ccttcgggactgttggggaggaacacaacagggctccctgttttcccattccttccccctttccca accccaaccgccgtatctggtggcggcaagacacacgggtctttccctctaaagcacaattgtgtg tgtgtcccaggtcctcctgcgtacggtgcgggagtgctcccacccaactgttgtaagcctgtccaa cgcgtcgtcctggcaagactatgacgtcgcatgttccgctgcggatgccgaccgggtaaccggtt ccccagtgtgtgtagtgcgatcttccaggtcctcctggttggcgttgtccagaaactgcttcaggtaa gtggggtgtgcccaatccctacaaaggttgattctttcaccaccttaggaatgctccggaggtaccc cagcaacagctgggatctgaccggaggctaattgtctacgggtggtgtttccttttctttcacacaa ctctactgctgacaactcactgactatccacttgctctcttgtgcctttctgctctggttcaagttccttg attgttttgactgcttttcactgcttttcttctcacaatccttgctcagttcaaagtc
674	염소 코브바이러스 d20	ccccctcaccctcttttccggtggccacgcccgggccaccgatacttcccttcactccttcgggact gttggggaggaacacaacagggctccctgttttcccattccttccccctttcccaaccccaaccg ccgtatctggtggcggcaagacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccagg tcctcctgcgtacggtgcgggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcct ggcaagactatgacgtcgcatgttccgctgcggatgccgaccgggtaaccggttccccagtgtgt gtagtgcgatcttccaggtcctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtg cccaatccctacaaaggttgattctttcaccaccttaggaatgctccggaggtaccccagcaacag ctgggatctgaccggaggctaattgtctacgggtggtgtttccttttctttcacacaactctactgct gacaactcactgactatccacttgctctcttgtgcctttctgctctggttcaagttccttgattgtttttga ctgcttttcactgcttttcttctcacaatccttgctcagttcaaagtc
675	염소 코브바이러스	ctcttttccggtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggagg aacacaacagggctcccctgttttcccattccttccccctttcccaaccccaaccgccgtatctggt

[1756]

	d30	ggcggcaagacacacgggtctttccctctaaagcacaattgtgtgtgtcccaggtcctcctgcgt acggtgcgggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactat gacgtcgcatgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatct tccaggtcctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctac aaaggttgattctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgac cggaggctaattgtctacgggtggtgtttccttttttctttttcacacaactctactgctgacaactcactg actatccacttgctctcttgtgcctttctgctctggttcaagttccttgattgtttttgactgcttttcactgc ttttcttctcacaatccttgctcagttcaaagtc
676	염소 코브바이러스 d40	gtggccacgcccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaac agggctcccctgttttcccattccttccccctttttcccaaccccaaccgccgtatctggtggcggcaa gacacacgggtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcg ggagtgctcccacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtc ctcctggttggcgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttg attctttcaccaccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggct aattgtctacgggtggtgtttccttttttctttttcacacaactctactgctgacaactcactgactatccact tgctctcttgtgcctttctgctctggttcaagttccttgattgtttttgactgctttcactgctttcttctca caatccttgctcagttcaaagtc
677	염소 코브바이러스 d50	ccgggccaccgatacttcccttcactccttcgggactgttggggaggaacacaacagggctcccc tgttttcccattccttccccctttcccaaccccaaccgccgtatctggtggcggcaagacacacgg gtctttccctctaaagcacaattgtgtgtgtgtcccaggtcctcctgcgtacggtgcgggagtgctcc cacccaactgttgtaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcatgttccgctg cggatgccgaccgggtaaccggttccccagtgtgtgtagtgcgatcttccaggtcctcctggttgg cgttgtccagaaactgcttcaggtaagtggggtgtgcccaatccctacaaaggttgattctttcacca ccttaggaatgctccggaggtaccccagcaacagctgggatctgaccggaggctaattgtctacg ggtggtgtttccttttttctttttcacacaactctactgctgacaactcactgactatccacttgctctcttgt gcctttctgctctggttcaagttccttgattgtttttgactgctttcactgctttcttctcacaatccttgct cagttcaaagtc
678	피코르나비랄레스 종 단리물 RtMruf-PicoV	tttgctcagcgtaacttctccgggttacgtggagaccaaaaggctacggagactcgggctacggc cctggagcacctaggtgctcctaaagacgttagaagttgtacaaactcgcccaatagggccccccc aaccaggggggtagcgggcaagcacttctgtttccccggtatgatctcataggctgtacccacgg ctgaaagagagattatcgttacccgcctcactacttcgagaagcccagtaatggttcatgaagttgat ctcgttgacccggtgtttcccccacaccagaaacctgtgatgggggtggtcatcccggtcatggcg acatgacggacctccccgcgccggcacagggcctcttcggaggacgagtgacatggattcaacc gtgaagagcctattgagctagtgttgattcctccgcccccgtgaatgcggctaatcccaactccgga gcaggcgggcccaaaccagggtctggcctgtcgtaacgcgaaagtctggagcggaaccgacta ctttcgggaaggcgtgtttccttttgttccttttatcaagttttatggtgacaactcctggtagacgtttttat tgcgtttattgagagatttccaacaattgaacagactagaaccacttgttttatcaaaccctcacagaaa taagataaca
679	아포데루스 아그라리우스피코르나바이러스 균주 통루안-Aa118	ttactcagcgtaactactccgggttacgtgatgaagaagaggctacggagattctcgggctacggc cctggagccactccggctcctaaagatttagaagtttgagcacacccgcccactagggccccca tccaggggggcaacgggcaagcacttctgtttccccggtatgatctgataggctgtaaccacggct gaaacagagattatcgttatccgcttcactacttcgagaagcctagtaatgatgggtgaaattgaatc cgttgatccggtgtctcccccacaccagaaactcatgatgagggttgccatcccggctacggcga cgtagcgggcatccctgcgctggcatgaggcctcttaggaggacggatgatatggatcttgtcgtg aagagcctattgagctagtgtcgactcctccgcccccgtgaatgcggctaatcctaaccccggagc aggtgggtccaatccagggcctggcctgtcgtaatgcgtaagtctgggacggaaccgactactttc gggaaggcgtgtttccatttgttcattatttgtgtgtttatggtgacaactctgggtaaacgttctattgc gtttattgagagattcccaacaattgaacaaacgagaactacctgttttattaaatttacacagagaag aattaca
680	니비벤터 콘푸시아누스피코르나바이러스	ccctttcataacccccccctttttaacccaaccctttcgtaaccgtacgcttcactcgcctttgggtatag cggcccaatgtgctgaagaaaaggatacgctataaggggccaacgggtggtggcccttaagacc acccaacctagaagcttgtacactcgggcaatagtgaggcccacatccagtgggtcaagcccaaa gcattcttgttccccggtatgatctcataagctgtacccacggctgaaagagtgattatcgttatccca

[1757]

		ctcagtacttcggagagcctagtacaccacttggaaatggaagtctgtgatccggggttgaccctg aaccccagaaactcatgatgaggctaaccttcccgaacacggcgacgtgtggttagcctgcgctg gcatgaggcctctttgtaggcagactgaaatggaagggtgacgaagagccgactgagctactgttt tattcctccggccccctgaatgcggctaatcctaactcctggtccagtacttgtaacccaacaggtg gctggtcgtaatgcgtaagccgggagcggaaccgactactttggggcgtccgtgtttctcaatatta ttcatttctagcttatggtgacaatttatgattgcagagattgtgctgtatttgtgtctgagagaagaagt aacaat
681	박쥐 피코르나바이러스 단리물 BtRs-PicoV	tttcaaaaggccctgggcatacggcgttattcgtaacgtcgtatgtccagggcggtagcatcaggc caaggcctgatgctaccacgtgtggactaaaccacacactcttcttgtgacacgttgtgtcacctatc cctttcttggtaacttagaagcttgtacacttacgcacgtaggtgccccacatccagtggggtttgtg caaagcaatcttgttcccggtaaaccctgataggctgtaaccacggccgaaacaaggtttgtcgtt acccgactcactactacacaaagcctagtaaagttcaatgaaagtgcgcagcgtgatccggtcaaa acccccttgaccagaaacacatgatgagggtcaccaaccccactggcgacagtgtggtgtccct gcgttggcatgtggcctcgtagaggcgttgcaatctggatttgctccgaagagccccgtgtgctagt gtttatacctccggcccctgaatgcggctaatcctaacccccgagcatgtacacacaagccagtgt gtagcatgtcgtaatgagcaatttggggatggaaccgactactttagggtgtccgtgtttctcattatt ctttgtttgatgtttt
682	관박쥐 피코르나바이러스 균주 구이저우-Rr100	ttttttttctcaggcggtagcatccagccaaggcctgatgctaccaacgtgtgactaaaccacactct ctttttgtgatacattgtgtcacctatccctttcttggtaacttagaagcttgtacacccacgcacgtag gtaccccacatccagtgggtttgtgcaaagcattcttgttcccggtaaaccctgataggctgtaa ccacggctgaaacaaggtttgtcgttacccgactcactactacgcaaagcctagtaaagttcaatga aagtgcgcagcgtgatccggtcaaaacccccttgaccagaaacacatgatgagggtcaccaacc cccactggcgacagtgtggtgtccctgcgttggcatgtggcctcatagaggcgttgcaatctggatt tgctccgaagagccccgtgtgctagtgtttatacctccggcccctgaatgcggctaatcctaaccc ccgagcatgtacacacacgccagtgtgcagcatgtcgtaatgagcaatttggggatggaaccgac tactttagggtgtccgtgtttctcattattctttgtttgatgttttatggtgacaaca
683	관박쥐 피코르나바이러스 균주 헤난-Rf265	cggaacgttgtatgctcagggcgtaggcaccacccacgggtggtgcctacacgtgtggactaaac cacacactcttttcagcacttagtgctgctatctcttttgtaacttagaagtttgtacacaatgcgttag ggccacacatccagtgtggtatcgcaaagcacttctgtttccccggtgctagtaggagggtggctg ctccacggccactgccgaacccatcgttacccgactcattacttcgcaaagcctagtaacccagtt gaagcaagcccggcgtgttccggtcaggaaaaaccccccctggccagaaacatgtgatgagggt gggctatcccccactggtgacagtgagccctccctgcgttggcacatggcccgatctgggcgtggtt cttgtggatgctgccgaagagccccgtgagctagtgtttataccgccggcctcgtgaatgcggcta accctaaccccggagcagaggctactgaagccacagtagtcgctgtcgtaacgagtaattctggg atgggaccgactactttcgagtgtccgtgtttccttttattcttttattgttgtttatggtgacaaac
684	인간 엔테로바이러스 C105	cccctaggatccactggatgtcagtacactggtatcgtggtaccttgtacgcctgtttttataccccctt ccccgcaactttagaagcatcaaaagcaccgctcaatagtcaccacaccccagtgtggtttcgag caagcacttctgttttcccggttgcgtcccatatgctgtgcaaacggcaaaaagggacaatatcgtta cccgcttgtatactacgggaaacctagtaccaccattgattgtgttgagagttgcgctcatcacctttc cccggtgtagctcaggccgatgaggctcagaatcccccacaggtgactgtgtctgagcctgcgtt ggcggcctgcctcgccttatggcgtgggacgcttgatacatgacatggtgcgaagagtctactgt gctatgcaagagtcctccggcccctgaatgtggctaatcctaaccactgatcccacgcacgcaaac cagtgtgtagtgggtcgtaacgcgcaagtcggtggcggaaccaactactttgggtgaccgtgtttc ctttattacttattgaatgtttatggtgacaattgtttgattcagttgttgccattctctacattcatttaccca gcatcaaaccaattgaactgttaca
685	인간 폴리오바이러스 1 균주 NIE1116623	agtctggacatccctcaccggcgacggtggtccaggctgcgctggcggcctacctgtggtccaaa gccacgggacgctacatgtgaacaaggtgtgaagagcctattgagctacaaaagagtcctccgg cccctgaatgcggctaatcccaaccacggatcaagggtgcacaaaccagtgtacaccttgtcgta acgcgcaagtctgtggcggaaccgactactttgggtgtccgtgtttccttttaattttgatggctgctt atggtgacaatcatagattgttatcataaagctaattggattggccatccggtgagagtgaaatatatt gtttacctccctgttgggtttactctaactaacttctccatttataaacttgtcatcacagttttaataatta gaagtgcagtttaca
686	인간 엔테로바이러스	tttaaaacagcctgggggttgttcccacccccagggcccactgggcgttagtactctggtatcgcg gtaccttagtatgcctgttttatgtctcctttccccccgcaactttagaagtaatcaagttatggctcaaca

[1758]

	109	gtcgccacacccccagtgtggttccgagcaagcacttctgttccccggttgcgtcttatatgctgtgt gaacggcagaaagggacaatatcgttatccgctcaactactacgggaagcctagtaccaccatgg attgacctgaaagttgcgttcagcgcaccccccagcgcagctcaggccgatgaggctccgaatacc ccacgggcgaccgtgtcggagcctgcgttggcggcctgcccacgttgcaaaacgtgggacgctc atttcatgacatggtgcgaagagcctactgtgctagttgagagtcctccggccctgaatgtggata atcctaaccactgaacctacgggcgcaaaccagcgtctggtaggccgtaacgcgcaagtcggtg gcggaaccaactactttgggtgtccgtgtttccttttatcttttgaatgtttatggtgacaattgttgtgta cagttgttaccatagtttgcattcagaaataaacctaacactttccaattatttgttaca
687	인간 폴리오바이러스 2 균주 NIE0811460	ttgtgcgcctgttttatattccccccccgcaacttagaagcacgaaaccaagttcaatagaaggggg tacaaaccagtaccaccacgaacaagcacttctgtttccccggtgacattgcatagactgctcacgc ggttgaaagtgatcgatccgttacccgcttgtgtacttcgaaaagcctagtattgccttggaatcttcg acgcgttgcgctcagcacccgaccccggggtgtagcttaggctgatgagtctggacattcctcacc ggtgacggtggtccaggctgcgttggcggcctacctatggctaatgccataggacgctagatgtg aacaaggtgtgaagagcctattgagctacataagagtcctccggcccctgaatgcggctaatccta accacggagcaggcggtcgcaaaccagtgactagcttgtcgtaacgcgcaagtctgtggcggaa ccgactactttgggtgtccgtgtttccttattttattatggctgcttatggtgacaatcagagattgtt atcataaagcgaattggattggccatccggtgagtgttgtgtcaggtatacaactgtttgttggaacc actgtgttagtttaacctctctttcaaccaattagtcaaaaacaatacgaagatagaacaacaatacta ca
688	소 피코르나바이러 스	ttttctcccctcccctccaactacctttttcccctctgtaacgctagaagtttgtgcaaaccgcctgt agggtactgcaatccagcagtgcataggctaagctttttcttgttaccccaccccacattatactgagg aggattgtgaaattgtgttagtatgggttagtagcggtgacccgggtaaccccaacccagaaactc acggatgagatgaacaggaccccacatggtaacgtgtgtgttcgtctgccccgcaaggtgaggcc gtgagagctttgcacgcgaaaaccttgaaaacccaaaagtaccttgagctcttcgctattttgtgtttc ctccaggaccctgaatgcggctaaacctaacccgcgatccgcacgtagcaacccagctagagtgt ggtcgtaatgcgcaagttgcgggcggtaccgactactttggtgttcctgtgtttccttttattttattttga atttttatggtgacaacagctagaaaataagagtgaac
689	인간 폴리오바이러스 1 균주 EQG1419328	acccttgtacgcctgtttttatactcccctccccgtaacttagaagaaacaaaataagttcaataggag gggggtacaaaccagtaccaccacgaacaagcacttctgtctccccggtgacattgcatagactgtc ccccacggttgaaagcaattgatccgttacccgctcttgtacttcgagaagcctagtaccatcttggaa tcatcgatgcgttgcgctccacactcagtcccagagtgtagcttaggctgatgagtctggacattcct caccggcgacggtggtccaggctgcgttggcggcctacctgtggcccaaagccacaggacgct agatgtgaacaaggtgtgaagagcctattgagctatagagagtcctccggcccctgaatgcggc taatcccaaccacggatcaaggggtgcacgaaccagtgtatacctttgtcgtaacgcgcaagtccgtg gcggaaccgactactttgggtgaccgtgtttccttttattatttcaatggctgcttatggtgacaatcatt gattgttatcataaagcgaattggactggccatccggtgaaagtgaaacatattgtttgcctcctcgtt gggtctacttcaaccaatctttacttacaatcttaccactacagttttgctggttagaagtgtgtttcacg
690	인간 폴리오바이러스 2 단리물 IS_061	ttgtgcgcctgttttatactccccctcccgcaacttagaagcacgaaaccaagttcaatagaaggggg tacaaaccagtaccactacgaacaagcacttctgtttccccggtgacattgcatagactgctcacgc ggttgaaagtgatcgatccgttacccgcttgtgtacttcgaaaagcctagtatcgccttggaatcttc gacgcgttgcgctcagcacccgaccccggggtgtagcttaggccgatgagtctggacattcctca ccggtgacggtggtccaggctgcgttggcggcctacctatggctaacgccataggacgttagatg tgaacaaggtgtgaagagcctattgagctacataagagtcctccggcccctgaatgcggctaatcc taaccacggagcaggcggtcgcgaaccagtgactggcttgtcgtaacgcgcaagtctgtggcgg aaccgactactttgggtgtccgtgtttccttttatttttatcatggctgcttatggtgacaatcagagatt gttatcataaagcgaattggattggccatccggtgagtgttgtgtcaggtatacaactgtttgttggaa ccactgtgttagctttgcttctcatttaaccaattaatcaaaaacaatacgaggataaaacaacaatac taca
691	콕사키바이러스 B5	cctttgtgcgcctgttttatgccccctccccaattgaaacttagaagttacacacaccgatcaacag cgggcgtggcataccagccgcgtcttgatcaagcactcctgtttccccggaccgagtatcaataga ctgctcacgcggttgaaggagaaaacgttcgttacccggctaactacttcgagaaacctagtagca tcatgaaagttgcgaagcgtttcgctcagcacatccccagtgtagatcaggtcgatgagtcaccgc attcccacgggcgaccgtggcggtggctgcgttggcggcctgcctacggggcaacccgtagg acgcttcaatacagacatggtgcgaagagtcgattgagctagttagtagtcctccggcccctgaatc

[1759]

		cggctaatcctaactgcggagcacataccctcaacccaggggcattgtgtcgtaacgggtaact ctgcagcggaaccgactactttgggtgtccgtgtttccttttattcttataatggctgcttatggtgaca attgaaagattgttaccatatagctattggattggccatccggtgtctaacagagctattatatacctct ttgttggatttgtaccacttgatctaaaggaagtcaagacactacaattcatcatacaattgaacacag caaa
692	콕사키바이러스 A10	tttgtgcgcctgttttacaacccttccccaacttgtaacgtagaagtaatacacactactgatcaatag caggcatggcgccgccagtcatgtctcgatcaagcacttctgttcccccggactgagtatcaatага ctgctcacgcggttgaaggagaaaacgttcgttacccggctaactacttcgagaaacctagtagca ccatagaagctgcagagtgtttcgctcagcacttcccccgtgtagatcaggctgatgagtcactgca atccccacgggtgaccgtggcagtggctgcgttggcggcctgcctatggggcaacccataggac gctctaatgtggacatggtgcgaagagtctattgagctagttagtagtcctccggcccctgaatgcg gctaatcctaactgcggagcacatgccttcaacccagaaggtagtgtgtcgtaacgggcaactctg cagcggaaccgactactttgggtgtccgtgtttcttttattcctatattggctgcttatggtgacaatca cggaattgttgccatatagctattggattggccatccggtgtctaatagagctattgtgtacctatttgtt ggatttactccgctatcacataaatctctgaacactttgtgctttatattgaacttaaacacccgaaa

[1760]

[1761] 일부 실시형태에서, 본 발명의 IRES는 표 17에 열거된 서열(서열번호 1 내지 72 및 348 내지 389)을 갖는 IRES이다. 일부 실시형태에서, IRES는 살리바이러스 IRES이다. 일부 실시형태에서, IRES는 살리바이러스 SZ1 IRES이다. 일부 실시형태에서, IRES는 AP1.0(서열번호 348)이다. 일부 실시형태에서, IRES는 CK1.0(서열번호 349)이다. 일부 실시형태에서, IRES는 PV1.0(서열번호 350)이다. 일부 실시형태에서, IRES는 SV1.0(서열번호 351)이다.

표 18

아나베나 순열 부위 5' 인드론 단편 서열.

서열번호	순열 부위	서열
73	L2-1	GAAGAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACC TAAATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAG TAGTAATTAGTAAGTTAACAATAGATGACTTACAACTAATCG GAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGAC GAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGT AGCGAAAGCTGCAAGAGAATGAAAATCCGT
74	L2-2	AAGAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCT AAATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGT AGTAATTAGTAAGTTAACAATAGATGACTTACAACTAATCGG AAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACG AGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTA GCGAAAGCTGCAAGAGAATGAAAATCCGT
75	L2-3	AGAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTA AATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTA GTAATTAGTAAGTTAACAATAGATGACTTACAACTAATCGGA AGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGA GGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAG CGAAAGCTGCAAGAGAATGAAAATCCGT
76	L5-1	GTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATT AGTAAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGC AGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAA AGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAG CTGCAAGAGAATGAAAATCCGT
77	L5-2	TTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCA GAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAA GAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGC TGCAAGAGAATGAAAATCCGT
78	L5-3	TATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAG TAAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAG AGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAG AGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCT GCAAGAGAATGAAAATCCGT
79	L5-4	ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTG CAAGAGAATGAAAATCCGT
80	L5-5	TAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTA AGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGAG ACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGT
81	L6-1	ACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACTCG ACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAG TCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGA GAATGAAAATCCGT
82	L6-2	CAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACTCGA

[1762]

		CGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
83	L6-3	AATAGATGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
84	L6-4	ATAGATGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
85	L6-5	TAGATGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
86	L6-6	AGATGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
87	L6-7	GATGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
88	L6-8	ATGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
89	L6-9	TGACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
90	L8-1	CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
91	L8-2	AAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
92	L8-3	AGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
93	L8-4	GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
94	L8-5	ACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
95	L9a-1	AATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
96	L9a-2	ATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
97	L9a-3	TAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
98	L9a-4	AGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
99	L9a-5	GGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
100	L9-1	GAAAGCTGCAAGAGAATGAAAATCCGT
101	L9-2	AAAGCTGCAAGAGAATGAAAATCCGT
102	L9-3	AAGCTGCAAGAGAATGAAAATCCGT
103	L9-4	AGCTGCAAGAGAATGAAAATCCGT
104	L9-5	GCTGCAAGAGAATGAAAATCCGT
105	L9-6	CTGCAAGAGAATGAAAATCCGT
106	L9-7	AAGAGAATGAAAATCCGT
107	L9-8	AGAGAATGAAAATCCGT

[1763]

108	L9-9	GAGAATGAAAATCCGT
109	L9a-6	GCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
110	L9a-7	AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
111	L9a-8	GTAGCGAAAGCTGCAAGAGAATGAAAATCCGT

[1764]

[1765] 일부 실시형태에서, 5' 인트론 단편은 표 18에 열거된 서열을 갖는 단편이다. 전형적으로, 표 18에 열거된 5' 인트론 단편을 함유하는 작제물은 표 19에 열거된 바와 같은 상응하는 3' 인트론 단편을 함유할 것이다(예를 들어, 둘 다 L9a-8 순열 부위를 갖는 단편을 나타냄).

표 19

아나베나 순열 부위 3' 인트론 단편 서열.

서열번호	순열 부위	서열
112	L2-1	ACGGACTTAAATAATTGAGCCTTAAA
113	L2-2	ACGGACTTAAATAATTGAGCCTTAAAG
114	L2-3	ACGGACTTAAATAATTGAGCCTTAAAGA
115	L5-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTA
116	L5-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAG
117	L5-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGT
118	L5-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT
119	L5-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTA
120	L6-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT A
121	L6-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AA
122	L6-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AAC
123	L6-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACA
124	L6-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAA
125	L6-6	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAAT
126	L6-7	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATA
127	L6-8	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAG
128	L6-9	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT

[1766]

		AACAATAGA
129	L8-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGT
130	L8-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTC
131	L8-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCA
132	L8-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAA
133	L8-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAG
134	L9a-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCC
135	L9a-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCA
136	L9a-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAA
137	L9a-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAAT
138	L9a-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT

[1767]

		CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATA
139	L9-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGC
140	L9-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCG
141	L9-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGA
142	L9-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAA
143	L9-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAA
144	L9-6	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAG
145	L9-7	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC
146	L9-8	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCA
147	L9-9	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT

[1768]

		CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCA A
148	L9a-6	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAG
149	L9a-7	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGC
150	L9a-8	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCA

[1769]

[1770] 일부 실시형태에서, 3' 인트론 단편은 표 19에 열거된 서열을 갖는 단편이다. 일부 실시형태에서, 표 19에 열거된 3' 인트론 단편을 함유하는 작제물은 표 18에 열거된 바와 같은 상응하는 5' 인트론 단편을 함유할 것이다 (예를 들어, 둘 다 L9a-8 순열 부위를 갖는 단편을 나타냄).

표 20

비-아나베나 순열 부위 5' 인트론 단편 서열.

서열번호	인트론	서열
151	Azop1	tgcgccgatgaaggtgtagagactagacggcacccacctaaggcaaacgctatggtgaaggcatagtccagggagtggcgaaagtcacacaaaccggaatccgt
152	Azop2	ccgggcgtatggcaacgccgagccaagcttcggcgcctgcgccgatgaaggtgtagagactagacggcacccacctaaggcaaacgctatggtgaaggcatagtccagggagtggcgaaagtcacacaaaccggaatccgt
153	Azop3	acggcacccacctaaggcaaacgctatggtgaaggcatagtccagggagtggcgaaagtcacacaaaccggaatccgt
154	Azop4	acgctatggtgaaggcatagtccagggagtggcgaaagtcacacaaaccggaatccgt
155	S795p1	attaaagttatagaattatcagagaatgatatagtccaagccttatggtaacatgagggcacttgaccctggtag
156	Twortp1	aagatgtaggcaatcctgagctaagctcttagtaataagagaaagtgcaacgactattccgataggaagtaggtcaagtgactcgaaatggggattacccttctagggtagtgatatagtctgaacatatatggaaacatataggaaggataggagtaacgaacctattcgtaacataattgaacttttagttat
157	Twortp2	taataagagaaagtgcaacgactattccgataggaagtagggtcaagtgactcgaaatggggattacccttctagggtagtgatatagtctgaacatatatggaaacatatagaaggataggagtaacgaacctattcgtaacataattgaacttttagttat
158	Twortp3	taggaagtagggtcaagtgactcgaaatggggattacccttctagggtagtgatatagtctgaacatatatggaaacatatagaaggataggagtaacgaacctattcgtaacataattgaacttttagttat
159	Twortp4	ctagggtagtgatatagtctgaacatatatggaaacatatagaaggataggagtaacgaacctattcgtaacataattgaacttttagttat
160	LSUp1	agttaataaagatgatgaaatagtctgaaccattttgagaaaagtggaaataaaagaaaatcttttatgataacataaattgaacaggctaa
161	Phip1	caaagactgatgatatagtccgacactcctagtaataggagaatacagaaaggatgaaatcc
162	Nostoc	agtcgagggtaaagggagagtccaattctcaaagcctattggcagtagcgaaagctgcgggagaatgaaaatccgt
163	Nostoc	agccgagggtaaagggagagtccaattctcaaagccaataggcagtagcgaaagctgcgggagaatgaaaatccgt
164	Nodularia	agccgagggtaaagggagagtccaattctcaaagccgaaggttattaaaacctggcagcagtgaaagctgcgggagaatgaaaatccgt
165	Pleurocapsa	agctgagggtaaagagagagtccaattctcaaagccagcagatggcagtagcgaaagctgcgggagaatgaaaatccgt
166	Planktothrix	agccgagggtaaagagagagtccaattctcaaagccaattggtagtagcgaaagctacgggagaatgaaaatccgt

[1771]

[1772] 일부 실시형태에서, 5' 인트론 단편은 표 20에 열거된 서열을 갖는 단편이다. 표 20에 열거된 5' 인트론 단편을 함유하는 작제물은 표 21의 상응하는 3' 인트론 단편을 함유할 것이다(예를 들어, 둘 다 Azop1 인트론을 갖는 단편을 나타냄).

표 21

비-아나베나 순열 부위 3' 인트론 단편 서열.

서열번호	인트론	서열
167	Azop1	gcggactcatatttcgatgtgccttgcgccgggaaaccacgcaagggatggtgtcaaattcggcgaaac ctaagcgcccgcccgggcgtatggcaacgccgagccaagcttcggcgcc
168	Azop2	gcggactcatatttcgatgtgccttgcgccgggaaaccacgcaagggatggtgtcaaattcggcgaaac ctaagcgcccgc
169	Azop3	gcggactcatatttcgatgtgccttgcgccgggaaaccacgcaagggatggtgtcaaattcggcgaaac ctaagcgcccgcccgggcgtatggcaacgccgagccaagcttcggcgcctgcgccgatgaaggtgta gagactag
170	Azop4	gcggactcatatttcgatgtgccttgcgccgggaaaccacgcaagggatggtgtcaaattcggcgaaac ctaagcgcccgcccgggcgtatggcaacgccgagccaagcttcggcgcctgcgccgatgaaggtgta gagactagacggcacccacctaaggcaa
171	S795p1	aggattagatactacactaagtgtcccccagactggtgacagtctggtgtgcatccagctatatcggtgaa accccattggggtaataccgagggaagctatattatatatattaataaatagccccgtagagactatgta ggtaaggagatagaagatgataaaatcaaaatcatc
172	Twortp1	actactgaaagcataaataattgtgcctttatacagtaatgtatatcgaaaaatcctctaattcagggaacac ctaaacaaact
173	Twortp2	actactgaaagcataaataattgtgcctttatacagtaatgtatatcgaaaaatcctctaattcagggaacac ctaaacaaactagatgtaggcaatcctgagctaagctcttag
174	Twortp3	actactgaaagcataaataattgtgcctttatacagtaatgtatatcgaaaaatcctctaattcagggaacac ctaaacaaactaagatgtaggcaatcctgagctaagctcttagtaataagagaaagtgcaacgactattcc ga
175	Twortp4	actactgaaagcataaataattgtgcctttatacagtaatgtatatcgaaaaatcctctaattcagggaacac ctaaacaaactaagatgtaggcaatcctgagctaagctcttagtaataagagaaagtgcaacgactattcc gataggaagtagggtcaagtgactcgaaatggggattaccctt
176	LSUp1	cgctagggatttataactgtgagtcctccaatattataaaatgttggtaatatattgggtaaatttcaaagaca acttttctccacgtcaggatatagtgtatttgaagcgaaacttattttagcagtgaaaaagcaaataaggac gttcaacgactaaaaggtgagtattgctaacaataatccttttttttaatgcccaacatctttattaact
177	Phip1	gtgggtgcataaactatttcattgtgcacattaaatctggtgaactcggtgaaaccctaatggggcaatacc gagccaagccatagggaggatatatgagaggcaagaagttaattcttgaggccactgagactggctgta tcatccctacgtcacacaaacttaatgccgatggttatttcagaaagaaaaccaatggcgtcttagagatgt atcacagaacggtgtggaaggagcataacggagacatacctgatggcttcgagatagaccataagtgtc gcaatagggcttgctgtaatatagagcatttacagatgcttgagggtacagcccacactgttaagaccaat

[1773]

		cgtgaacgctacgcagacagaaaggaaacagctagggaatactggctggagactggatgtaccggcc tagcactcggtgagaagtttggtgtgtcgttctcttctgcttgtaagtggattagagaatggaaggcgtaga gactatccgaaaggagtagggccgagggtgagactccctcgtaacccgaagcgccagacagtcaact
178	Nostoc	acggacttaagtaattgagccttaaagaagaaattctttaagtggcagctctcaaactcagggaaacctaa atctgttcacagacaaggcaatcctgagccaagccgaaagagtcatgagtgctgagtagtgagtaaaat aaaagctcacaactcagaggttgtaactctaagctagtcggaaggtgcagagactcgacgggagctac cctaacgtaa
179	Nostoc	acggacttaaactgaattgagccttagagaagaaattctttaagtgtcagctctcaaactcagggaaacct aaatctgttgacagacaaggcaatcctgagccaagccgagaactctaagttattcggaaggtgcagaga ctcgacgggagctaccctaacgtca
180	Nodularia	acggacttagaaaactgagccttgatcgagaaatctttcaagtggaagctctcaaattcagggaaacctaa atctgtttacagatatggcaatcctgagccaagccgaaacaagtcctgagtgttaaagctcataactcatc ggaaggtgcagagactcgacgggagctaccctaacgtta
181	Pleurocapsa	acggacttaaaaaaattgagccttggcagagaaatctgtcatgcgaacgctctcaaattcagggaaacct aagtctggcaacagatatggcaatcctgagccaagccttaatcaaggaaaaaacatttttacctttacctt gaaaggaaggtgcagagactcaacgggagctaccctaacaggtca
182	Planktothrix	acggacttaaagataaattgagccttgaggcgagaaatctctcaagtgtaagctgtcaaattcagggaaa cctaaatctgtaaattcagacaaggcaatcctgagccaagcctaggggtattagaaatgagggagtttcc ccaatctaagatcaatacctaggaaggtgcagagactcgacgggagctaccctaacgtta

[1774]

[1775] 일부 실시형태에서, 3' 인트론 단편은 표 21에 열거된 서열을 갖는 단편이다. 표 21에 열거된 3' 인트론 단편을

함유하는 작제물은 표 20에 열거된 바와 같은 상응하는 5' 인트론 단편을 함유할 것이다(예를 들어, 둘 다 Azop1 인트론을 갖는 단편을 나타냄).

표 22

스페이서 및 아나베나 5' 인트론 단편 서열.

서열번호	스페이서	서열
183	T25 L10	agtatataagaaacaaaccacTAGATGACTTACAACTAATCGGAAGGTGC AGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAA AGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAG CTGCAAGAGAATGAAAATCCGTggctcgcagc
184	T25 L20	ctgaaattatacttatactcaaacaaaccacTAGATGACTTACAACTAATCGGAA GGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGC GAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
185	T25 L30 (I80-10) [대조군]	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
186	T25 L40	catcaacaatatgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTAC AACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAA CGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCA ATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggct cgcagc
187	T25 L50	catcaacaatatgaaactatacttatactcagtatatgaagcattatcgcaaacaaaccacTAGATG ACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTA CCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCA AAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAAT CCGTggctcgcagc
188	T50 L10	tagcgtcagcaaacaaacaaaTAGATGACTTACAACTAATCGGAAGGTGC AGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAA AGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAG CTGCAAGAGAATGAAAATCCGTggctcgcagc
189	T50 L20	atactcatactagcgtcagcaaacaaacaaaTAGATGACTTACAACTAATCGGA AGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGA GGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAG CGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
190	T50 L30	gtgtgaagctatactcatactagcgtcagcaaacaaacaaaTAGATGACTTACAACTA ATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCA AGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
191	T50 L40	cctcacctgagtgtgaagctatactcatactagcgtcagcaaacaaacaaaTAGATGACTTA CAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTA ACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCC AATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgg ctcgcagc
192	T50 L50	ccgaatgatgcctcacctgagtgtgaagctatactcatactagcgtcagcaaacaaacaaaTAGAT GACTTACAACTAATCGGAAGGTGCAGAGACTCGACGGGAGCT ACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTC AAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAA TCCGTggctcgcagc
193	T75 L10	cggtgcgagcaaacaaacaaaTAGATGACTTACAACTAATCGGAAGGTG CAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTA AAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAA GCTGCAAGAGAATGAAAATCCGTggctcgcagc

[1776]

194	T75 L20	cgctccgacccagtgcgagcaaacaaacaaaTAGATGACTTACAACTAATCGG AAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACG AGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTA GCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
195	T25 L30 1MM	ctgaaattatactAatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
196	T25 L30 3MM	ctgaaaAtatactAatactcaCtatatgacaaacaaaccacTAGATGACTTACAACTA ATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCA AGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
197	T25 L30 5MM	ctgaTaAtataGtAatactcaCtatatgacaaacaaaccacTAGATGACTTACAACT AATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTC AAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAG GCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
198	T25 L30 8MM	ctgaTaAtaAaGtAatacAcaCtataAgacaaacaaaccacTAGATGACTTACAA CTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACG TCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAAT AGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgc agc
199	T25 L30 오프타깃 10	ctgaaattatacttatactctctaagttacaaacaaaccacTAGATGACTTACAACTAAT CGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAG ACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCA GTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
200	T25 L30 오프타깃 20	ctgaaattatgtgtgttacAtctaagttacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
201	T25 L30 오프타깃 30	gttgatcggtgtgtgttacAtctaagttacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
202	T25 L30 I25-10	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattaaacag
203	T25 L30 I25-20	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattcacaatataaa ttacg
204	T25 L30 I50-10	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcatagc
205	T25 L30 I50-20	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcgcagcataa tatccg
206	T25 L30 I80-20	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC

[1777]

		AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagcgcg cctaccg
207	T25 L30 I80-20x2	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagcgcg cctaccgaaagccggcgtcgacgttagcgc
208	T25 L30 I50-20x2	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcgcagcataa tatccgaaacgaggatacaagtgacatgc
209	T25 L30 I25-20x2	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattcacaatctaaa ttacgaaacgataaatgataactctaac
210	T0 L0	aaacaaaccacTAGATGACTTACAACTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAA GAGAATGAAAATCCGTggctcgcagc
211	T100 L5	cgggcaaacaaacaaaTAGATGACTTACAACTAATCGGAAGGTGCAG AGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAG AGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCT GCAAGAGAATGAAAATCCGTggctcgcagc
212	T75 L30	cgctccgacgagcttccggccagtgcgagcaaacaaacaaaTAGATGACTTACAACT AATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTC AAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAG GCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
213	T0 L0a	aaacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCC GTggctcgcagc
214	T25 L10a	agtatataagaaacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgcagc
215	T25 L20a	ctgaaattatacttatactcaaacaaaccacGGCAGTAGCGAAAGCTGCAAGAG AATGAAAATCCGTggctcgcagc
216	T25 L30a (I80-10) [대조군]	ctgaaattatacttatactcagtatatgacaaacaaaccacGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggctcgcagc
217	T50 L10a	tagcgtcagcaaacaaacaaaGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgcagc
218	T50 L20a	atactcatactagcgtcagcaaacaaacaaaGGCAGTAGCGAAAGCTGCAAGAG AATGAAAATCCGTggctcgcagc
219	T50 L30a	gtgtgaagctatactcatactagcgtcagcaaacaaacaaaGGCAGTAGCGAAAGCTG CAAGAGAATGAAAATCCGTggctcgcagc
220	T75 L10a	cggtgcgagcaaacaaacaaaGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgcagc
221	T75 L20a	cgctccgacccagtgcgagcaaacaaacaaaGGCAGTAGCGAAAGCTGCAAGA GAATGAAAATCCGTggctcgcagc
222	T75 L30a	cgctccgacgagcttccggccagtgcgagcaaacaaacaaaGGCAGTAGCGAAAGCT GCAAGAGAATGAAAATCCGTggctcgcagc
223	T0 L0b	aaacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAA GCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCC GTggctcgcagc
224	T25 L10b	agtatataagaaacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTC

[1778]

		TCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAA AATCCGTggctcgcagc
225	T25 L20b	ctgaaattatacttatactcaaacaaaccacAAGACGAGGGTAAAGAGAGAGTC CAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGA ATGAAAATCCGTggctcgcagc
226	T25 L30b (I80-10) [대조군]	ctgaaattatacttatactcagtatatgacaaacaaaccacAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggctcgcagc
227	T50 L10b	tagcgtcagcaaacaaacaaaAAGACGAGGGTAAAGAGAGAGTCCAATT CTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgcagc
228	T50 L20b	atactcatactagcgtcagcaaacaaacaaaAAGACGAGGGTAAAGAGAGAGT CCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAG AATGAAAATCCGTggctcgcagc
229	T50 L30b	gtgtgaagctatactcatactagcgtcagcaaacaaacaaaAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTG CAAGAGAATGAAAATCCGTggctcgcagc
230	T75 L10b	cggtgcgagcaaacaaacaaaAAGACGAGGGTAAAGAGAGAGTCCAATT CTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgcagc
231	T75 L20b	cgctccgacccagtgcgagcaaacaaacaaaAAGACGAGGGTAAAGAGAGAG TCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGA GAATGAAAATCCGTggctcgcagc
232	T75 L30b	cgctccgacgagcttccggccagtgcgagcaaacaaacaaaAAGACGAGGGTAAAG AGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCT GCAAGAGAATGAAAATCCGTggctcgcagc
233	T25 L30 I0-0	ctgaaattatacttatactcagtatatgacaaacaaaccacTAGATGACTTACAACTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
234	T25 L30a I0-0	ctgaaattatacttatactcagtatatgacaaacaaaccacGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGT
235	T25 L30a I25-10	ctgaaattatacttatactcagtatatgacaaacaaaccacGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTgattaaacag
236	T25 L30a I25-20	ctgaaattatacttatactcagtatatgacaaacaaaccacGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTgattcacaatataaattacg
237	T25 L30a I50-10	ctgaaattatacttatactcagtatatgacaaacaaaccacGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggatcatagc
238	T25 L30a I50-20	ctgaaattatacttatactcagtatatgacaaacaaaccacGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggatcgcagcataatatccg
239	T25 L30a I80-20	ctgaaattatacttatactcagtatatgacaaacaaaccacGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggctcgcagcgcgcctaccg
240	T25 L30b I0-0	ctgaaattatacttatactcagtatatgacaaacaaaccacAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGT
241	T25 L30b I25-10	ctgaaattatacttatactcagtatatgacaaacaaaccacAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTgattaaacag
242	T25 L30b I25-20	ctgaaattatacttatactcagtatatgacaaacaaaccacAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTgattcacaatataaattacg
243	T25 L30b I50-10	ctgaaattatacttatactcagtatatgacaaacaaaccacAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggatcatagc

[1779]

244	T25 L30b I50-20	ctgaaattatacttatactcagtatatgacaaacaaaccacAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggatcgcagcataatatccg
245	T25 L30b I80-20	ctgaaattatacttatactcagtatatgacaaacaaaccacAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggctcgcagcgcgcctaccg

[1780]

[1781] 일부 실시형태에서, 스페이서 및 5' 인트론 단편은 표 22에 열거된 바와 같은 서열을 갖는 스페이서 및 단편이다.

표 23

스페이서 및 아나베나 3' 인트론 단편 서열

서열번호	스페이서	서열
246	T25 L10	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaacttatatact
247	T25 L20	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagagtataagtatatttcag
248	T25 L30 (I80-10) [대조군]	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
249	T25 L40	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcatattgttgatg
250	T25 L50	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagcgataatgcttcatatactgagtataagtatagtttcatattg ttgatg
251	T50 L10	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagctgacgcta
252	T50 L20	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagctgacgctagtatgagtat
253	T50 L30	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagctgacgctagtatgagtatagcttcacac
254	T50 L40	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagctgacgctagtatgagtatagcttcacactcaggtgagg
255	T50 L50	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagctgacgctagtatgagtatagcttcacactcaggtgaggc atcattcgg
256	T75 L10	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagctcgcaccg
257	T75 L20	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagctcgcactgggtcggagcg

[1782]

258	T25 L30 1MM	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
259	T25 L30 3MM	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
260	T25 L30 5MM	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
261	T25 L30 8MM	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
262	T25 L30 오프타킷 10	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtaacttagagagtataagtataatttcag
263	T25 L30 오프타킷 20	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtaacttagaTgtaacacacataatttcag
264	T25 L30 오프타킷 30	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtaacttagaTgtaacacacaccgatcaac
265	T25 L30 I25-10	ctgtttaatcACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCT TTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
266	T25 L30 I25-20	cgtaatttatattgtgaatcACGGACTTAAATAATTGAGCCTTAAAGAAGA AATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATC TAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAA TTAGTAAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
267	T25 L30 I50-10	gctatgatccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
268	T25 L30 I50-20	cggatattatgctgcgatccACGGACTTAAATAATTGAGCCTTAAAGAAG AAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAA TCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGT AATTAGTAAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
269	T25 L30 I80-20	cggtaggcgcgctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAA GAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAA ATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAG TAATTAGTAAGTTAACAAcacaaacacaagtcatatactgagtataagtataatttcag
270	T25 L30 I80-20x2	gcgctaacgtcgacgccggcaaacggtaggcgcgctgcgagccACGGACTTAAATAA TTGAGCCTTAAAGAAGAAATTCTTTAAGTGGATGCTCTCAAAC TCAGGGAAACCTAAATCTAGTTATAGACAAGGCAATCCTGAG CCAAGCCGAAGTAGTAATTAGTAAGTTAACAAcacaaacacaagtcat atactgagtataagtataatttcag
271	T25 L30 I50-	gcatgtcacttgtatcctcgaaacggatattatgctgcgatccACGGACTTAAATAATTG

[1783]

	20x2	AGCCTTAAAGAAGAAATTCTTTAAGTGGATGCTCTCAAACTC AGGGAAACCTAAATCTAGTTATAGACAAGGCAATCCTGAGCC AAGCCGAAGTAGTAATTAGTAAGTTAACAAcacaaacacaagtcatatac tgagtataagtataatttcag
272	T25 L30 I25-20x2	gttagagttatcatttatcgaaacgtaatttagattgtgaatcACGGACTTAAATAATTGA GCCTTAAAGAAGAAATTCTTTAAGTGGATGCTCTCAAACTCA GGGAAACCTAAATCTAGTTATAGACAAGGCAATCCTGAGCCA AGCCGAAGTAGTAATTAGTAAGTTAACAAcacaaacacaagtcatatactg agtataagtataatttcag
273	T0 L0	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaaacacaa
274	T100 L5	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagcccg
275	T75 L30	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaaaaacaagctcgcactggccggaagctcgtcggagcg
276	T0 L0a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaaacacaa
277	T25 L10a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaaacacaacttatatact
278	T25 L20a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaaacacaagagtataagtataatttc ag
279	T25 L30a (I80-10) [대조군]	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaaacacaagtcatatactgagtataa gtataatttcag
280	T50 L10a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctgacgcta
281	T50 L20a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC

[1784]

		TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctgacgctagtatga gtat
282	T50 L30a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctgacgctagtatga gtatagcttcacac
283	T75 L10a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctcgcaccg
284	T75 L20a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctcgcactgggtcgg agcg
285	T75 L30a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctcgcactggccgga agctcgtcggagcg
286	T0 L0b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCcacaaacacaa
287	T25 L10b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCcacaaacacaacttatatact
288	T25 L20b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCcacaaacacaagagtataagtataatt tcag
289	T25 L30b (I80-10) [대조군]	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCcacaaacacaagtcatatactgagtat

[1785]

		aagtataatttcag
290	T50 L10b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgacgcta
291	T50 L20b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgacgctagtatg agtat
292	T50 L30b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgacgctagtatg agtatagcttcacac
293	T75 L10b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctcgcaccg
294	T75 L20b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctcgcactgggtcg gagcg
295	T75 L30b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctcgcactggccg gaagctcgtcggagcg
296	T25 L30 I0-0	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAAcacaaacacaagtcatatactgagtataagtataatttcag
297	T25 L30a I0-0	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACTC GACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGA GTCCAATTCTCAAAGCCAATAcacaaacacaagtcatatactgagtataagtataat ttcag
298	T25 L30a I25-10	ctgtttaatcACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCT TTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaaacacaagtcatatactgagtataa gtataatttcag
299	T25 L30a	cgtaatttatattgtgaatcACGGACTTAAATAATTGAGCCTTAAAGAAGA

[1786]

[1787]

	I25-20	AATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATC TAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAA TTAGTAAGTTAACAATAGATGACTTACAACTAATCGGAAGGT GCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGT AAAGAGAGAGTCCAATTCTCAAAGCCAATAcacaaacacaagtcatatac tgagtataagtataatttcag
300	T25 L30a I50-10	gctatgatccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaaacacaagtcatatactgagtataa gtataatttcag
301	T25 L30a I50-20	cggatattatgctgcgatccACGGACTTAAATAATTGAGCCTTAAAGAAG AAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAA TCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGT AATTAGTAAGTTAACAATAGATGACTTACAACTAATCGGAAG GTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGG GTAAAGAGAGAGTCCAATTCTCAAAGCCAATAcacaaacacaagtcat atactgagtataagtataatttcag
302	T25 L30a I80-20	cggtaggcgcgctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAA GAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAA ATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAG TAATTAGTAAGTTAACAATAGATGACTTACAACTAATCGGAA GGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAcacaaacacaagtc atatactgagtataagtataatttcag
303	T25 L30b I0-0	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCTTTAAG TGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACTAATCGGAAGGTGCAGAGACTC GACGGGAGCTACCCTAACGTCcacaaacacaagtcatatactgagtataagtataat ttcag
304	T25 L30b I25-10	ctgtttaatcACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCT TTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCcacaaacacaagtcatatactgagtat aagtataatttcag
305	T25 L30b I25-20	cgtaatttatattgtgaatcACGGACTTAAATAATTGAGCCTTAAAGAAGA AATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATC TAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAA TTAGTAAGTTAACAATAGATGACTTACAACTAATCGGAAGGT GCAGAGACTCGACGGGAGCTACCCTAACGTCcacaaacacaagtcatat actgagtataagtataatttcag
306	T25 L30b I50-10	gctatgatccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCcacaaacacaagtcatatactgagtat aagtataatttcag
307	T25 L30b I50-20	cggatattatgctgcgatccACGGACTTAAATAATTGAGCCTTAAAGAAG AAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAAA TCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGT

[1788]

		AATTAGTAAGTTAACAATAGATGACTTACAACTAATCGGAAG GTGCAGAGACTCGACGGGAGCTACCCTAACGTCcacaaacacaagtc atatactgagtataagtataatttcag
308	T25 L30b I80-20	cggtaggcgcgctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAA GAAATTCTTTAAGTGGATGCTCTCAAACTCAGGGAAACCTAA ATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAG TAATTAGTAAGTTAACAATAGATGACTTACAACTAATCGGAA GGTGCAGAGACTCGACGGGAGCTACCCTAACGTCcacaaacacaagt catatactgagtataagtataatttcag

[1789] 일부 실시형태에서, 스페이서 및 3' 인트론 단편은 표 23에 열거된 바와 같은 서열을 갖는 스페이서 및 인트론 단편이다.

표 24

CAR 서열

서열번호	CAR	서열
309	FMC63-4-1BB	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTGG CGAGGGATCTACCAAGGGCGAAGTGAAACTGCAAGAGTCTG GCCCTGGACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTG TACAGTCAGCGGAGTGTCCCTGCCTGATTACGGCGTGTCCTG GATCAGACAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGT GATCTGGGGCAGCGAGACAACCTACTACAACAGCGCCCTGAA GTCCCGGCTGACCATCATCAAGGACAACTCCAAGAGCCAGGT GTTCCTGAAGATGAACAGCCTGCAGACCGACGACACCGCCAT CTACTATTGCGCCAAGCACTACTACTACGGCGGCAGCTACGC CATGGATTATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCT GCCGCCGCTATCGAAGTGATGTACCCTCCTCCTTACCTGGACA ACGAGAAGTCCAACGGCACCATCATCCACGTGAAGGGCAAG CACCTGTGTCCTTCTCCACTGTTCCCCGGACCTAGCAAGCCTT TCTGGGTGCTCGTTGTTGTTGGCGGCGTGCTGGCCTGTTACAG CCTGCTGGTTACCGTGGCCTTCATCATCTTTTGGGTCAAGAGA GGCCGGAAGAAACTTCTTTATATATTCAAGCAGCCCTTTATGC GACCCGTTCAGACTACCCAAGAGGAAGATGGATGCAGTTGCC GCTTTCCAGAAGAGGAGGAGGGCGGGTGCGAACTGtaa
310	FMC63-CD28	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTGG CGAGGGATCTACCAAGGGCGAAGTGAAACTGCAAGAGTCTG GCCCTGGACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTG TACAGTCAGCGGAGTGTCCCTGCCTGATTACGGCGTGTCCTG GATCAGACAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGT GATCTGGGGCAGCGAGACAACCTACTACAACAGCGCCCTGAA GTCCCGGCTGACCATCATCAAGGACAACTCCAAGAGCCAGGT GTTCCTGAAGATGAACAGCCTGCAGACCGACGACACCGCCAT CTACTATTGCGCCAAGCACTACTACTACGGCGGCAGCTACGC CATGGATTATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCT GCCGCCGCTATCGAAGTGATGTACCCTCCTCCTTACCTGGACA ACGAGAAGTCCAACGGCACCATCATCCACGTGAAGGGCAAG

[1790]

		CACCTGTGTCCTTCTCCACTGTTCCCCGGACCTAGCAAGCCTT TCTGGGTGCTCGTTGTTGTTGGCGGCGTGCTGGCCTGTTACAG CCTGCTGGTTACCGTGGCCTTCATCATCTTTTGGGTCCGAAGC AAGCGGAGCCGGCTGCTGCACTCCGACTACATGAACATGACC CCTAGACGGCCCGGACCAACCAGAAAGCACTACCAGCCTTAC GCTCCTCCTAGAGACTTCGCCGCCTACCGGTCCtaa
311	FMC63- CD28-제타	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTGG CGAGGGATCTACCAAGGGCGAAGTGAAACTGCAAGAGTCTG GCCCTGGACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTG TACAGTCAGCGGAGTGTCCCTGCCTGATTACGGCGTGTCCTG GATCAGACAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGT GATCTGGGGCAGCGAGACAACCTACTACAACAGCGCCCTGAA GTCCCGGCTGACCATCATCAAGGACAACTCCAAGAGCCAGGT GTTCCTGAAGATGAACAGCCTGCAGACCGACGACACCGCCAT CTACTATTGCGCCAAGCACTACTACTACGGCGGCAGCTACGC CATGGATTATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCT GCCGCCGCTATCGAAGTGATGTACCCTCCTCCTTACCTGGACA ACGAGAAGTCCAACGGCACCATCATCCACGTGAAGGGCAAG CACCTGTGTCCTTCTCCACTGTTCCCCGGACCTAGCAAGCCTT TCTGGGTGCTCGTTGTTGTTGGCGGCGTGCTGGCCTGTTACAG CCTGCTGGTTACCGTGGCCTTCATCATCTTTTGGGTCCGAAGC AAGCGGAGCCGGCTGCTGCACTCCGACTACATGAACATGACC CCTAGACGGCCCGGACCAACCAGAAAGCACTACCAGCCTTAC GCTCCTCCTAGAGACTTCGCCGCCTACCGGTCCAGAGTGAAG TTCAGCAGATCCGCCGATGCTCCCGCCTATCAGCAGGGCCAA AACCAGCTGTACAACGAGCTGAACCTGGGGAGAAGAGAAGA GTACGACGTGCTGGACAAGCGGAGAGGCAGAGATCCTGAAA TGGGCGGCAAGCCCAGACGGAAGAATCCTCAAGAGGGCCTG TATAATGAGCTGCAGAAAGACAAGATGGCCGAGGCCTACAG CGAGATCGGAATGAAGGGCGAGCGCAGAAGAGGCAAGGGAC ACGATGGACTGTACCAGGGACTGAGCACCGCCACCAAGGATA CCTATGACGCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa
312	FMC63-제타	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTGG CGAGGGATCTACCAAGGGCGAAGTGAAACTGCAAGAGTCTG GCCCTGGACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTG TACAGTCAGCGGAGTGTCCCTGCCTGATTACGGCGTGTCCTG

[1791]

		GATCAGACAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGT GATCTGGGGCAGCGAGACAACCTACTACAACAGCGCCCTGAA GTCCCGGCTGACCATCATCAAGGACAACTCCAAGAGCCAGGT GTTCCTGAAGATGAACAGCCTGCAGACCGACGACACCGCCAT CTACTATTGCGCCAAGCACTACTACTACGGCGGCAGCTACGC CATGGATTATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCT GCCGCCGCTATCGAAGTGATGTACCCTCCTCCTTACCTGGACA ACGAGAAGTCCAACGGCACCATCATCCACGTGAAGGGCAAG CACCTGTGTCCTTCTCCACTGTTCCCCGGACCTAGCAAGCCTT TCTGGGTGCTCGTTGTTGTTGGCGGCGTGCTGGCCTGTTACAG CCTGCTGGTTACCGTGGCCTTCATCATCTTTTGGGTCAGAGTG AAGTTCAGCAGATCCGCCGATGCTCCCGCCTATCAGCAGGGC CAAAACCAGCTGTACAACGAGCTGAACCTGGGGAGAAGAGA AGAGTACGACGTGCTGGACAAGCGGAGAGGCAGAGATCCTG AAATGGGCGGCAAGCCCAGACGGAAGAATCCTCAAGAGGGC CTGTATAATGAGCTGCAGAAAGACAAGATGGCCGAGGCCTAC AGCGAGATCGGAATGAAGGGCGAGCGCAGAAGAGGCAAGGG ACACGATGGACTGTACCAGGGACTGAGCACCGCCACCAAGG ATACCTATGACGCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa
313	CircKymriah - Q388	ATGGCTCTCCCGGTCACAGCCCTTCTCCTGCCCCTGGCACTCT TGCTGCATGCGGCACGACCCGACATCCAGATGACCCAGACCA CAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCATCA GCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTGGT ATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTACC ACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTCTG GCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAACC TGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGGCA ACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGAAA TCACCGGTGGAGGTGGTTCTGGCGGAGGGGGATCTGGTGGAG GCGGTTCAGAAGTGAAACTGCAAGAGTCTGGCCCTGGACTGG TGGCCCCATCTCAGTCTCTGAGCGTGACCTGTACAGTCAGCG GAGTGTCCCTGCCTGATTACGGCGTGTCCTGGATCAGACAGC CTCCTCGGAAAGGCCTGGAATGGCTGGGAGTGATCTGGGGCA GCGAGACAACCTACTACAACAGCGCCCTGAAGTCCCGGCTGA CCATCATCAAGGACAACTCCAAGAGCCAGGTGTTCCTGAAGA TGAACAGCCTGCAGACCGACGACACCGCCATCTACTATTGCG CCAAGCACTACTACTACGGCGGCAGCTACGCCATGGATTATT GGGGCCAGGGCACCAGCGTGACCGTTTCTTCTACCACAACGC CCGCCCCGCGACCGCCTACTCCCGCTCCCACAATTGCATCACA ACCCCTGTCTTTGAGACCCGAAGCTTGTCGACCAGCTGCCGGT GGCGCGGTTCACACGCGGGGGCTCGATTTCGCCTGTGATATA TATATATGGGCCCCATTGGCTGGAACATGCGGAGTATTGCTTC TGAGCCTGGTGATTACCCTCTACTGTAAGAGAGGCCGGAAGA AACTTCTTTATATATTCAAGCAGCCCTTTATGCGACCCGTTCA GACTACCCAAGAGGAAGATGGATGCAGTTGCCGCTTTCCAGA AGAGGAGGAGGGCGGGTGCGAACTGAGAGTGAAGTTCAGCA GATCCGCCGATGCTCCCGCCTATCAGCAGGGCCAAAACCAGC TGTACAACGAGCTGAACCTGGGGAGAAGAGAAGAGTACGAC GTGCTGGACAAGCGGAGAGGCAGAGATCCTGAAATGGGCGG CAAGCCCAGACGGAAGAATCCTCAAGAGGGCCTGTATAATGA GCTGCAGAAAGACAAGATGGCCGAGGCCTACAGCGAGATCG GAATGAAGGGCGAGCGCAGAAGAGGCAAGGGACACGATGGA CTGTACCAGGGACTGAGCACCGCCACCAAGGATACCTATGAC GCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa
314	CircKymriah -	ATGGCTCTCCCGGTCACAGCCCTTCTCCTGCCCCTGGCACTCT

[1792]

	K388	TGCTGCATGCGGCACGACCCGACATCCAGATGACCCAGACCA CAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCATCA GCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTGGT ATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTACC ACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTCTG GCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAACC TGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGGCA ACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGAAA TCACCGGTGGAGGTGGTTCTGGCGGAGGGGGATCTGGTGGAG GCGGTTCAGAAGTGAAACTGCAAGAGTCTGGCCCTGGACTGG TGGCCCCATCTCAGTCTCTGAGCGTGACCTGTACAGTCAGCG GAGTGTCCCTGCCTGATTACGGCGTGTCCTGGATCAGACAGC CTCCTCGGAAAGGCCTGGAATGGCTGGGAGTGATCTGGGGCA GCGAGACAACCTACTACAACAGCGCCCTGAAGTCCCGGCTGA CCATCATCAAGGACAACTCCAAGAGCCAGGTGTTCCTGAAGA TGAACAGCCTGCAGACCGACGACACCGCCATCTACTATTGCG CCAAGCACTACTACTACGGCGGCAGCTACGCCATGGATTATT GGGGCCAGGGCACCAGCGTGACCGTTTCTTCTACCACAACGC CCGCCCCGCGACCGCCTACTCCCGCTCCCACAATTGCATCACA ACCCCTGTCTTTGAGACCCGAAGCTTGTCGACCAGCTGCCGGT GGCGCGGTTCACACGCGGGGGCTCGATTTCGCCTGTGATATA TATATATGGGCCCCATTGGCTGGAACATGCGGAGTATTGCTTC TGAGCCTGGTGATTACCCTCTACTGTAAGAGAGGCCGGAAGA AACTTCTTTATATATTCAAGCAGCCCTTTATGCGACCCGTTCA GACTACCCAAGAGGAAGATGGATGCAGTTGCCGCTTTCCAGA AGAGGAGGAGGGCGGGTGCGAACTGAGAGTGAAGTTCAGCA GATCCGCCGATGCTCCCGCCTATAAGCAGGGCCAAAACCAGC TGTACAACGAGCTGAACCTGGGGAGAAGAGGAAGAGTACGAC GTGCTGGACAAGCGGAGAGGCAGAGATCCTGAAATGGGCGG CAAGCCCAGACGGAAGAATCCTCAAGAGGGCCTGTATAATGA GCTGCAGAAAGACAAGATGGCCGAGGCCTACAGCGAGATCG GAATGAAGGGCGAGCGCAGAAGAGGCAAGGGACACGATGGA CTGTACCAGGGACTGAGCACCGCCACCAAGGATACCTATGAC GCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa
315	CircM971 - CD22	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCCAGGTTCAACTCCAGCAGTC TGGTCCCGGCCTCGTTAAACCAAGCCAGACTTTGTCTCTTACC TGTGCTATCAGTGGCGATAGCGTGTCTAGTAATTCAGCCGCAT GGAACTGGATCCGACAATCACCGAGTAGGGGACTTGAATGGC TGGGTAGAACCTATTACCGGTCCAAATGGTACAATGACTATG CAGTGTCTGTAAAAAGCAGGATCACGATCAACCCTGATACGT CTAAAAACCAGTTTTCTCTGCAACTTAATAGTGTGACCCCTGA AGACACCGCTGTGTATTACTGTGCACGGGAGGTTACCGGTGA TCTTGAAGATGCTTTTGATATATGGGGCCAAGGTACGATGGT CACGGTGTCTAGTgggggaggcggcagcGACATACAGATGACGCAG AGCCCATCCAGTCTCTCCGCGTCTGTTGGTGACAGAGTGACTA TTACATGTAGGGCGTCTCAGACCATTTGGTCTTACCTCAATTG GTATCAACAGCGACCAGGCAAAGCACCGAACTTGCTCATTTA CGCTGCCAGCTCACTCCAAAGTGGTGTGCCGTCCAGATTTAGT GGTAGGGGCAGTGGCACTGATTTCACTCTGACTATTTCAAGTC TTCAAGCTGAGGATTTTGCCACATACTACTGCCAGCAAAGTT ACTCAATACCTCAGACTTTTGGACAGGGGACAAAATTGGAGA TTAAAtccggaACCACAACGCCCGCCCCGCGACCGCCTACTCCC GCTCCCACAATTGCATCACAACCCCTGTCTTTGAGACCCGAA GCTTGTCGACCAGCTGCCGGTGGCGCGGTTCACACGCGGGGG

[1793]

		CTCGATTTCGCCTGTGATATATATATATGGGCCCCATTGGCTG GAACATGCGGAGTATTGCTTCTGAGCCTGGTGATTACCCTCTA CTGTAAGAGAGGCCGGAAGAAACTTCTTTATATATTCAAGCA GCCCTTTATGCGACCCGTTCAGACTACCCAAGAGGAAGATGG ATGCAGTTGCCGCTTTCCAGAAGAGGAGGAGGGCGGGTGCGA ACTGAGAGTGAAGTTCAGCAGATCCGCCGATGCTCCCGCCTA TAAGCAGGGCCAAAACCAGCTGTACAACGAGCTGAACCTGG GGAGAAGAGAAGAGTACGACGTGCTGGACAAGCGGAGAGGC AGAGATCCTGAAATGGGCGGCAAGCCCAGACGGAAGAATCC TCAAGAGGGCCTGTATAATGAGCTGCAGAAAGACAAGATGG CCGAGGCCTACAGCGAGATCGGAATGAAGGGCGAGCGCAGA AGAGGCAAGGGACACGATGGACTGTACCAGGGACTGAGCAC CGCCACCAAGGATACCTATGACGCCCTGCACATGCAGGCCCT GCCTCCAAGAtaa
316	CircCD19_22 이중특이적 29	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCggcggcggaggatccCAGGTTCAACTCCAGCAGTCTGGTC CCGGCCTCGTTAAACCAAGCCAGACTTTGTCTCTTACCTGTGC TATCAGTGGCGATAGCGTGTCTAGTAATTCAGCCGCATGGAA CTGGATCCGACAATCACCGAGTAGGGGACTTGAATGGCTGGG TAGAACCTATTACCGGTCCAAATGGTACAATGACTATGCAGT GTCTGTAAAAAGCAGGATCACGATCAACCCTGATACGTCTAA AAACCAGTTTTCTCTGCAACTTAATAGTGTGACCCCTGAAGAC ACCGCTGTGTATTACTGTGCACGGGAGGTTACCGGTGATCTTG AAGATGCTTTTGATATATGGGGCCAAGGTACGATGGTCACGG TGTCTAGTGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTG GCGAGGGATCTACCAAGGGCGACATACAGATGACGCAGAGC CCATCCAGTCTCTCCGCGTCTGTTGGTGACAGAGTGACTATTA CATGTAGGGCGTCTCAGACCATTTGGTCTTACCTCAATTGGTA TCAACAGCGACCAGGCAAAGCACCGAACTTGCTCATTTACGC TGCCAGCTCACTCCAAAGTGGTGTGCCGTCCAGATTTAGTGGT AGGGGCAGTGGCACTGATTTCACTCTGACTATTTCAAGTCTTC AAGCTGAGGATTTTGCCACATACTACTGCCAGCAAAGTTACT CAATACCTCAGACTTTTGGACAGGGGACAAAATTGGAGATTA AAggggaggcggcagcGAAGTGAAACTGCAAGAGTCTGGCCCTGG ACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTGTACAGTC AGCGGAGTGTCCCTGCCTGATTACGGCGTGTCCTGGATCAGA CAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGTGATCTGG GGCAGCGAGACAACCTACTACAACAGCGCCCTGAAGTCCCGG CTGACCATCATCAAGGACAACTCCAAGAGCCAGGTGTTCCTG AAGATGAACAGCCTGCAGACCGACGACACCGCCATCTACTAT TGCGCCAAGCACTACTACTACGGCGGCAGCTACGCCATGGAT TATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCTtccggaACC ACAACGCCCGCCCCGCGACCGCCTACTCCCGCTCCCACAATT GCATCACAACCCCTGTCTTTGAGACCCGAAGCTTGTCGACCA GCTGCCGGTGGCGCGGTTCACACGCGGGGGCTCGATTTCGCC TGTGATATATATATATGGGCCCCATTGGCTGGAACATGCGGA GTATTGCTTCTGAGCCTGGTGATTACCCTCTACTGTAAGAGAG

[1794]

		GCCGGAAGAAACTTCTTTATATATTCAAGCAGCCCTTTATGCG ACCCGTTCAGACTACCCAAGAGGAAGATGGATGCAGTTGCCG CTTTCCAGAAGAGGAGGAGGGCGGGTGCGAACTGAGAGTGA AGTTCAGCAGATCCGCCGATGCTCCCGCCTATAAGCAGGGCC AAAACCAGCTGTACAACGAGCTGAACCTGGGGAGAAGAGAA GAGTACGACGTGCTGGACAAGCGGAGAGGCAGAGATCCTGA AATGGGCGGCAAGCCCAGACGGAAGAATCCTCAAGAGGGCC TGTATAATGAGCTGCAGAAAGACAAGATGGCCGAGGCCTACA GCGAGATCGGAATGAAGGGCGAGCGCAGAAGAGGCAAGGGA CACGATGGACTGTACCAGGGACTGAGCACCGCCACCAAGGAT ACCTATGACGCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa
317	CircCD19_22 이중특이적 30	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCCAGGTTCAACTCCAGCAGTC TGGTCCCGGCCTCGTTAAACCAAGCCAGACTTTGTCTCTTACC TGTGCTATCAGTGGCGATAGCGTGTCTAGTAATTCAGCCGCAT GGAACTGGATCCGACAATCACCGAGTAGGGGACTTGAATGGC TGGGTAGAACCTATTACCGGTCCAAATGGTACAATGACTATG CAGTGTCTGTAAAAAGCAGGATCACGATCAACCCTGATACGT CTAAAAACCAGTTTTCTCTGCAACTTAATAGTGTGACCCCTGA AGACACCGCTGTGTATTACTGTGCACGGGAGGTTACCGGTGA TCTTGAAGATGCTTTTGATATATGGGGCCAAGGTACGATGGT CACGGTGTCTAGTggggaggcggcagcGACATACAGATGACGCAG AGCCCATCCAGTCTCTCCGCGTCTGTTGGTGACAGAGTGACTA TTACATGTAGGGCGTCTCAGACCATTTGGTCTTACCTCAATTG GTATCAACAGCGACCAGGCAAAGCACCGAACTTGCTCATTTA CGCTGCCAGCTCACTCCAAAGTGGTGTGCCGTCCAGATTTAGT GGTAGGGGCAGTGGCACTGATTTCACTCTGACTATTTCAAGTC TTCAAGCTGAGGATTTTGCCACATACTACTGCCAGCAAAGTT ACTCAATACCTCAGACTTTTGGACAGGGGACAAAATTGGAGA TTAAAGGGGGAGGCGGATCCGGCGGTGGTGGCTCCGGCGGTG GTGGTTCTGGAGGCGGCGGAAGCGGTGGGGGTGGTAGCGAC ATCCAGATGACCCAGACCACAAGCAGCCTGTCTGCCAGCCTG GGCGATAGAGTGACCATCAGCTGTAGAGCCAGCCAGGACATC AGCAAGTACCTGAACTGGTATCAGCAAAAGCCCGACGGCACC GTGAAGCTGCTGATCTACCACACCAGCAGACTGCACAGCGGC GTGCCAAGCAGATTTTCTGGCAGCGGCTCTGGCACCGACTAC AGCCTGACAATCAGCAACCTGGAACAAGAGGATATCGCTACC TACTTCTGCCAGCAAGGCAACACCCTGCCTTACACCTTTGGCG GAGGCACCAAGCTGGAAATCACCGGCTCTACAAGCGGCAGC GGCAAACCTGGATCTGGCGAGGGATCTACCAAGGGCGAAGT GAAACTGCAAGAGTCTGGCCCTGGACTGGTGGCCCCATCTCA GTCTCTGAGCGTGACCTGTACAGTCAGCGGAGTGTCCCTGCCT GATTACGGCGTGTCCTGGATCAGACAGCCTCCTCGGAAAGGC CTGGAATGGCTGGGAGTGATCTGGGGCAGCGAGACAACCTAC TACAACAGCGCCCTGAAGTCCCGGCTGACCATCATCAAGGAC AACTCCAAGAGCCAGGTGTTCCTGAAGATGAACAGCCTGCAG ACCGACGACACCGCCATCTACTATTGCGCCAAGCACTACTAC TACGGCGGCAGCTACGCCATGGATTATTGGGGCCAGGGCACC AGCGTGACCGTTTCTTCTtccggaACCACAACGCCCGCCCCGCGA CCGCCTACTCCCGCTCCCACAATTGCATCACAACCCCTGTCTT TGAGACCCGAAGCTTGTCGACCAGCTGCCGGTGGCGCGGTTC ACACGCGGGGGCTCGATTTCGCCTGTGATATATATATATGGG CCCCATTGGCTGGAACATGCGGAGTATTGCTTCTGAGCCTGGT GATTACCCTCTACTGTAAGAGAGGCCGGAAGAAACTTCTTTA TATATTCAAGCAGCCCTTTATGCGACCCGTTCAGACTACCCAA

[1795]

		GAGGAAGATGGATGCAGTTGCCGCTTTCCAGAAGAGGAGGA GGGCGGGTGCGAACTGAGAGTGAAGTTCAGCAGATCCGCCG ATGCTCCCGCCTATAAGCAGGGCCAAAACCAGCTGTACAACG AGCTGAACCTGGGGAGAAGAGAAGAGTACGACGTGCTGGAC AAGCGGAGAGGCAGAGATCCTGAAATGGGCGGCAAGCCCAG ACGGAAGAATCCTCAAGAGGGCCTGTATAATGAGCTGCAGAA AGACAAGATGGCCGAGGCCTACAGCGAGATCGGAATGAAGG GCGAGCGCAGAAGAGGCAAGGGACACGATGGACTGTACCAG GGACTGAGCACCGCCACCAAGGATACCTATGACGCCCTGCAC ATGCAGGCCCTGCCTCCAAGAtaa
318	FMC63 CD28 항-CD19 인간 CAR	MLLVTSLLLCELPHPAFLLIPDIQMTQTTSSLSASLGDRVTISCR ASQDISKYLNWYQQKPDGTVKLLIYHTSRLHSGVPSRFSGSGSG TDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGTKLEITGSTSGSG KPGSGEGSTKGEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYG VSWIRQPPRKGLEWLGVIWGSETTYYNSALKSRLTIIKDNSKSQV FLKMNSLQTDDTAIYYCAKHYYYGGSYAMDYWGQGTSVTVSS AAAIEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPLFPGPSKPFWV LVVVGGVLACYSLLVTVAFIIFWVRSKRSRLLHSDYMNMTPRRP GPTRKHYQPYAPPRDFAAYRSRVKFSRSADAPAYQQGQNQLYN ELNLGRREEYDVLDKRRGRDPEMGGKPRRKNPQEGLYNELQKD KMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQ ALPPR
319	FMC63 4-1BB 항-CD19 인간 CAR	MALPVTALLLPLALLLHAARPDIQMTQTTSSLSASLGDRVTISCR ASQDISKYLNWYQQKPDGTVKLLIYHTSRLHSGVPSRFSGSGSG TDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGTKLEITGGGGSG GGGSGGGGSEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVS WIRQPPRKGLEWLGVIWGSETTYYNSALKSRLTIIKDNSKSQVFL KMNSLQTDDTAIYYCAKHYYYGGSYAMDYWGQGTSVTVSSTT TPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACDIYI WAPLAGTCGVLLLSLVITLYCKRGRKKLLYIFKQPFMRPVQTTQ EEDGCSCRFPEEEEGGCELRVKFSRSADAPAYKQGQNQLYNELN LGRREEYDVLDKRRGRDPEMGGKPRRKNPQEGLYNELQKDKM AEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQAL PPR
320	항-BCMA 인간 CAR	MALPVTALLLPLALLLHAARPDIVLTQSPASLAVSLGERATINCR ASESVSVIGAHLIHWYQQKPGQPPKLLIYLASNLETGVPARFSGS GSGTDFTLTISSLQAEDAAIYYCLQSRIFPRTFGQGTKLEIKGSTS GSGKPGSGEGSTKGQVQLVQSGSELKKPGASVKVSCKASGYTFT DYSINWVRQAPGQGLEWMGWINTETREPAYAYDFRGRFVFSLD TSVSTAYLQISSLKAEDTAVYYCARDYSYAMDYWGQGTLVTVS SAAATTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLD FACDIYIWAPLAGTCGVLLLSLVITLYCKRGRKKLLYIFKQPFMR PVQTTQEEDGCSCRFPEEEEGGCELRVKFSRSADAPAYQQGQNQ LYNELNLGRREEYDVLDKRRGRDPEMGGKPRRKNPQEGLYNEL QKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDAL HMQALPPR
321	1D3ζ 마우스 CD19 CAR; 염소 코브바이러스 IRES	MGVPTQLLGLLLLWITDAICDIQMTQSPASLSTSLGETVTIQCQA SEDIYSGLAWYQQKPGKSPQLLIYGASDLQDGVPSRFSGSGSGT QYSLKITSMQTEDEGVYFCQQGLTYPRTFGGGTKLELKGGGGSG GGGSGGGGSEVQLQQSGAELVRPGTSVKLSCKVSGDTITFYYM HFVKQRPGQGLEWIGRIDPEDESTKYSEKFKNKATLTADTSSNT AYLKLSSLTSEDTATYFCIYGGYYFDYWGQGVMVTVSSIEFMYP PPYLDNERSNGTIIHIKEKHLCHTQSSPKLFWALVVVAGVLFCYG

[1796]

		LLVTVALCVIWTNSRRNRGGQSDYMNMTPRRPGLTRKPYQPYA PARDFAAYRPRAKFSRSAETAANLQDPNQLFNELNLGRREEFDV LEKKRARDPEMGGKQQRRRNPQEGVYNALQKDKMAEAYSEIG TKGERRRGKGHDGLFQGLSTATKDTFDALHMQTLAPR
322	WTζ (야생형 제타) 마우스 CAR	MGVPTQLLGLLLLWITDAICDIQMTQSPASLSTSLGETVTIQCQA SEDIYSGLAWYQQKPGKSPQLLIYGASDLQDGVPSRFSGSGSGT QYSLKITSMQTEDEGVYFCQQGLTYPRTFGGGTKLELKGGGGSG GGGSGGGGSEVQLQQSGAELVRPGTSVKLSCKVSGDTITFYYM HFVKQRPGQGLEWIGRIDPEDESTKYSEKFKNKATLTADTSSNT AYLKLSSLTSEDTATYFCIYGGYYFDYWGQGVMVTVSSIEFMYP PPYLDNERSNGTIIHIKEKHLCHTQSSPKLFWALVVVAGVLFCYG LLVTVALCVIWTNSRRNRGGQSDYMNMTPRRPGLTRKPYQPYA PARDFAAYRPRAKFSRSAETAANLQDPNQLYNELNLGRREEYD VLEKKRARDPEMGGKQQRRRNPQEGVYNALQKDKMAEAYSEI GTKGERRRGKGHDGLYQGLSTATKDTYDALHMQTLAPR

[1797]

[1798] 일부 실시형태에서, CAR은 표 24에 열거된 바와 같은 뉴클레오타이드 서열에 의해서 암호화된다.

표 25

CAR 도메인 서열.

서열번호	단백질	서열
318	4-1BB	KRGRKKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEGGCEL
319	CD3ζ 세포내 도메인	RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRRGRDPE MGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGKGHD GLYQGLSTATKDTYDALHMQALPPR
320	CD28 세포내 신호전달 도메인	QVQLVQSGAEVEKPGASVKVSCKASGYTFTDYYMHWVRQAPGQ GLEWMGWINPNSGGTNYAQKFQGRVTMTRDTSISTAYMELSRLR SDDTAVYYCASGWDFDYWGQGTLVTVSSGGGGSGGGGSGGGGS GGGGSDIVMTQSPSSLSASVGDRVTITCRASQSIRYYLSWYQQKP GKAPKLLIYTASILQNGVPSRFSGSGSGTDFTLTISSLQPEDFATYY CLQTYTTPDFGPGTKVEIK
321	FMC63 VH	EVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIRQPPRKGLE WLGVIWGSETTYYNSALKSRLTIIKDNSKSQVFLKMNSLQTDDTA IYYCAKHYYYGGSYAMDYWGQGTSVTVSS
322	FMC63 VL	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPDGTVK LLIYHTSRLHSGVPSRFSGSGSGTDYSLTISNLEQEDIATYFCQQGN TLPYTFGGGTKLEIT

[1799]

[1800] 일부 실시형태에서, 본 발명의 폴리뉴클레오타이드에 의해서 암호화된 CAR 도메인은 표 25에 열거된 바와 같은 서열을 갖는다.

표 26

PD-1 또는 PD-L1 서열.

서열번호	설명	서열
323	펨브롤리주맙 중쇄	QVQLVQSGVEVKKPGASVKVSCKASGYTFTNYYMYWV RQAPGQGLEWMGGINPSNGGTNFNEKFKNRVTLTTDSST TTAYMELKSLQFDDTAVYYCARRDYRFDMGFDYWGQG TTVTVSSASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFP EPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSS SLGTKTYTCNVDHKPSNTKVDKRVESKYGPPCPPCPAPE FLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEV QFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQ DWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPREPQVYT LPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN NYKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVM HEALHNHYTQKSLSLSLGK
324	펨브롤리주맙	EIVLTQSPATLSLSPGERATLSCRASKGVSTSGYSYLHWY QQKPGQAPRLLIYLASYLESGVPARFSGSGSGTDFTLTISS LEPEDFAVYYCQHSRDLPLTFGGGTKVEIKRTVAAPSVFI FPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS GNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACE VTHQGLSSPVTKSFNRGEC
325	니볼루맙 중쇄	QVQLVESGGGVVQPGRSLRLDCKASGITFSNSGMHWVR QAPGKGLEWVAVIWYDGSKRYYADSVKGRFTISRDNSK NTLFLQMNSLRAEDTAVYYCATNDDYWGQGTLVTVSSA STKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSW NSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTKTY TCNVDHKPSNTKVDKRVESKYGPPCPPCPAPEFLGGPSVF LFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVD GVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKE YKCKVSNKGLPSSIEKTISKAKGQPREPQVYTLPPSQEEM TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV LDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNH YTQKSLSLSLGK
326	니볼루맙 경쇄	EIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKP GQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTISSLEPE DFAVYYCQQSSNWPRTFGQGTKVEIKRTVAAPSVFIFPPS DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNS QESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTH QGLSSPVTKSFNRGEC
327	아테졸리주맙 중쇄	EVQLVESGGGLVQPGGSLRLSCAASGFTFSDSWIHWVRQ APGKGLEWVAWISPYGGSTYYADSVKGRFTISADTSKNT AYLQMNSLRAEDTAVYYCARRHWPGGFDYWGQGTLVT VSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPV TVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLG TQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPE LLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV KFNWYVDGVEVHNAKTKPREEQYASTYRVVSVLTVLHQ DWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYT LPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN NYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVM HEALHNHYTQKSLSLSPGK

[1801]

328	아테졸리주맙 경쇄	DIQMTQSPSSLSASVGDRVTITCRASQDVSTAVAWYQQK PGKAPKLLIYSASFLYSGVPSRFSGSGSGTDFTLTISSLQPE DFATYYCQQYLYHPATFGQGTKVEIKRTVAAPSVFIFPPS DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNS QESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTH QGLSSPVTKSFNRGEC
329	아벨루맙 중쇄	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYIMMWVRQ APGKGLEWVSSIYPSGGITFYADTVKGRFTISRDNSKNTL YLQMNSLRAEDTAVYYCARIKLGTVTTVDYWGQGTLVT VSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPV TVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLG TQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPE LLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV KFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQ DWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYT LPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN YKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMH EALHNHYTQKSLSLSPGK
330	아벨루맙 경쇄	QSALTQPASVSGSPGQSITISCTGTSSDVGGYNYVSWYQQ HPGKAPKLMIYDVSNRPSGVSNRFSGSKSGNTASLTISGL QAEDEADYYCSSYTSSSTRVFGTGTKVTVLGQPKANPTV TLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADGSPV KAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQ VTHEGSTVEKTVAPTECS
331	더발루맙 중쇄	EVQLVESGGGLVQPGGSLRLSCAASGFTFSRYWMSWVR QAPGKGLEWVANIKQDGSEKYYVDSVKGRFTISRDNAK NSLYLQMNSLRAEDTAVYYCAREGGWFGELAFDYWGQ GTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDY FPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVP SSSLGTQTYICNVNHKPSNTKVDKRVEPKSCDKTHTCPPC PAPEFEGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHE DPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLT VLHQDWLNGKEYKCKVSNKALPASIEKTISKAKGQPREP QVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNG QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS CSVMHEALHNHYTQKSLSLSPGK
332	더발루맙 경쇄	EIVLTQSPGTLSLSPGERATLSCRASQRVSSSYLAWYQQK PGQAPRLLIYDASSRATGIPDRFSGSGSGTDFTLTISRLEPE DFAVYYCQQYGSLPWTFGQGTKVEIKRTVAAPSVFIFPPS DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNS QESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTH QGLSSPVTKSFNRGEC

[1802]

[1803] 일부 실시형태에서, 본 발명의 폴리뉴클레오타이드에 의해서 암호화된 발현 서열을 분리하는 절단 부위는 표 26에 열거된 서열을 갖는다.

표 27

사이토카인 서열.

서열번호	사이토카인	서열
333	IL-2	APTSSSTKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFY MPKKATELKHLQCLEEELKPLEEVLNLAQSKNFHLRPRDLISNINV IVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTLT
334	IL-12A	RNLPVATPDPGMFPCLHHSQNLLRAVSNMLQKARQTLEFYPCTSE EIDHEDITKDKTSTVEACLPLELTKNESCLNSRETSFITNGSCLASR KTSFMMALCLSSIYEDLKMYQVEFKTMNAKLLMDPKRQIFLDQN MLAVIDELMQALNFNSETVPQKSSLEEPDFYKTKIKLCILLHAFRIR AVTIDRVMSYLNAS
335	IL-12B	IWELKKDVYVVELDWYPDAPGEMVVLTCDTPEEDGITWTLDQSS EVLGSGKTLTIQVKEFGDAGQYTCHKGGEVLSHSLLLLHKKEDGI WSTDILKDQKEPKNKTFLRCEAKNYSGRFTCWWLTTISTDLTFSV KSSRGSSDPQGVTCGAATLSAERVRGDNKEYEYSVECQEDSACPA AEESLPIEVMVDAVHKLKYENYTSSFFIRDIIKPDPPKNLQLKPLKN SRQVEVSWEYPDTWSTPHSYFSLTFCVQVQGKSKREKKDRVFTD KTSATVICRKNASISVRAQDRYYSSSWSEWASVPCS
336	IL-7	DCDIEGKDGKQYESVLMVSIDQLLDSMKEIGSNCLNNEFNFFKRHI CDANKEGMFLFRAARKLRQFLKMNSTGDFDLHLLKVSEGTTILLN CTGQVKGRKPAALGEAQPTKSLEENKSLKEQKKLNDLCFLKRLL QEIKTCWNKILMGTKEH
337	IL-10	SPGQGTQSENSCTHFPGNLPNMLRDLRDAFSRVKTFFQMKDQLD NLLLKESLLEDFKGYLGCQALSEMIQFYLEEVMPQAENQDPDIKA HVNSLGENLKTLRLRLRRCHRFLPCENKSKAVEQVKNAFNKLQE KGIYKAMSEFDIFINYIEAYMTMKIRN
338	IL-15	NWVNVISDLKKIEDLIQSMHIDATLYTESDVHPSCKVTAMKCFLL ELQVISLESGDASIHDTVENLIILANNSLSSNGNVTESGCKECEELE EKNIKEFLQSFVHIVQMFINTS
339	IL-18	YFGKLESKLSVIRNLNDQVLFIDQGNRPLFEDMTDSDCRDNAPRTI FIISMYKDSQPRGMAVTISVKCEKISTLSCENKIISFKEMNPPDNIKD TKSDIIFFQRSVPGHDNKMQFESSSYEGYFLACEKERDLFKLILKK EDELGDRSIMFTVQNED
340	IL-27베타	RKGPPAALTLPRVQCRASRYPIAVDCSWTLPPAPNSTSPVSFIATY RLGMAARGHSWPCLQQTPTSTSCTITDVQLFSMAPYVLNVTAVH PWGSSSSFVPFITEHIIKPDPPEGVRLSPLAERQLQVQWEPPGSWPF PEIFSLKYWIRYKRQGAARFHRVGPIEATSFILRAVRPRARYYVQV AAQDLTDYGELSDWSLPATATMSLGK
341	IFN감마	QDPYVKEAENLKKYFNAGHSDVADNGTLFLGILKNWKEESDRKI MQSQIVSFYFKLFKNFKDDQSIQKSVETIKEDMNVKFFNSNKKKR DDFEKLTNYSVTDLNVQRKAIHELIQVMAELSPAAKTGKRKRSQ MLFRG
342	TGF베타1	ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFC LGPCPYIWSLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIV YYVGRKPKVEQLSNMIVRSCKCS

[1804]

[1805] 일부 실시형태에서, 본 발명의 폴리뉴클레오타이드에 의해서 암호화된 사이토카인은 표 27에 열거된 바와 같은 서열을 갖는다.

표 28

전사 인자 서열

서열번호	전사 인자	서열
343	FOXP3	MPNPRPGKPSAPSLALGPSPGASPSWRAAPKASDLLGARGPGGT FQGRDLRGGAHASSSSLNPMPPSQLQLPTLPLVMVAPSGARLGP LPHLQALLQDRPHFMHQLSTVDAHARTPVLQVHPLESPAMISLT PPTTATGVFSLKARPGLPPGINVASLEWVSREPALLCTFPNPSAPR KDSTLSAVPQSSYPLLANGVCKWPGCEKVFEEPEDFLKHCQADH LLDEKGRAQCLLQREMVQSLEQQLVLEKEKLSAMQAHLAGKM ALTKASSVASSDKGSCCIVAAGSQGPVVPAWSGPREAPDSLFAV RRHLWGSHGNSTFPEFLHNMDYFKFHNMRPPFTYATLIRWAILE APEKQRTLNEIYHWFTRMFAFFRNHPATWKNAIRHNLSLHKCFV RVESEKGAVWTVDELEFRKKRSQRPSRCSNPTPGP
344	FOXP3	MPNPRPGKPSAPSLALGPSPGASPSWRAAPKASDLLGARGPGGT FQGRDLRGGAHASSSSLNPMPPSQLQLPTLPLVMVAPSGARLGP LPHLQALLQDRPHFMHQLSTVDAHARTPVLQVHPLESPAMISLT PPTTATGVFSLKARPGLPPGINVASLEWVSREPALLCTFPNPSAPR KDSTLSAVPQSSYPLLANGVCKWPGCEKVFEEPEDFLKHCQADH LLDEKGRAQCLLQREMVQSLEQQLVLEKEKLSAMQAHLAGKM ALTKASSVASSDKGSCCIVAAGSQGPVVPAWSGPREAPDSLFAV RRHLWGSHGNSTFPEFLHNMDYFKFHNMRPPFTYATLIRWAILE APEKQRTLNEIYHWFTRMFAFFRNHPATWKNAIRHNLSLHKCFV RVESEKGAVWTVDELEFRKKR
345	FOXP3	GGAHASSSSLNPMPPSQLQLPTLPLVMVAPSGARLGPLPHLQAL LQDRPHFMHQLSTVDAHARTPVLQVHPLESPAMISLTPPTTATG VFSLKARPGLPPGINVASLEWVSREPALLCTFPNPSAPRKDSTLS AVPQSSYPLLANGVCKWPGCEKVFEEPEDFLKHCQADHLLDEK GRAQCLLQREMVQSLEQQLVLEKEKLSAMQAHLAGKMALTKA SSVASSDKGSCCIVAAGSQGPVVPAWSGPREAPDSLFAVRRHLW GSHGNSTFPEFLHNMDYFKFHNMRPPFTYATLIRWAILEAPEKQ RTLNEIYHWFTRMFAFFRNHPATWKNAIRHNLSLHKCFVRVESE KGAVWTVDELEFRKKR
346	STAT5B	MAVWIQAQQLQGEALHQMQALYGQHFPIEVRHYLSQWIESQA WDSVDLDNPQENIKATQLLEGLVQELQKKAEHQVGEDGFLLKI KLGHYATQLQNTYDRCPMELVRCIRHILYNEQRLVREANNGSSP AGSLADAMSQKHLQINQTFEELRLVTQDTENELKKLQQTQEYFI QYQESLRIQAQFGPLAQLSPQERLSRETALQQKQVSLEAWLQRE AQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQLA GNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG PVEEMLAEVNATITDIISALVTSTFIIEKQPPQVLKTQTKFAATVR LLVGGKLNVHMNPPQVKATIISEQQAKSLLKNENTRNDYSGEIL NNCCVMEYHQATGTLSAHFRNMSLKRIKRSDRRGAESVTEEKF TILFESQFSVGGNELVFQVKTLSLPVVVIVHGSQDNNATATVLW DNAFAEPGRVPFAVPDKVLWPQLCEALNMKFKAEVQSNRGLTK ENLVFLAQKLFNNSSSHLEDYSGLSVSWSQFNRENLPGRNYTFW QWFDGVMEVLKKHLKPHWNDGAILGFVNKQQAHDLLINKPDG TFLLRFSDSEIGGITIAWKFDSQERMFWNLMPFTTRDFSIRSLADR LGDLNYLIYVFPDRPKDEVYSKYYTPVPCESATAKAVDGYVKPQ IKQVVPEFVNASADAGGGSATYMDQAPSPAVCPQAHYNMYPQ NPDSVLDTDGDFDLEDTMDVARRVEELLGRPMDSQWIPHAQS
347	HELIOS	METEAIDGYITCDNELSPEREHSNMAIDLTSSTPNGQHASPSHMT STNSVKLEMQSDEECDRKPLSREDEIRGHDEGSSLEEPLIESSEVA DNRKVQELQGEGGIRLPNGKLKCDVCGMVCIGPNVLMVHKRSH

[1806]

		TGERPFHCNQCGASFTQKGNLLRHIKLHSGEKPFKCPFCSYACRR RDALTGHLRTHSVGKPHKCNYCGRSYKQRSSLEEHKERCHNYL QNVSMEAAGQVMSHHVPPMEDCKEQEPIMDNNISLVPFERPAVI EKLTGNMGKRKSSTPQKFVGEKLMRFSYPDIHFDMNLTYEKEA ELMQSHMMDQAINNAITYLGAEALHPLMQHPPSTIAEVAPVISS AYSQVYHPNRIERPISRETADSHENNMDGPISLIRPKSRPQEREAS PSNSCLDSTDSESSHDDHQSYQGHPALNPKRKQSPAYMKEDVK ALDTTKAPKGSLKDIYKVFNGEGEQIRAFKCEHCRVLFLDHVMY TIHMGCHGYRDPLECNICGYRSQDRYEFSSHIVRGEHTFH

[1807]

[1808] 일부 실시형태에서, 본 발명의 폴리뉴클레오타이드에 의해서 암호화된 전사 인자는 표 28에 열거된 바와 같은 서열을 갖는다.

표 29

추가 액세서리 서열

서열번호	IRES	서열
390	CK 3' UTR Scr	ccctgcagccgtcaccgtaagtttgaagttaccgcatatcagcctctgcttcccagcgcgtccaatt cctgttcttattgtttcccctccaggcgttacgcgtgacgacgaactgtgtcgcagctaccacattatt ccggagccttcattctcgcggctctgatcgt
391	CK 3' UTR S2M	ggagaccgcggccacgccgagtaggatcgagggtacagtctcc
392	CK 3' UTR	gacaccaggatcactcttgctctgacccgccctgtgtagaatagactcatgcttccctaagacctgg atttcttcccaggcactttcacccgcctgccctgctccttcagtggactgcacccagggaggcggtc tctgactgtcctttactttctattctggattgc
393	CK 5' UTR 1	AAACCCCCCTAAGCCGCCGCCGCCGCCACC
394	CK 5' UTR 2	CCCCCCCAACCCGTCACG
395	CK 5' UTR 3	GTCACG
396	SZ1 3' UTR Scr	tctgcgcactcgtaatcagtactaacccccctttgtcggacactatgcgataatcgatccgccttttc accgccttcggaattttatttacctcaactgatcctggagtctctcttggttttcacggaggcctccgcc ca
397	SZ1 S2M	ggagaccgcggccacgccgagtaggatcgagggtacagtctcc
398	SZ1 3' UTR	ccccttgaaaccccgccccaggttcagtctctcttcatccctctgtcctgcatggtgatacaaagac cctttgtggaccctaagccatgtagttgctgctccctccttccagttgtgaatattggtttctgttaatca ca
399	SZ1 5' UTR 1	AAACCCCCCTAAGCCGCCGCCGCCGCCACC
400	SZ1 5' UTR 2	CCCCCCCAACCCGTCACG
401	SZ1 5' UTR 3	GTCACG
402	UTR1	gTcacG
403	UTR2	AATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGC CACC
404	UTR3	cgaactagtattcttctggtccccacagactcagagagaacccgccacc
405	UTR4	Agccacc
406	STOP1	tgatAGctAaCtaG
407	STOP2	tagtAGctAaCtaG
408	STOP3	tGatGActGaGtGA
409	STOP4	tagtagctagGtag
410	STOP5	taa
411	STOP6	taatagCtaaCtag
412	STOP7	taaCtagCtaaCtag
718	CCL1	MQIITTALVCLLLAGMWPEDVDSKSMQVPFSRCCFSFAEQEIP LRAILCYRNTSSICSNEGLIFKLKRGKEACALDTVGWVQRHRK MLRHCPSKRK
719	CCL2	MKVSAALLCLLLIAATFIPQGLAQPDAINAPVTCCYNFTNRKIS VQRLAS
720	CCL3	MQVSTAALAVLLCTMALCNQFSASLAADTPTACCFSYTSRQI PQNFIADYFETSSQCSKPGVIFLTKRSRQVCADPSEEWVQKYV SDLELSA
721	CCL4	MKLCVTVLSLLMLVAAFCSPALSAPMGSDPPTACCFSYTARK LPRNFVVDYYETSSLCSQPAVVFQTKRSKQVCADPSESWVQE YVYDLELN
722	CCL5	MKVSAAALAVILIATALCAPASASPYSSDTTPCCFAYIARPLPR AHIKEYFYTSGKCSNPAVVFVTRKNRQVCANPEKKWVREYIN SLEMS

[1809]

723	CCL6	MRNSKTAISFFILVAVLGSQAGLIQEMEKEDRRYNPPIIHQGFQ DTSSDCCFSYATQIPCKRFIYYFPTSGGCIKPGIIFISRRGTQVCA DPSDRRVQRCLSTLKQGPRSGNKVIA
724	CCL7	MKASAALLCLLLTAAAFSPQGLAQPVGINTSTTCCYRFINKKI PKQRLESYRRTTSSHCPREAVIFKTKLDKEICADPTQKWVQDF MKHLDKKTQTPKL
725	CCL8	MKVSAALLCLLLMAATFSPQGLAQPDSVSIPITCCFNVINRKIP IQRLESYTRITNIQCPKEAVIFKTKRGKEVCADPKERWVRDSM KHLDQIFQNLKP
726	CCL9/CCL10	MKPFHTALSFLILTTALGIWAQITHATETKEVQSSLKAQQGLEI EMFHMGFQDSSDCCLSYNSRIQCSRFIGYFPTSGGCTRPGIIFIS KRGFQVCANPSDRRVQRCIERLEQNSQPRTYKQ
727	CCL11	MKVSAALLWLLLIAAAFSPQGLAGPASVPTTCCFNLANRKIPL QRLESYRRITSGKCPQKAVIFKTKLAKDICADPKKKWVQDSM KYLDQKSPTPKP
728	CCL12	MKISTLLCLLLIATTISPQVLAGPDAVSTPVTCCYNVVKQKIHV RKLKSYRRITSSQCPREAVIFRTILDKEICADPKEKWVKNSINH LDKTSQTFILEPSCLG
729	CCL13	MKVSAVLLCLLLMTAAFNPQGLAQPDALNVPSTCCFTFSSKKI SLQRLKSYVITTSRCPQKAVIFRTKLGKEICADPKEKWVQNY MKHLGRKAHTLKT
730	CCL14	MKISVAAIPFFLLITIALGTKTESSSRGPYHPSECCFTYTTYKIPR QRIMDYYETNSQCSKPGIVFITKRGHSVCTNPSDKWVQDYIKD MKEN
731	CCL15	MKVSVAALSCLMLVAVLGSQAQFINDAETELMMSKLPLENP VVLNSFHFAADCCTSYISQSIPCSLMKSYFETSSECSKPGVIFLT KKGRQVCAKPSGPGVQDCMKKLKPYSI
732	CCL16	MKVSEAALSLLVLILIITSASRSQPKVPEWVNTPSTCCLKYYEK VLPRRLVVGYRKALNCHLPAIIFVTKRNREVCTNPNDDWVQE YIKDPNLPLLPTRNLSTVKIITAKNGQPQLLNSQ
733	CCL17	MAPLKMLALVTLLLGASLQHIHAARGTNVGRECCLEYFKGAI PLRKLKTWYQTSEDCSRDAIVFVTVQGRAICSDPNNKRVKNA VKYLQSLERS
734	CCL18	MKGLAAALLVLVCTMALCSCAQVGTNKELCCLVYTSWQIPQ KFIVDYSETSPQCPKPGVILLTKRGRQICADPNKKWVQKYISD LKLNA
735	CCL19	MALLLALSLLVLWTSPAPTLSGTNDAEDCCLSVTQKPIPGYIV RNFHYLLIKDGCRVPAVVFTTLRGRQLCAPPDQPWVERIIQRL QRTSAKMKRRSS
736	CCL20	MCCTKSLLLAALMSVLLLHLCGESEAASNFDCCLGYTDRILH PKFIVGFTRQLANEGCDINAIIFHTKKKLSVCANPKQTWVKYI VRLLSKKVKNM
737	CCL21	MAQSLALSLLILVLAFGIPRTQGSDGGAQDCCLKYSQRKIPAK VVRSYRKQEPSLGCSIPAILFLPRKRSQAELCADPKELWVQQL MQHLDKTPSPQKPAQGCRKDRGASKTGKKGKGSKGCKRTER SQTPKGP
738	CCL22	MDRLQTALLVVLVLLAVALQATEAGPYGANMEDSVCCRDY VRYRLPLRVVKHFYWTSDSCPRPGVVLLTFRDKEICADPRVP WVKMILNKLSQ
739	CCL23	MKVSVAALSCLMLVTALGSQARVTKDAETEFMMSKLPLENP VLLDRFHATSADCCISYTPRSIPCSLLESYFETNSECSKPGVIFL TKKGRRFCANPSDKQVQVCVRMLKLDTRIKTRKN

[1810]

740	CCL24	MAGLMTIVTSLLFLGVCAHHIIPTGSVVIPSPCCMFFVSKRIPE NRVVSYQLSSRSTCLKAGVIFTTKKGQQFCGDPKQEWVQRY MKNLDAKQKKASPRARAVAVKGPVQRYPGNQTTC
741	CCL25	MNLWLLACLVAGFLGAWAPAVHTQGVFEDCCLAYHYPIGW AVLRRAWTYRIQEVSGSCNLPAAIFYLPKRHRKVCGNPKSRE VQRAMKLLDARNKVFAKLHHNTQTFQAGPHAVKKLSSGNSK LSSSKFSNPISSSKRNVSLLISANSGL
742	CCL26	MMGLSLASAVLLASLLSLHLGTATRGSDISKTCCFQYSHKPLP WTWVRSYEFTSNSCSQRAVIFTTKRGKKVCTHPRKKWVQKYI SLLKTPKQL
743	CCL27	MKGPPTFCSLLLLSLLLSPDPTAAFLLPPSTACCTQLYRKPLSD KLLRKVIQVELQEADGDCHLQAFVLHLAQRSICIHPQNPSLSQ WFEHQERKLHGTLPKLNFGMLRKMG
744	CCL28	MQQRGLAIVALAVCAALHASEAILPIASSCCTEVSHHISRRLLE RVNMCRIQRADGDCDLAAVILHVKRRRICVSPHNHTVKQWM KVQAAKKNGKGNVCHRKKHHGKRNSNRAHQGKHETYGHK TPY
745	CXCL1	MARAALSAAPSNPRLLRVALLLLLLVAAGRRAAGASVATELR CQCLQTLQGIHPKNIQSVNVKSPGPHCAQTEVIATLKNGRKAC LNPASPIVKKIIEKMLNSDKSN
746	CXCL2	MARATLSAAPSNPRLLRVALLLLLLVAASRRAAGAPLATELR CQCLQTLQGIHLKNIQSVKVKSPGPHCAQTEVIATLKNGQKA CLNPASPMVKKIIEKMLKNGKSN
747	CXCL3	MAHATLSAAPSNPRLLRVALLLLLLVAASRRAAGASVVTELR CQCLQTLQGIHLKNIQSVNVRSPGPHCAQTEVIATLKNGKKAC LNPASPMVQKIIEKILNKGSTN
748	CXCL4	MAYYEESIIFDTDNSSEESGDFDVDFQGPCERAPNYDFHKIFLP TVFGIIFILGIFGNGLVVIVMGFQNKCKTSMTDKYRLHLSVAD LMFVLTLPFWAVDAASSWYFGGFLCKAVHMIYTVNLYSSVLI LAFISLDRYLAVVRATATSSQATRKLLAGKVIYVCVWLPAAIL TVPDLVFASSFEVELDMGGSRMICQRIYPLESNHIWVAAFRFQ HILVGFVLPGLVILTCYCIIISKLSQGSKGQALKRKALKTTVILI LCFFCCWLPYCAGIFVDTLMNLEVIPHSCALEQGVMKWI 부위 VLAYFHCCLNPILYAFLGVKFKKSARNALTFSSRSSHKMLTKK RGAISSASTESESSSVLYS
749	CXCL5	MSLLSSRAARVPGPSSSLCALLVLLLLLTQPGPIASAGPAAAVL RELRCVCLQTTQGVHPKMISNLQVFAIGPQCSKVEVVASLKN GKEICLDPEAPFLKKVIQKILDGGNKEN
750	CXCL6	MSLPSSRAARVPGPSGSLCALLALLLLLTPPGPLASAGPVSAV LTELRCTCLRVTLRVNPKTIGKLQVFPAGPQCSKVEVVASLKN GKQVCLDPEAPFLKKVIQKILDSGNKKN
751	CXCL7	MSLRLDTTPSCNSARPLHALQVLLLLSLLLTALASSTKGQTKR NLAKGKEESLDSDLYAELRCMCIKTTSGIHPKNIQSLEVIGKGT HCNQVEVIATLKDGRKICLDPDAPRIKKIVQKKLAGDESAD
752	CXCL8	MTSKLAVALLAAFLISAALCEGAVLPRSAKELRCQCIKTYSKP FHPKFIKELRVIESGPHCANTEIIVKLSDGRELCLDPKENWVQR VVEKFLKRAENS
753	CXCL9	MKKSGVLFLLGIILLVLIGVQGTPVVRKGRCSCISTNQGTIHLQ SLKDLKQFAPSPSCEKIEIIATLKNGVQTCLNPDSADVKELIKK WEKQVSQKKKQKNGKKHQKKKVLKVRKSQRSRQKKTT
754	CXCL10	MNQTAILICCLIFLTLSGIQGVPLSRTVRCTCISISNQPVNPRSLE KLEIIPASQFCPRVEIIATMKKKGEKRCLNPESKAIKNLLKAVS KERSKRSP

[1811]

755	CXCL11	MSVKGMAIALAVILCATVVQGFPMFKRGRCLCIGPGVKAVK VADIEKASIMYPSNNCDKIEVIITLKENKGQRCLNPKSKQARLI IKKVERKNF
756	CXCL12	MNAKVVVVLVLVLTALCLSDGKPVSLSYRCPCRFFESHVARA NVKHLKILNTPNCALQIVARLKNNNRQVCIDPKLKWIQEYLE KALNKRFKM
757	CXCL13	MKFISTSLLLMLLVSSLSPVQGVLEVYYTSLRCRCVQESSVFIP RRFIDRIQILPRGNGCPRKEIIVWKKNKSIVCVDPQAEWIQRM MEVLRKRSSSTLPVPVFKRKIP
758	CXCL14	MSLLPRRAPPVSMRLLAAALLLLLLALYTARVDGSKCKCSRK GPKIRYSDVKKLEMKPKYPHCEEKMVIITTKSVSRYRGQEHCL HPKLQSTKRFIKWYNAWNEKRRVYEE
759	CXCL15	MAAQGWSMLLLAVLNLGIFVRPCDTQELRCLCIQEHSEFIPLK LIKNIMVIFETIYCNRKEVIAVPKNGSMICLDPDAPWVKATVG PITNRFLPEDLKQKEFPPAMKLLYSVEHEKPLYLSFGRPENKRI FPFPIRETSRHFADLAHNSDRNFLRDSSEVSLTGSDA
760	CXCL16	MGRDLRPGSRVLLLLLLLLLVYLTQPGNGNEGSVTGSCYCGK RISSDSPPSVQFMNRLRKHLRAYHRCLYYTRFQLLSWSVCGG NKDPWVQELMSCLDLKECGHAYSGIVAHQKHLLPTSPPISQA SEGASSDIHTPAQMLLSTLQSTQRPTLPVGSLSSDKELTRPNET TIHTAGHSLAAGPEAGENQKQPEKNAGPTARTSATVPVLCLL AIIFILTAALSYVLCKRRRGQSPQSSPDLPVHYIPVAP
761	CXCL17	MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRW LQEGGQECECKDWFLRAPRRKFMTVSGLPKKQCPCDHFKGN VKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL
762	XCL1	MRLLILALLGICSLTAYIVEGVGSEVSDKRTCVSLTTQRLPVSR IKTYTITEGSLRAVIFITKRGLKVCADPQATWVRDVVRSMDRK SNTRNNMIQTKPTGTQQSTNTAVTLTG
763	XCL2	MRLLILALLGICSLTAYIVEGVGSEVSHRRTCVSLTTQRLPVSR IKTYTITEGSLRAVIFITKRGLKVCADPQATWVRDVVRSMDRK SNTRNNMIQTKPTGTQQSTNTAVTLTG
764	CX3CL1	MAPISLSWLLRLATFCHLTVLLAGQHHGVTKCNITCSKMTSKI PVALLIHYQQNQASCGKRAIILETRQHRLFCADPKEQWVKDA MQHLDRQAAALTRNGGTFEKQIGEVKPRTTPAAGGMDESVV LEPEATGESSSLEPTPSSQEAQRALGTSPELPTGVTGSSGTRLPP TPKAQDGGPVGTELFRVPPVSTAATWQSSAPHQPGPSLWAEA KTSEAPSTQDPSTQASTASSPAPEENAPSEGQRVWGQGQSPRP ENSLEREEMGPVPAHTDAFQDWGPGSMAHVSVVPVSSEGTPS REPVASGSWTPKAEEPIHATMDPQRLGVLITPVPDAQAATRR QAVGLLAFLGLLFCLGVAMFTYQSLQGCPRKMAGEMAEGLR YIPRSCGSNSYVLVPV
765	Tbet	HLLWSKFNQHQTEMIITKQGRRMFPFLSFTVAGLEPTSHYRMF VDVVLVDQHHWRYQSGKWVQCGKAEGSMPGNRLYVHPDSP NTGAHWMRQEVSFGKLKLTNNKGASNNVTQMIVLQSLHKY QPRLHIVEVNEGEPETVCNASNTHIFTFQETQFIAVTAYQNAEI TQLKIDNNPFAKGFRENFESMYASVDTSVPSPPGPNCQLLGGD PYSPLLSNQYPVP
766	GATA3	MEVTADQPRWVSHHHPAVLNGQHPDTHHPGLSHSYMDAAQ YPLPEEVDVLFNIDGQGNHVPPYYGNSVRATVQRYPPTHHGS QVCRPPLLHGSLPWLDGGKALGSHHTASPWNLSPFSKTSIHH GSPGPLSVYPPASSSSLSGGHASPHLFTFPPTPPKDVSPDPSLST PGSAGSARQDEKECLKYQVPLPDSMKLESSHSRGSMTALGGA SSSTHHPITTYPPYVPEYSSGLFPPSSLLGGSPTGFGCKSRPKAR SSTGRECVNCGATSTPLWRRDGTGHYLCNACGLYHKMNGQN

[1812]

		RPLIKPKRRLSAARRAGTSCANCQTTTTTLWRRNANGDPVCN ACGLYYKLHNINRPLTMKKEGIQTRNRKMSSKSKKCKKVHD SLEDFPKNSSFNPAALSRHMSSLSHISPFSHSSHMLTTPTPMHP PSSLSFGPHHPSSMVTAMG
767	RORgt	MDRAPQRQHRASRELLAAKKTHTSQIEVIPCKICGDKSSGIHY GVITCEGCKGFFRRSQRCNAAYSCTRQQNCPIDRTSRNRCQHC RLQKCLALGMSRDAVKFGRMSKKQRDSLHAEVQKQLQQRQ QQQQEPVVKTPPAGAQGADTLTYTLGLPDGQLPLGSSPDLPE ASACPPGLLKASGSGPSYSNNLAKAGLNGASCHLEYSPERGK AEGRESFYSTGSQLTPDRCGLRFEEHRHPGLGELGQGPDSYGS PSFRSTPEAPYASLTEIEHLVQSVCKSYRETCQLRLEDLLRQRS NIFSREEVTGYQRKSMWEMWERCAHHLTEAIQYVVEFAKRL SGFMELCQNDQIVLLKAGAMEVVLVRMCRAYNADNRTVFFE GKYGGMELFRALGCSELISSIFDFSHSLSALHFSEDEIALYTAL VLINAHRPGLQEKRKVEQLQYNLELAFHHHLCKTHRQSILAK LPPKGKLRSLCSQHVERLQIFQHLHPIVVQAAFPPLYKELFSTE TESPVGLSK
768	Cd25	MEPCLLMWGILTFITVSGYTTDLCDDDPPNLKHATFKALTYK TGTVLNCDCERGFRRISSYMHCTGNSSHASWENKCQCKSISPE NRKGKVTTKPEEQKGENPTEMQSQTPPMDEVDLVGHCREPPP WEHENSKRIYHFVVGQTLHYQCMQGFTALHRGPAKSICKTIF GKTRWTQPPLKCISESQFPDDEELQASTDAPAGRDTSSPFITTS TPDFHKHTEVATTMESFIFTTEYQIAVASCVLLLISIVLVSGLT WQRRRRKSRTI
769	Skp1	MPSIKLQSSDGEIFEVDVEIAKQSVTIKTMLEDLGMDDEGDDD PVPLPNVNAAILKKVIQWCTHHKDDPPPPEDDENKEKRTDDIP VWDQEFLKVDQGTLFELILAANYLDIKGLLDVTCKTVANMIK GKTPEEIRKTFNIKNDFTEEEEAQVRKENQWCEEK
770	Spy	MRKLTALFVASTLALGAANLAHAADTTTAAPADAKPMMHH KGKFGPHQDMMFKDLNLTDAQKQQIREIMKGQRDQMKRPPL EERRAMHDIIASDTFDKAKAEAQIAKMEEQRKANMLAHMET QNKIYNILTPEQKKQFNANFEKRLTERPAAKGKMPATAE
771	Fkpa	MKSLFKVTLLATTMAVALHAPITFAAEAAKPATAADSKAAFK NDDQKSAYALGASLGRYMENSLKEQEKLGIKLDKDQLIAGV QDAFADKSKLSDQEIEQTLQAFEARVKSSAQAKMEKDAADN EAKGKEYREKFAKEKGVKTSSTGLVYQVVEAGKGEAPKDSD TVVVNYKGTLIDGKEFDNSYTRGEPLSFRLDGVIPGWTEGLK NIKKGGKIKLVIPPELAYGKAGVPGIPPNSTLVFDVELLDVKPA PKADAKPEADAKAADSAKK
772	Sura	MKNWKTLLLGIAMIANTSFAAPQVVDKVAAVVNNGVVLESD VDGLMQSVKLNAAQARQQLPDDATLRHQIMERLIMDQIILQM GQKMGVKISDEQLDQAIANIAKQNNMTLDQMRSRLAYDGLN YNTYRNQIRKEMIISEVRNNEVRRRITILPQEVESLAQQVGNQ NDASTELNLSHILIPLPENPTSDQVNEAESQARAIVDQARNGA DFGKLAIAHSADQQALNGGQMGWGRIQELPGIFAQALSTAKK GDIVGPIRSGVGFHILKVNDLRGESKNISVTEVHARHILLKPSPI MTDEQARVKLEQIAADIKSGKTTFAAAAKEFSQDPGSANQGG DLGWATPDIFDPAFRDALTRLNKGQMSAPVHSSFGWHLIELL DTRNVDKTDAAQKDRAYRMLMNRKFSEEAASWMQEQRASA YVKILSN
773	Hsp60	MLRVNSKSSIKTFVRHLSHKELKFGVEGRAALLKGVNTLADA VSVTLGPKGRNVLIEQQFGAPKITKDGVTVAKAITLEDKFEDL GAKLLQEVASKTNESAGDGTTSATVLGRSIFTESVKNVAAGC NPMDLRRGSQAAVEAVIEFLQKNKKEITTSEEIAQVATISANG

[1813]

		DKHIGDLLANAMEKVGKEGVITVKEGKTLEDELEVTEGMKF DRGFISPYFITNTKTGKVEFENPLILLSEKKISSIQDILPSLELSN QTRRPLLIIAEDVDGEALAACILNKLRGQVQVCAVKAPGFGD NRKNTLGDIAILSGGTVFTEELDIKPENATIEQLGSAGAVTITK EDTVLLNGEGSKDNLEARCEQIRSVIADVHTTEYEKEKLQERL AKLSGGVAVIKVGGASEVEVGEKKDRYEDALNATRAAVEEG ILPGGGTALIKATKILDEVKEKAVNFDQKLGVDTIRAAITKPA KRIIENAGEEGAVIVGKIYDEPEFNKGYDSQKGEFTDMIAAGII DPFKVVKNGLVDASGVASLLATTECAIVDAPQPKGSPAAPPA PGMGGMPGMF
774	Hsp70	MAPAVGIDLGTTYSCVGIFRDDRIEIIANDQGNRTTPSFVAFTD TERLIGDAAKNQVAMNPANTVFDAKRLIGRKFADPEVQADM KHFPFKITDKGGKPVIQVEFKGETKEFTPEEISSMVLTKMRETA EAYLGGTVNNAVVTVPAYFNDSQRQATKDAGLIAGLNVLRII NEPTAAAIAYGLDKKADGERNVLIFDLGGGTFDVSLLTIEEGIF EVKSTAGDTHLGGEDFDNRLVNHFVSEFKRKFKKISPAERAR ALRRSPTACERAKRTLSSAAQTSIEIDSLYEGIDFYTSITRARFE ELCQDLFRSTMEPVERVLRDAKIDKSSVHEIVLVGGSTRIPRIQ KLVSDFFNGKEPNKSINPDEAVAYGAAVQAAILSGDTSSKSTN EILLLDVAPLSLGIETAGGVMTPLIKRNTTIPTKKSETFSTFSDN QPGVLIQVFEGERARTKDNNLLGKFELTGIPRARGVPQIEVTF DVDANGIMNVSALEKGTRKTNKIVITNDKGRLSKEEIERMLA EAEKYKAEDEAEASRIRPKNGLESYAYSLRNSLRHSKVDEKL EAGDKEKLKSEIDKTVQWLDENQTATKEEYESQQKELEAVA NPIMMKFYAGGEGAPGGFPGAGGPGGFPGGPGAGHASGGGD DGPTVEEVDLKFPMLPLPWQLSVRKMHRPFFLFLLFLIFLIFLIL FLFYFFLPVRFNESCFS
775	GroEL	MAAKDVKFGNDARVKMLRGVNVLADAVKVTLGPKGRNVV LDKSFGAPTITKDGVSVAREIELEDKFENMGAQMVKEVASKA NDAAGDGTTTATVLAQAIITEGLKAVAAGMNPMDLKRGIDK AVTAAVEELKALSVPCSDSKAIAQVGTISANSDETVGKLIAEA MDKVGKEGVITVEDGTGLQDELDVVEGMQFDRGYLSPYFIN KPETGAVELESPFILLADKKISNIREMLPVLEAVAKAGKPLLIIA EDVEGEALATLVVNTMRGIVKVAAVKAPGFGDRRKAMLQDI ATLTGGTVISEEIGMELEKATLEDLGQAKRVVINKDTTTIIDGV GEEAAIQGRVAQIRQQIEEATSDYDREKLQERVAKLAGGVAV IKVGAATEVEMKEKKARVEDALHATRAAVEEGVVAGGGVA LIRVASKLADLRGQNEDQNVGIKVALRAMEAPLRQIVLNCGE EPSVVANTVKGGDGNYGYNAATEEYGNMIDMGILDPTKVTR SALQYAASVAGLMITTECMVTDLPKNDAADLGAAGGMGGM GGMGGMM
776	GroES	MNIRPLHDRVIVKRKEVETKSAGGIVLTGSAAAKSTRGEVLA VGNGRILENGEVKPLDVKVGDIVIFNDGYGVKSEKIDNEEVLI MSESDILAIVEA
777	Hsp90-알파	MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIF LRELISNSSDALDKIRYESLTDPSKLDSGKELHINLIPNKQDRTL TIVDTGIGMTKADLINNLGTIAKSGTKAFMEALQAGADISMIG QFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRT DTGEPMGRGTKVILHLKEDQTEYLEE RRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEDKEEE KEKEEKESEDKPEIEDVGSDEEEEKKDGDKKKKKKIKEKYIDQ EELNKTKPIWTRNPDDITNEEYGEFYKSLTNDWEDHLAVKHF SVEGQLEFRALLFVPRRAPFDLFENRKKKNNIKLYVRRVFIMD NCEELIPEYLNFIRGVVDSEDLPLNISR

[1814]

		EMLQQSKILKVIRKNLVKKCLELFTELAEDKENYKKFYEQFSK NIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMK ENQKHIYYITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEY CVQQLKEFEGKTLVSVTKEGLELPEDEEEKKKQEEKKTKFEN LCKIMKDILEKKVEKVVVSNRLVTSPCCIV TSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHS IIETLRQKAEADKNDKSVKDLVILLYETALLSSGFSLEDPQTHA NRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPLEGDDDTSR MEEVD
778	HtpG	MKGQETRGFQSEVKQLLHLMIHSLYSNKEIFLRELISNASDAA DKLRFRALSNPDLYEGDGELRVRVSFDKDKRTLTISDNGVGM TRDEVIDHLGTIAKSGTKSFLESLGSDQAKDSQLIGQFGVGFYS AFIVADKVTVRTRAAGEKPENGVFWESAGEGEYTVADITKED RGTEITLHLREGEDEFLDDWRVRSIISKY SDHIALPVEIEKREEKDGETVISWEKINKAQALWTRNKSEITDE EYKEFYKHIAHDFNDPLTWSHNRVEGKQEYTSLLYIPSQAPW DMWNRDHKHGLKLYVQRVFIMDDAEQFMPNYLRFVRGLIDS SDLPLNVSREILQDSTVTRNLRN ALTKRVLQMLEKLAKDDAEKYQTFWQQFGLVLKEGPAEDFA NQEAIAKLLRFASTHTDSSAQTVSLEDYVSRMKEGQEKIYYIT ADSYAAAKSSPHLELLRKKGIEVLLLSDRIDEWMMNYLTEFD GKPFQSVSKVDESLEKLADEVDESAKEAEKALTPFIDRVKALL GERVKDVRLTHRLTDTPAIVSTDADEMSTQM AKLFAAAGQKVPEVKYIFELNPDHVLVKRAADTEDEAKFSE WVELLLDQALLAERGTLEDPNLFIRRMNQLLVS
779	Hsp40	MGKDYYQTLGLARGASDEEIKRAYRRQALRYHPDKNKEPGA EEKFKEIAEAYDVLSDPRKREIFDRYGEEGLKGSGPSGGSGGG ANGTSFSYTFHGDPHAMFAEFFGGRNPFDTFFGQRNGEEGMD IDDPFSGFPMGMGGFTNVNFGRSRSAQEPARKKQDPPVTHDL RVSLEEIYSGCTKKMKISHKRLNPDGKSIRNEDKILTIEVKKG WKEGTKITFPKEGDQTSNNIPADIVFVLKDKPHNIFKRDGSDVI YPARISLREALCGCTVNVPTLDGRTIPVVFKDVIRPGMRRKVP GEGLPLPKTPEKRGDLIIEFEVIFPERIPQTSRTVLEQVLPI
780	ClpP	MWPGILVGGARVASCRYPALGPRLAAHFPAQRPPQRTLQNGL ALQRCLHATATRALPLIPIVVEQTGRGERAYDIYSRLLRERIVC VMGPIDDSVASLVIAQLLFLQSESNKKPIHMYINSPGGVVTAG LAIYDTMQYILNPICTWCVGQAASMGSLLLAAGTPGMRHSLP NSRIMIHQPSGGARGQATDIAIQAEEIMKLKKQLYNIYAKHTK QSLQVIESAMERDRYMSPMEAQEFGILDKVLVHPPQDGEDEP TLVQKEPVEAAPAAEPVPAST
781	ClpX	MPSCGACTCGAAAVRLITSSLASAQRGISGGRIHMSVLGRLGT FETQILQRAPLRSFTETPAYFASKDGISKDGSGDGNKKSASEGS SKKSGSGNSGKGG NQLRCPKCGDLCTHVETFVSSTRFVKCEKCHHFFVVLSEADS KKSIIKEPESAAEAVKLAFQQKPPPPPKKIYNYLDKYVVGQSF AKKVLSVAVYNHYKRIYNNIPANLRQQAEVEKQTSLTPRELEI RRREDEYRFTKLLQIAGISPHG NALGASMQQQVNQQIPQEKRGGEVLDSSHDDIKLEKSNILLL GPTGSGKTLLAQTLAKCLDVPFAICDCTTLTQAGYVGEDIESV IAKLLQDANYNVEKAQQGIVFLDEVDKIGSVPGIHQLRDVGG EGVQQGLLKLLEGTIVNVPEKNSRKLRGETVQVDTTNILFVAS GAFNGLDRIISRRKNEKYLGFGTPSNLGKG RRAAAAADLANRSGESNTHQDIEEKDRLLRHVEARDLIEFGM IPEFVGRLPVVVPLHSLDEKTLVQILTEPRNAVIPQYQALFSMD

[1815]

		KCELNVTEDALKAIARLALERKTGARGLRSIMEKLLLEPMFEV PNSDIVCVEVDKEVVEGKKEPGYIRAPTKESSEEEYDSGVEEE GWPRQADAANS
782	CDKN1C	MSDASLRSTSTMERLVARGTFPVLVRTSACRSLFGPVDHEELS RELQARLAELNAEDQNRWDYDFQQDMPLRGPGRLQWTEVD SDSVPAFYRETVQVGRCRLLLAPRPVAVAVAVSPPLEPAAESL DGLEEAPEQLPSVPVPAPASTPPPVPVLAPAPAPAPAPVAAPV AAPVAVAVLAPAPAPAPAPAPAPAPVAAPAPAPAPAPAPAPA PAPAPDAAPQESAEQGANQGQRGQEPLADQLHSGISGRPAAG TAAASANGAAIKKLSGPLISDFFAKRKRSAPEKSSGDVPAPCPS PSAAPGVGSVEQTPRKRLR
783	BAX 저해제	MNIFDRKINFDALLKFSHITPSTQQHLKKVYASFALCMFVAAA GAYVHMVTHFIQAGLLSALGSLILMIWLMATPHSHETEQKRL GLLAGFAFLTGVGLGPALEFCIAVNPSILPTAFMGTAMIFTCFT LSALYARRRSYLFLGGILMSALSLLLLSSLGNVFFGSIWLFQAN LYVGLVVMCGFVLFDTQLIIEKAEHGD
784	codA	MSNNALQTIINARLPGEEGLWQIHLQDGKISAIDAQSGVMPIT ENSLDAEQGLVIPPFVEPHIHLDTTQTAGQPNWNQSGTLFEGIE RWAERKALLTHDDVKQRAWQTLKWQIANGIQHVRTHVDVS DATLTALKAMLEVKQEVAPWIDLQIVAFPQEGIQDYIWHCIDL FLDFITVFRKLMMILAMNEKDKKKEKKLSYPNGEALLEEALR LGADVVGAIPHFEFTREYGVESLHKTFALAQKYDRLIDVHCD EIDDEQSRFVETVAALAHHEGMGARVTASHTTAMHSYNGAY TSRLFRLLKMSGINFVANPLVNIHLQGRFDTYPKRRGITRVKE MLESGINVCFGHDDVFDPWYPLGTANMLQVLHMGLHVCQL MGYGQINDGLNLITHHSARTLNLQDYGIAAGNSANLIILPAEN GFDALRRQVPVRYSVRGGKVIASTQPAQTTVYLEQPEAIDYK R

[1816]

[1817] 일부 실시형태에서, 본 명세서에 개시된 원형 RNA 또는 전구체 RNA(예를 들어, 선형 전구체 RNA)는 표 29에 열거된 서열을 포함한다.

[1818] 일부 실시형태에서, 폴리뉴클레오타이드 또는 폴리뉴클레오타이드에 의해서 암호화된 단백질은 본 명세서에 개시된 하나 이상의 서열과 적어도 약 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 99.5% 유사성을 갖는 서열을 함유한다. 일부 실시형태에서, 폴리뉴클레오타이드 또는 폴리뉴클레오타이드에 의해서 암호화된 단백질은 본 명세서에 개시된 하나 이상의 서열과 동일한 서열을 함유한다.

[1819] 바람직한 실시형태가 본 명세서에 기재되어 있다. 이러한 바람직한 실시형태의 변형은 전술한 설명을 읽을 때 당업자에게 명백해질 수 있다. 본 발명자들은 숙련된 기술자가 이러한 변형을 적절하게 사용할 것으로 예상하고, 본 발명자들은 본 발명이 본 명세서에 구체적으로 기재된 것과는 달리 실시되기를 의도한다. 따라서, 본 발명은 적용 가능한 법률이 허용하는 바에 따라 본 명세서에 첨부된 청구범위에 인용된 발명 대상의 모든 변형 및 등가물을 포함한다. 또한, 본 명세서에서 달리 나타내지 않거나 문맥상 명백히 모순되지 않는 한, 모든 가능한 변형에서 상기에 기재된 요소의 임의의 조합이 본 발명에 포함된다.

[1820] **실시예**

[1821] 문헌[Wesselhoeft *et al.* (2019) RNA Circularization Diminishes Immunogenicity and Can Extend Translation Duration *In Vivo*. Molecular Cell. 74(3), 508-520 및 Wesselhoeft *et al.* (2018) Engineering circular RNA for Potent and Stable Translation in Eukaryotic Cells. Nature Communications. 9, 2629]은 이의 전문이 참조에 의해 포함된다.

[1822] 본 발명은 하기 실시예를 참조하여 더욱 상세하게 설명되지만, 하기 실시예에 제한되는 것으로 의도되지 않는다. 이들 실시예는 본 발명을 만들고 사용하는 방법에 대한 완전한 개시 및 설명을 당업자에게 제공하려는 의도로 예시의 임의의 변형 및 모든 변형을 포함하며, 본 발명으로 간주되는 것의 범주를 제한하려는 것은 아니다.

[1823] **실시예 1**

[1824] ***실시예 1A:*** *외부 상동성 영역은 순열 인트론 엑손(PIE) 원형화 전략을 사용하여 긴 전구체 RNA의 원형화를 허용한다.*

[1825] 전장 뇌심근염 바이러스(EMCV) IRES, 가우시아 루시퍼라제(GLuc) 발현 서열 및 순열 인트론-엑손(PIE) 작제물의 2개의 짧은 엑손 단편을 함유하는 1,100nt 서열을 T4 파지의 티미딜레이트 합성효소(Td) 유전자 내의 순열 그룹 I 촉매적 인트론의 3' 인트론과 5' 인트론 사이에 삽입하였다. 전구체 RNA를 유출 전사에 의해 합성하였다. 마그네슘 이온 및 GTP의 존재 하에서 전구체 RNA를 가열함으로써 원형화를 시도하였지만, 스플라이싱 생성물을 얻지 못했다.

[1826] 완벽하게 상보적인 9개 뉴클레오타이드 및 19개 뉴클레오타이드 길이 상동성 영역을 설계하고, 전구체 RNA의 5' 단부 및 3' 단부에 추가하였다. 이러한 상동성 아암의 추가는 전구체 RNA 밴드의 소실에 의해 평가된 바와 같이 9개 뉴클레오타이드 상동성 영역의 경우 0에서 16%로, 19개 뉴클레오타이드 상동성 영역의 경우 48%로 스플라이싱 효율을 증가시켰다.

[1827] 스플라이싱 산물을 RNase R로 처리하였다. RNase R-처리된 스플라이싱 반응의 추정 스플라이스 접합부 전체의 서열결정은 결찰된 엑손을 나타내었고, 올리고뉴클레오타이드-표적화 RNase H로의 RNase R-처리된 스플라이싱 반응의 소화는 RNase H-소화된 선형 전구체에 의해 생성된 2개의 밴드와 대조적으로 단일 밴드를 생성시켰다. 이것은 원형 RNA가 9개 또는 19개 뉴클레오타이드 길이의 외부 상동성 영역을 함유하는 전구체 RNA의 스플라이싱 반응의 주요 산물이라는 것을 나타낸다.

[1828] ***실시예 1B:*** *IRES 및 PIE 스플라이스 부위의 2차 구조를 보존하는 스페이서는 원형화 효율을 증가시킨다.*

[1829] 일련의 스페이서를 설계하고, 3' PIE 스플라이스 부위와 IRES 사이에 삽입하였다. 이 스페이서는 IRES, 3' PIE 스플라이스 부위 및/또는 5' 스플라이스 부위의 인트론 서열 내의 2차 구조를 보존하거나 파괴하도록 설계되었다. 2차 구조를 보존하도록 설계된 스페이서 서열을 추가하면 스플라이싱 효율이 87%인 반면, 파괴적인 스페이서 서열을 추가하면 스플라이싱이 검출되지 않았다.

[1830] **실시예 2**

[1831] ***실시예 2A:*** *외부 상동성 영역에 더하여 내부 상동성 영역은 스플라이싱 버블을 생성시키고, 여러 발현 서열의 번역을 허용한다.*

[1832] 구조화되지 않고, 인트론 및 IRES 서열과 비-상동성이고, 스페이서-스페이서 상동성 영역을 함유하도록 스페이

서를 설계하였다. 이들을 외부 상동성 영역, EMCV IRES 및 가우시아 루시퍼라제(총 길이: 1289nt), 파이어플라이 루시퍼라제(2384nt)에 대한 발현 서열, eGFP(1451nt), 인간 에리쓰로포이에틴(1313 nt) 및 Cas9 엔도뉴클레아제(4934nt)를 함유하는 작제물에서 5' 엑손과 IRES 사이 그리고 3' 엑손과 발현 서열 사이에 삽입하였다. 5개 작제물 모두의 원형화를 달성하였다. T4 파지 및 아나베나 인트론을 사용한 작제물의 원형화는 거의 동일하였다. 원형화 효율은 더 짧은 서열에서 더 높았다. 번역을 측정하기 위해, 각각의 작제물을 HEK293 세포에 형질주입시켰다. 가우시아 및 파이어플라이 루시퍼라제 형질주입된 세포는 발광에 의해 측정된 바와 같이 강력한 반응을 생성시켰고, 인간 에리쓰로포이에틴은 에리쓰로포이에틴 circRNA이 형질주입된 세포의 배지에서 검출 가능하였고, EGFP 형광은 EGFP circRNA가 형질주입된 세포에서 관찰되었다. GFP를 구성적으로 발현하는 세포에 GFP에 지향되는 sgRNA와 함께 Cas9 circRNA를 동시에 형질주입시키는 것은 sgRNA-단독 대조군과 비교하여 최대 97%의 세포에서 형광을 제거하였다.

[1833] ***실시예 2B:*** *CVB3 IRES의 사용은 단백질 생산을 증가시킨다.*

[1834] 내부 및 외부 상동성 영역 및 가우시아 루시퍼라제 또는 파이어플라이 루시퍼라제 발현 서열을 함유하는 상이한 IRES를 갖는 작제물을 제조하였다. 형질주입 24시간 후에 HEK293 세포의 상청액에서 발광에 의해 단백질 생산을 측정하였다. 콕사키바이러스 B3(CVB3) IRES 작제물은 두 경우 모두에서 가장 많은 단백질을 생산하였다.

[1835] ***실시예 2C:*** *polyA 또는 polyAC 스페이서의 사용은 단백질 생산을 증가시킨다.*

[1836] 30개의 뉴클레오타이드 길이의 polyA 또는 polyAC 스페이서를 실시예 2B에서 단백질을 생산한 각각의 IRES를 갖는 작제물의 IRES와 스플라이스 접합부 사이에 추가하였다. 형질주입 24시간 후에 HEK293 세포의 상청액에서 발광에 의해 가우시아 루시퍼라제 활성을 측정하였다. 두 스페이서 모두 스페이서가 없는 대조군 작제물에 비해 모든 작제물에서 발현을 개선시켰다.

[1837] **실시예 3**

[1838] *원형 RNA가 형질주입된 HEK293 또는 HeLa 세포는 대등한 비변형 또는 변형된 선형 RNA가 형질주입된 세포보다 더 많은 단백질을 생산한다.*

[1839] HPLC-정제된 가우시아 루시퍼라제-암호화 circRNA(CVB3-GLuc-pAC)를 표준(canonical) 비변형된 5' 메틸구아노신-캡핑 및 3' polyA-테일 선형 GLuc mRNA 및 상업적으로 입수 가능한 뉴클레오사이드-변형(슈도우리딘, 5-메틸사이토신) 선형 GLuc mRNA(트라이링크사(Trilink) 제품)와 비교하였다. 형질주입 24시간 후에 발광을 측정하였는데, 이는 circRNA가 HEK293 세포에서 비변형된 선형 mRNA보다 811.2% 더 많은 단백질을 생산하고, 변형된 mRNA보다 54.5% 더 많은 단백질을 생산한다는 것을 나타내었다. HeLa 세포에서 유사한 결과가 얻어졌으며, 인간 에리쓰로포이에틴을 암호화하는 최적화된 circRNA를 5-메톡시우리딘으로 변형된 선형 mRNA와 비교하였다.

[1840] 발광 데이터를 6일에 걸쳐 수집하였다. HEK293 세포에서, circRNA 형질주입은 80시간의 단백질 생산 반감기를 초래하였고, 이는 비변형된 선형 mRNA의 43시간 및 변형된 선형 mRNA의 45시간과 비교된다. HeLa 세포에서, circRNA 형질주입은 116시간의 단백질 생산 반감기를 초래하였고, 이는 비변형된 선형 mRNA의 44시간 및 변형된 선형 mRNA의 49시간과 비교된다. CircRNA는 두 세포 유형 모두에서 수명 동안 비변형된 선형 mRNA 및 변형된 선형 mRNA 둘 다보다 실질적으로 더 많은 단백질을 생산하였다.

[1841] **실시예 4**

[1842] ***실시예 4A:*** *RNase 소화, HPLC 정제 및 포스파타제 처리에 의한 circRNA의 정제는 면역원성을 감소시킨다. 완전히 정제된 원형 RNA는 비정제되거나 부분적으로 정제된 원형 RNA보다 면역원성이 현저히 낮다. 단백질 발현 안정성 및 세포 생존력은 세포 유형 및 원형 RNA 순도에 좌우된다.*

[1843] 인간 배아 신장 293(HEK293) 및 인간 폐 암종 A549 세포를 다음으로 형질주입시켰다:

[1844] 비정제된 GLuc 원형 RNA 스플라이싱 반응의 산물,

[1845] • 스플라이싱 반응의 RNase R 소화의 산물,

[1846] • 스플라이싱 반응의 RNase R 소화 및 HPLC 정제의 산물,

[1847] • 스플라이싱 반응의 RNase 소화, HPLC 정제 및 포스파타제 처리의 산물.

[1848] • 스플라이싱 반응의 RNase R 소화는 형질주입되지 않은 대조군과 비교하여 A549 세포에서 사이토카인 방출을

방지하기에 충분하지 않았다.

[1849] 비정제된 스플라이싱 반응과 비교하여 인터류킨-6(IL-6)의 상당한 감소 및 인터페론-α1(IFN-α1)의 상당한 증가가 존재하였지만, HPLC 정제의 추가는 사이토카인 방출을 방지하기에 충분하지 않았다.

[1850] HPLC 정제 후 그리고 RNase R 소화 전에 포스파타제 처리의 추가는 A549 세포에서 평가된 모든 상향조절된 사이토카인의 발현을 극적으로 감소시켰다. 분비된 단핵구 화학유인물질 단백질 1(MCP1), IL-6, IFN-α1, 종양 괴사 인자 α(TNFα) 및 IFNγ 유도성 단백질-10(IP-10)은 검출 불가능하거나 형질주입되지 않은 기준선 수준으로 떨어졌다.

[1851] HEK293 세포에서는 어떠한 실질적인 사이토카인 방출도 존재하지 않았다. A549 세포는 더 높은 순도의 원형 RNA가 형질주입되었을 때 GLuc 발현 안정성 및 세포 생존력이 증가하였다. 완전히 정제된 원형 RNA는 형질주입된 293 세포와 유사한 안정성 표현형을 가졌다.

[1852] ***실시예 4B****: 원형 RNA는 유의한 면역원성을 유발하지 않고, RIG-I 리간드가 아니다.*

[1853] A549 세포에 스플라이싱 반응 산물을 형질주입시켰다.

[1854] A549 세포를 다음으로 형질주입시켰다:

[1855] • 비정제된 원형 RNA,

[1856] • 고분자량(선형 및 원형 연결부) RNA,

[1857] • 원형(니킹된(nicked)) RNA,

[1858] • 정제된 원형 RNA의 초기 분획(니킹된 RNA 피크와 더 많이 중첩됨),

[1859] • 정제된 원형 RNA의 후기 분획(니킹된 RNA 피크와 더 적게 중첩됨),

[1860] • 원형화 동안 절단된 인트론, 또는

[1861] • 비히클(즉, 비 형질주입되지 않은 대조군).

[1862] 스플라이싱 반응으로부터 적절하게 순수한 선형 전구체 RNA를 얻는 것이 어렵기 때문에, 전구체 RNA를 스플라이스 부위 결실 돌연변이체(splice site arrival mutation: DS)의 형태로 별도로 합성하고, 정제시켰다. 각각의 경우에 사이토카인 방출 및 세포 생존력을 측정하였다.

[1863] 스플라이싱 반응 내에 존재하는 대부분의 종뿐만 아니라 전구체 RNA에 대한 응답으로 강력한 IL-6, RANTES 및 IP-10 방출이 관찰되었다. 초기 circRNA 분획은 다른 비-circRNA 분획과 대등한 사이토카인 반응을 유도했는데, 이는 상대적으로 소량의 선형 RNA 오염물이라도 A549 세포에서 실질적인 세포 면역 반응을 유도할 수 있다는 것을 나타낸다. 후기 circRNA 분획은 형질주입되지 않은 대조군의 것보다 더 많은 사이토카인 반응을 도출하지 못했다. 형질주입 36시간 후 A549 세포 생존력은 다른 모든 분획과 비교하여 후기 circRNA 분획에서 유의하게 더 컸다.

[1864] 후기 circRNA HPLC 분획, 전구체 RNA 또는 비정제된 스플라이싱 반응을 사용한 A549 세포의 형질주입 시 RIG-I 및 IFN-β1 전사체 유도를 분석하였다. RIG-I 및 IFN-β1 전사체 둘 다의 유도는 전구체 RNA 및 비정제된 스플라이싱 반응보다 후기 circRNA 분획에 대해 더 약했다. 스플라이싱 반응의 RNase R 처리만으로는 이 효과를 제거하기에 충분하지 않았다. 매우 적은 양의 RIG-I 리간드 3p-hpRNA를 원형 RNA에 추가하면 상당한 RIG-I 전사가 유도되었다. HeLa 세포에서, RNase R-소화 스플라이싱 반응의 형질주입은 RIG-I 및 IFN-β1을 유도했지만, 정제된 circRNA는 그렇지 않았다. 전반적으로, HeLa 세포는 A549 세포보다 RNA 종을 오염시키는 데 덜 민감하였다.

[1865] 스플라이싱 반응 또는 완전히 정제된 circRNA를 A549 세포에 형질주입시킨 후 처음 8시간 이내에 RIG-I, IFN-β1, IL-6 및 RANTES 전사체 유도를 모니터링하는 시간 경과 실험은 circRNA에 대한 일시적인 반응을 나타내지 않았다. 정제된 circRNA는 유사하게 RAW264.7 뮤린 대식세포에서 전염증성 전사체를 유도하는 데 실패하였다.

[1866] A549 세포에 EMCV IRES 및 EGFP 발현 서열을 함유하는 정제된 circRNA를 형질주입시켰다. 이것은 전염증성 전사체의 실질적인 유도를 생성하는 데 실패하였다. 이러한 데이터는, 스플라이싱 반응의 비-원형 성분이 이전 연구에서 관찰된 면역원성에 대한 책임이 있으며, circRNA가 RIG-I에 대한 자연 리간드가 아님을 입증한다.

[1867] **실시예 5**

[1868] *원형 RNA는 TLR에 의한 검출을 회피한다.*

[1869] TLR 3, 7 및 8 리포터 세포주에 다중 선형 또는 원형 RNA 작제물로 형질주입시키고, 분비된 배아 알칼리 포스파타제(SEAP)를 측정하였다.

[1870] 인트론 및 상동성 아암 서열을 결실시켜 선형화된 RNA를 작제하였다. 그런 다음 선형 RNA 작제물을 포스파타제로 처리하고(캡핑된 RNA의 경우 캡핑 후) HPLC로 정제시켰다.

[1871] 시도된 형질주입 중 어느 것도 TLR7 리포터 세포에서 반응을 일으키지 않았다. TLR3 및 TLR8 리포터 세포를 캡핑된 선형화된 RNA, 폴리아데닐화된 선형화된 RNA, 닉킹된 circRNA HPLC 분획 및 초기 circRNA 분획에 의해 활성화시켰다. 후기 circRNA 분획 및 m1ψ-mRNA는 어떤 세포주에서도 TLR 매개 반응을 유발하지 않았다.

[1872] 제2 실험에서, circRNA를 2가지 방법을 사용하여 선형화시켰다: 마그네슘 이온의 존재 하에 열을 사용한 circRNA의 처리 및 DNA 올리고뉴클레오타이드-가이드 RNase H 소화. 두 방법 모두 소량의 온전한 circRNA를 포함하는 대부분의 전장 선형 RNA를 산출하였다. TLR3, 7, 8 리포터 세포에 원형 RNA, 열에 의해 분해된 원형 RNA 또는 RNase H에 의해 분해된 원형 RNA를 형질주입시키고, 형질주입 36시간 후에 SEAP 분비를 측정하였다. TLR8 리포터 세포는 두 가지 형태 모두의 분해된 원형 RNA에 대한 응답으로 SEAP를 분비했지만, 모의 형질주입보다 원형 RNA 형질주입에 대한 더 큰 응답을 생성하지 않았다. 시험관내 전사된 선형화된 RNA에 의한 TLR3의 활성화에도 불구하고, TLR3 및 TLR7 리포터 세포에서는 분해 조건 또는 온전한 조건에서 활성화가 관찰되지 않았다.

[1873] **실시예 6**

[1874] *비변형된 원형 RNA는 선형 RNA보다 증가된 지속적인 생체내 단백질 발현을 생성한다.*

[1875] 마우스에 주입하고, HEK293 세포에 비변형된 및 m1ψ-변형된 인간 에리쓰로포이에틴(hEpo) 선형 mRNA 및 circRNA를 형질주입시켰다. m1ψ-mRNA 및 비변형된 circRNA의 등몰 형질주입은 HEK293 세포에서 강력한 단백질 발현을 초래하였다. hEpo 선형 mRNA 및 circRNA는 HEK293 및 A549 세포의 동일한 중량 형질주입 시 GLuc 선형 mRNA 및 circRNA와 비교하여 유사한 상대적 단백질 발현 패턴 및 세포 생존력을 나타냈다.

[1876] 마우스에서, hEpo circRNA 또는 선형 mRNA를 내장 지방에 주입한 후 혈청에서 hEpo가 검출되었다. 비변형된 circRNA의 주입 후 검출된 hEpo는 비변형된 또는 m1ψ-mRNA로부터의 hEpo보다 더 느리게 붕괴되었고, 주입 후 42시간 동안 여전히 존재하였다. 혈청 hEpo는 비정제된 circRNA 스플라이싱 반응 또는 비변형된 선형 mRNA의 주입 시 신속하게 감소하였다. 비정제된 스플라이싱 반응의 주입은 정제된 circRNA를 포함한 다른 RNA에서는 관찰되지 않은 혈청에서 검출 가능한 사이토카인 반응을 생성시켰다.

[1877] **실시예 7**

[1878] *원형 RNA는 생체내에서 또는 시험관내에서 지질 나노입자를 통해서 효과적으로 전달될 수 있다.*

[1879] 정제된 원형 RNA를 이온화 가능한 리피도이드 cKK-E12를 사용하여 지질 나노입자(LNP)로 제형하였다(문헌[Dong *et al.*, 2014; Kauffman *et al.*, 2015]). 입자는 5moU로 변형된 상업적으로 이용 가능한 대조군 선형 mRNA를 함유하는 입자의 것과 유사한 평균 크기, 다분산 지수 및 캡슐화 효율을 갖는 균일한 다중층 구조를 형성하였다.

[1880] 정제된 hEpo circRNA는 LNP에 캡슐화되고 HEK293 세포에 첨가될 때 5moU-mRNA보다 더 큰 발현을 나타냈다. HEK293 세포에서 LNP-RNA의 발현 안정성은 5moU-mRNA 및 circRNA 둘 다에 대해서 약간의 붕괴 지연을 제외하고는 형질주입 시약에 의해 전달된 RNA의 것과 유사하였다. 비변형된 circRNA 및 5moU-mRNA 둘 다는 시험관내에서 RIG-I/IFN-β1을 활성화시키는 데 실패하였다.

[1881] 마우스에서, LNP-RNA는 내장 지방 조직으로의 국소 주사 또는 간으로의 정맥내 전달에 의해 전달되었다. circRNA로부터의 혈청 hEpo 발현은 더 낮았지만, 두 경우 모두 전달 6시간 후 5moU-mRNA의 발현과 대등하였다. 비변형된 LNP-circRNA의 지방 주입 후 검출된 혈청 hEpo는 LNP-5moU-mRNA로부터의 것보다 더 느리게 붕괴되었고, 혈청에 존재하는 발현 붕괴의 지연은 시험관내에서 언급된 것과 유사하지만, LNP-circRNA 또는 LNP-5moU-mRNA의 정맥내 주입 후 혈청 hEpo는 거의 동일한 속도로 붕괴되었다. 이들 중 어떠한 경우에도 혈청 사이토카인 또는 국소 RIG-I, TNFα 또는 IL-6 전사체 유도의 증가는 존재하지 않았다.

[1882] **실시예 8**

[1883] ***실시예 8A:*** *HEK293, HepG2 및 1C1C7 세포에서 IRES에 의한 발현 및 기능적 안정성.*

[1884] 아나베 인트론/엑손 영역, 가우시아 알루시퍼라제 발현 서열 및 다양한 IRES를 포함하는 작제물을 원형화하였다. 각각의 원형화 반응 100ng을 Lipofect아민 MessengerMax를 사용하여 20,000개의 HEK293 세포, HepG2 세포 및 1C1C7 세포에 개별적으로 형질주입시켰다. 각각의 상청액의 발광을 단백질 발현의 척도로서 24시간 후에 평가하였다. HEK293 세포에서, 크로히바이러스 B, 살리바이러스 FHB, 아이치 바이러스, 살리바이러스 HG-J1 및 엔테로바이러스 J IRES를 포함하는 작제물이 24시간에 가장 큰 발광을 생성시켰다(도 1A). HepG2 세포에서, 아이치 바이러스, 살리바이러스 FHB, EMCV-Cf 및 CVA3 IRES를 포함하는 포함한 작제물이 24시간에 높은 발광을 생성시켰다(도 1B). 1C1C7 세포에서, 살리바이러스 FHB, 아이치 바이러스, 살리바이러스 NG-J1 및 살리바이러스 A SZ-1 IRES를 포함하는 작제물이 24시간에 높은 발광을 생성시켰다(도 1C).

[1885] 24시간에 더 큰 발광을 생성시키는 더 큰 IRES의 경향이 관찰되었다. 더 짧은 총 서열 길이는 원형화 효율을 증가시키는 경향이 있으므로, 높은 발현 및 상대적으로 짧은 IRES를 선택하면 개선된 작제물을 생성시킬 수 있다. HEK293 세포에서, 크로히바이러스 B IRES를 사용한 작제물이 특히 유사한 길이의 다른 IRES와 비교하여 가장 높은 발광을 생성시켰다(**도 2A**). IRES 크기에 대해 플로팅된 HepG2 및 1C1C7 세포의 IRES 작제물로부터의 발현을 도 2B 및 도 2C에 도시한다.

[1886] HepG2 및 1C1C7 세포에서 선택된 IRES 작제물의 기능적 안정성을 3일에 걸쳐 측정하였다. 20,000개 세포에 100ng의 각각의 원형화 반응을 형질주입시키고, 완전 배지 교체 후 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 24시간마다 측정하였다. 살리바이러스 A GUT 및 살리바이러스 FHB는 HepG2 세포에서 가장 높은 기능적 안정성을 나타내었고, 살리바이러스 N-J1 및 살리바이러스 FHB는 1C1C7 세포에서 가장 안정적인 발현을 생성시켰다(도 3A 및 도 3B).

[1887] ***실시예 8B:*** *추가 IRES의 스크리닝*

[1888] HEK293 세포에서 추가 IRES 작제물의 기능적 안정성을 측정하였다. 간략하면, 관심 5' 미번역 영역(UTR)을 GenBank로부터 식별하였다. 선택된 UTR UTR을 5' 단부로부터 675nt로 절단하고, *가우시아* 루시퍼라제(Gluc)를 암호화하는 원형 RNA 골격 작제물 내에 그리고 Gluc 암호 영역 앞에 삽입하였다. 원형 RNA를 HEK293 세포에 형질주입하였다. 24시간 후, 상청액을 수집하고, 분비된 Gluc 단백질로부터의 발광을 상업적으로 입수 가능한 시약을 사용하여 측정하였다. 결과를 **도 1D** 및 **도 1E** 및 표 30에 제시하며, 이는 다수의 자연 IRES 서열이 원형 RNA 맥락에서 단백질 발현을 향상시킨다는 것을 시사한다.

표 30

서열번호	IRES	발현	서열번호	IRES	발현
413	RhPV	1.10E+05	553	프롱호른 안테로프 페스티바이러스	1.35E+06
414	할라스타비 아르바 (1x mut)	9.46E+04	554	돼지 페스티바이러스 단리물 분고와나	1.10E+07
415	오스시바이러스	4.55E+07	555	돼지 페스티바이러스 1	9.46E+04
416	카디시바이러스 B	2.10E+05	556	페스티바이러스 지라페-1	4.72E+05
417	PSIV (Xba1의 경우 2x mut)	9.70E+04	557	돼지 열병 바이러스	3.16E+05
418	PSIV IGR	1.01E+05	558	인간 페기바이러스 단리물 JD2B1I	6.85E+05
419	PV 마호니	1.09E+05	559	인간 페기바이러스 단리물 GBV-C-ZJ	N/A
420	REV A	9.44E+04	560	인간 페기바이러스 단리물 JD2B8C	5.36E+05
421	트로피바이러스 A	9.52E+04	561	간염 GB 바이러스 A	N/A
422	시마피바이러스 A	1.27E+05	562	유인원 페기바이러스	8.56E+04
423	사코브바이러스 A FFUP1 (1x mut)	8.82E+06	563	페기바이러스 I	8.02E+04
424	토사바이러스 C NFSM6F	6.84E+05	564	페기바이러스 K	8.07E+04
425	토사바이러스 2 GA7403	5.05E+06	565	테일러병-연관 바이러스	7.84E+04
426	리마바이러스 A	8.42E+05	566	설치류 페기바이러스	1.79E+05
427	라피바이러스 LPXYC222841	2.22E+05	567	인간 페기바이러스 2	3.14E+05
428	라피바이러스 WHWGGF74766	4.53E+06	568	GB 바이러스 C/ G 형 간염 바이러스	1.36E+05
429	포에시바이러스 BCCH-449	3.43E+05	569	말 페기바이러스 1	8.80E+04
430	메기리바이러스 A LY	1.80E+06	570	콜렉스 테일레리 플라비바이러스	8.52E+04
431	메기리바이러스 E	1.10E+07	571	부수루아라 바이러스	8.20E+04
432	메기리바이러스 C	1.24E+05	572	지카 바이러스	8.61E+04
433	루도피바이러스	1.05E+05	573	요코세 바이러스	8.55E+04
434	리부피 바이러스	2.10E+05	574	웨셀스브론 바이러스	N/A
435	아이치바이러스 A FSS693	6.25E+07	575	말 헤파시바이러스	8.40E+04
436	아이치바이러스 KVGH	1.72E+07	576	헤파시바이러스 B	8.84E+04

[1889]

서열번호	IRES	발현	서열번호	IRES	발현
437	아이치바이러스 DV	7.79E+07	577	헤파시바이러스 I	7.50E+04
438	뮤린 코브바이러스 1	1.60E+07	578	헤파시바이러스 J	7.65E+04
439	돼지 코브바이러스 K-30	N/A	579	헤파시바이러스 K	8.91E+04
440	돼지 코브바이러스 XX	1.32E+07	580	이카바이러스	4.41E+06
441	염소 코브바이러스 12Q108	2.87E+08	581	안타크틱 펭귄 바이러스 A	8.42E+04
442	토끼 코브바이러스	3.73E+07	582	포레스트 파우치츠 자이언트 래트 아르테리바이러스	N/A
443	아알리바이러스	2.65E+05	583	아비시바이러스 Pf-CHK1	1.19E+05
444	그루소피바이러스 A	1.09E+05	584	조류 파라믹소바이러스 펭귄	9.91E+04
445	그루소피바이러스 B	2.12E+05	585	뉴캐슬병 바이러스	8.86E+04
446	얀청 오스벡스 그레나디어 앤초비 피코르나바이러스	1.57E+06	586	박쥐 Hp-베타코로나바이러스	8.47E+04
447	칠면조 갈리바이러스 M176	4.37E+05	587	바셀라 알바 엔도나바이러스	7.65E+04
448	팔코바이러스 A1	1.48E+05	588	볼 피톤 니도바이러스	8.25E+04
449	트레모바이러스 B	1.31E+05	589	박쥐 사펠로바이러스	8.05E+04
450	디델피스 아우리타 HAV	1.38E+05	590	박쥐 피코르나바이러스 3	N/A
451	헤파토바이러스 G1	1.41E+05	591	박쥐 피코르나바이러스 2	7.99E+07
452	헤파토바이러스 D	1.47E+06	592	박쥐 피코르나바이러스 1	1.85E+07
453	헤파토바이러스 H2	1.08E+05	593	박쥐 이플라바이러스	9.76E+04
454	헤파토바이러스 I	8.79E+05	594	박쥐 디시바바이러스	7.43E+04
455	헤파토바이러스 C	5.08E+05	595	베타코로나바이러스 HKU24	8.96E+04
456	피피바이러스 A	2.69E+05	596	베타코로나바이러스 잉글랜드 1	8.74E+04
457	피피바이러스 C	1.09E+05	597	분 카디오바이러스 1	2.62E+06

[1890]

서열번호	IRES	발현	서열번호	IRES	발현
458	피피바이러스 E	1.10E+05	598	브레다 바이러스	1.16E+05
459	아루아마바이러스	4.51E+06	599	소 바이러스성 설사 바이러스 3	2.70E+06
460	아비시바이러스 A	1.91E+05	600	소 비염 A 바이러스	3.62E+06
461	아비시바이러스 B	8.68E+04	601	소 피코르나바이러스 단리물 TCH6	1.21E+05
462	크로히바이러스 A	9.96E+04	602	소 니도바이러스 TCH5	1.17E+05
463	분사기시바이러스 B	8.01E+04	603	소 헤파시바이러스	1.89E+05
464	림니피바이러스 A	8.30E+04	604	보트리티스 시네레아 미토바이러스 4 RdRp	9.68E+04
465	림니피바이러스 C	1.35E+05	605	보트리티스 시네레아 미토바이러스 2 RdRp	8.73E+04
466	오리바이러스	6.09E+05	606	개 피코디시스트로바이러스 균주 209	2.79E+06
467	HAV FH1	1.24E+05	607	개 디스템퍼 바이러스	3.02E+05
468	HAV HM175	4.96E+05	608	개 코브바이러스	1.48E+08
469	파레코바이러스 F	6.56E+05	609	낙타 알파코로나바이러스	2.48E+05
470	파레코바이러스 D	3.10E+05	610	크리파바이러스	1.95E+05
471	파레코바이러스 C	1.24E+06	611	인간 콕사키바이러스 A2	7.75E+07
472	르준간 바이러스 87-012	2.00E+06	612	코로나바이러스 AcCoV-JC34	1.82E+05
473	파레코바이러스 A2	1.80E+07	613	담 피코르나바이러스 3	9.13E+04
474	파레코바이러스 A3	3.58E+06	614	담 피코르나바이러스 1	1.21E+05
475	파레코바이러스 A8	1.61E+07	615	담 오리바이러스 1	3.16E+05
476	파레코바이러스 A17	1.20E+06	616	담 갈리바이러스 1	1.51E+07
477	포타미피바이러스 A	8.43E+05	617	담 칼시바이러스	1.28E+05
478	포타미피바이러스 B	7.20E+05	618	잉어 피코르나바이러스 1	1.13E+05
479	베이하이 콘거 피코르나바이러스	1.15E+06	619	매 피코르나바이러스	3.08E+06
480	돼지 사펠로바이러스 JD2011	N/A	620	말 비염 B 바이러스 1	1.01E+05

[1891]

서열번호	IRES	발현	서열번호	IRES	발현
481	돼지 사펠로바이러스 A2	4.34E+06	621	말 비염 A 바이러스	3.73E+05
482	유인원 사펠로바이러스 1	6.55E+07	622	말 아르테리티스 바이러스	1.89E+05
483	유인원 사펠로바이러스 2	4.24E+07	623	엔테로바이러스 종 단리물 CPML	6.83E+07
484	라보바이러스 C	2.49E+06	624	엔테로바이러스 AN12	3.87E+06
485	라보바이러스 A NYC-B10	1.24E+06	625	돌고래 모빌리바이러스	1.22E+05
486	파라보바이러스 C	1.83E+07	626	디안케 바이러스	1.35E+05
487	파라보바이러스 B	7.85E+06	627	게레자 헤파시바이러스	1.38E+05
488	파라보바이러스 A3	2.44E+08	628	그라페빈 연관 나나바이러스-1	1.30E+05
489	펠리피바이러스 127F	8.92E+06	629	염소 토로바이러스	1.19E+05
490	부세피바이러스 A	7.07E+07	630	구제역 바이러스 O 단리물	1.12E+05
491	부세피바이러스 B	1.17E+08	631	고영이 전염성 복막염 바이러스	1.35E+05
492	파코바이러스 Pf-CHK1	5.87E+06	632	파밍턴 바이러스	1.22E+05
493	HRVC3 QPM	1.64E+07	633	조류 전염성 기관지염 바이러스	2.84E+05
494	HRVB27	2.04E+08	634	인간 리노바이러스 1	7.40E+07
495	HRVA73	1.08E+08	635	EV22	1.95E+07
496	EV L	6.49E+07	636	인간 TMEV-유사 카디오바이러스	4.48E+07
497	EV K	7.52E+07	637	인간 코로나바이러스 229E	N/A
498	EV J 1631	9.88E+07	638	후베이 자오바이러스-유사 바이러스 1	1.03E+05
499	EV J N125	2.90E+07	639	후베이 톰버스-유사 바이러스 9	9.28E+04
500	EV I	1.31E+08	640	후베이 톰버스-유사 바이러스 32	9.23E+04
501	EV F1 BEV 261	1.12E+07	641	후베이 소베모-유사 바이러스 3	1.17E+05
502	EV D94	9.25E+07	642	후베이 피코르나-유사 바이러스 2	1.95E+05
503	PV3	1.25E+08	643	헤파시바이러스 P	6.04E+05
504	EV C102	8.85E+07	644	해리어 피코르나바이러스 1	1.47E+05

[1892]

서열번호	IRES	발현	서열번호	IRES	발현
505	EV 30	5.48E+06	645	푼사기시바이러스 1	4.15E+05
506	SA5	1.61E+08	646	카고시마-2-24-KoV	9.30E+07
507	EV A114	1.50E+08	647	캐시미어 비 바이러스	1.65E+05
508	모보바이러스 A	3.44E+06	648	징먼 피코르나-유사 바이러스	9.32E+04
509	부르펜가리 바이러스	1.09E+07	649	멈프스 바이러스	1.47E+05
510	허니바이러스 A1	1.61E+06	650	마우스 모사바이러스	9.00E+04
511	허니바이러스 A2	6.38E+06	651	미니옵테루스 슈라이버시이 피코르나바이러스 1	6.05E+06
512	Ia Io	1.35E+06	652	린다 바이러스	7.37E+05
513	타우라 증후군 바이러스	8.30E+05	653	레사바이러스 2	3.67E+07
514	ABPV	6.48E+05	654	레사바이러스 1	6.37E+06
515	BRAV-2	3.98E+06	655	포피바이러스 균주 뉴잉글랜드	1.06E+05
516	BRBV-1	3.34E+06	656	페스티바이러스 균주 아이딘	3.11E+06
517	ERAV-1 U188	N/A	657	메추라기 피코르나바이러스 QPV1	6.55E+07
518	GFTV	1.23E+06	658	돼지 사펠로바이러스 1	N/A
519	SAFV V13C	9.32E+07	659	돼지 생식기 호흡기 증후군 바이러스 2	1.29E+05
520	SAV P-113	4.37E+07	660	돼지 엔테로바이러스 9	3.20E+07
521	VHEV	1.74E+08	661	비둘기 피코르나바이러스 B	1.24E+05
522	TRV NGS910	3.84E+07	662	피코르나바이러스 HK21	4.09E+05
523	EMCV2 RD1338	1.97E+06	663	피코르나비랄레스 토토리-HG1	9.54E+04
524	EMCV1 JZ1203	N/A	664	설치류 헤파토바이러스	1.39E+05
525	EMCV1 AnrB-3741	2.55E+06	665	린더케스트 바이러스	4.26E+05
526	코사바이러스 D1	2.11E+06	666	라보바이러스 A	2.88E+06
527	코사바이러스 B1	1.91E+06	667	싱글백 니도바이러스 1	2.62E+05
528	코사바이러스 A SH1	2.16E+06	668	세네카 밸리 바이러스	1.46E+07
529	말라가시바이러스 B	5.05E+06	669	스클레로티니아 스클레로티오럼 dsRNA 마이코바이러스-L	1.69E+05

[1893]

서열번호	IRES	발현	서열번호	IRES	발현
530	모사바이러스 A2 SZAL6	8.27E+06	670	야크 엔테로바이러스	6.19E+06
531	SVV	1.06E+06	671	워블리 포섬 질병 바이러스	2.60E+05
532	PTV A	7.29E+05	672	조류 오쏘레오바이러스 분절 S1	4.37E+05
533	PTV B	6.02E+06	673	염소 코브바이러스 d10	2.20E+08
534	토트오리바이러스	2.76E+07	674	염소 코브바이러스 d20	2.00E+08
535	포사바이러스 1	1.55E+06	675	염소 코브바이러스 d30	1.87E+08
536	A105-675	2.18E+07	676	염소 코브바이러스 d40	2.15E+08
537	A110-675	1.24E+08	677	염소 코브바이러스 d50	9.65E+07
538	18-675	6.04E+07	678	피코르나비랄레스 종 단리물 RtMruf-PicoV	2.26E+08
539	A115-675	5.93E+07	679	아포데루스 아그라리우스피코르나바이러스 균주 봉루안-Aa118	1.90E+08
540	A73-675	1.30E+08	680	니비벤터 콘푸시아누스피코르나바이러스	6.10E+07
541	코브바이러스 16317	2.03E+07	681	박쥐 피코르나바이러스 단리물 BtRs-PicoV	1.13E+06
542	아이치바이러스 Chshc7	1.87E+07	682	관박쥐 피코르나바이러스 균주 구이저우-Rr100	N/A
543	아이치바이러스 고이아니아	1.66E+07	683	관박쥐 피코르나바이러스 균주 헤난-Rf265	3.85E+05
544	아이치바이러스 ETHP4	1.78E+07	684	인간 엔테로바이러스 C105	5.49E+05
545	아이치바이러스 DVI2169	2.98E+06	685	인간 폴리오바이러스 1 균주 NIE1116623	3.94E+05
546	아이치바이러스 DVI2321	6.63E+07	686	인간 엔테로바이러스 109	4.92E+05
547	아이치바이러스 rat08	3.51E+07	687	인간 폴리오바이러스 2 균주 NIE0811460	2.59E+07
548	아이치바이러스 Rt386	5.71E+07	688	소 피코르나바이러스	3.82E+06
549	노르웨이 래트 페스티바이러스	N/A	689	인간 폴리오바이러스 1 균주 EQG1419328	2.44E+05
550	돼지 코브바이러스 GS2	44200000	690	인간 폴리오바이러스 2 단리물 IS_061	5.84E+06
551	코브바이러스 SZAL6	98850000	691	콕사키바이러스 B5	N/A
552	코브바이러스 쉽 TB3	N/A	692	콕사키바이러스 A10	N/A

[1894]

[1895] **실시예 9**

[1896] *주카트 세포에서 IRES에 의한 발현 및 기능성 안정성.*

[1897] 아나베나 인트론/엑손 영역, 가우시아 루시퍼라제 발현 서열 및 이전에 시험된 IRES의 하위세트를 포함하는 2세트의 작제물을 원형화하였다. 60,000개의 주카트 세포에 각각의 원형화 반응 1㎍을 전기천공하였다. 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 전기천공 24시간 후에 측정하였다. 세트 간 비교 및 이전에 정의된 IRES 효능과의 비교를 위해 CVB3 IRES 작제물을 두 세트에 모두에 포함시켰다. CVB1 및 살리바이러스 A SZ1 IRES 작제물은 24시간에 가장 노은 발현을 생성시켰다. 데이터는 도 4A 및 4B에서 찾을 수 있다.

[1898] 전기천공된 주카트 세포의 각각의 라운드에서 IRES 작제물의 기능적 안정성을 3일에 걸쳐 측정하였다. 60,000개 세포에 1㎍의 각각의 원형화 반응을 전기천공시키고, 그 다음 완전 배지 교체 후 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 24시간마다 측정하였다(도 5A 및 도 5B).

[1899] 살리바이러스 A SZ1 및 살리바이러스 A BN2 IRES 작제물은 다른 작제물에 비해 높은 기능적 안정성을 가졌다.

[1900] **실시예 10**

[1901] *주카트 세포에서 원형 RNA 및 선형 RNA의 발현, 기능적 안정성 및 사이토카인 방출.*

[1902] 아나베나 인트론/엑손 영역, 가우시아 루시퍼라제 발현 서열 및 살리바이러스 FHB IRES를 포함하는 작제물을 원형화하였다. 가우시아 루시퍼라제 발현 서열 및 약 150nt polyA 테일을 포함하고, 100%의 우리딘을 5-메톡시 우리딘(5moU)으로 대체하도록 변형된 mRNA를 상업적으로 입수하며, 트라이링크사로부터 구입하였다. 5moU 뉴클레

오타이드 변형은 mRNA 안정성 및 발현을 개선시키는 것으로 밝혀져 있다(문헌[Bioconjug Chem. 2016 Mar 16;27(3):849-53]). 주카트 세포에서 변형된 mRNA, 원형화 반응(순수하지 않음) 및 크기 배제 HPLC에 의해 정제된 circRNA(순수함)의 발현을 측정하고, 비교하였다(도 6A). 60,000개 세포에 1㎍의 각각의 RNA 종을 전기천공한 지 24시간 후에 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 측정하였다.

[1903] 60,000개 세포에 1ug의 각각의 RNA 종을 전기천공하고, 그 다음 완전 배지 교체 후 24시간마다 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 측정하였다. 3일 동안 주카트 세포에서 변형된 mRNA 및 circRNA의 기능적 안정성 데이터의 비교를 도 6B에 도시한다.

[1904] 60,000개 주카트 세포에 상기에 기재된 각각의 RNA 종 및 3p-hpRNA(공지된 RIG-I 효능제인 5' 트라이포스페이트 헤어핀 RNA) 1㎍를 전기천공한 지 18시간 후에 IFNɣ(도 7A), IL-6(도 7B), IL-2(도 7C), RIG-I(도 7D), IFN-β1(도 7E), 및 TNFα(도 7F) 전사 유도를 측정하였다.

[1905] **실시예 11**

[1906] *단핵구 및 대식세포에서 원형 RNA 및 선형 RNA의 발현.*

[1907] 아나베나 인트론/엑손 영역, 가우시아 루시퍼라제 발현 서열 및 살리바이러스 FHB IRES를 포함하는 작제물을 원형화하였다. 가우시아 루시퍼라제 발현 서열 및 약 150nt polyA 테일을 포함하고, 100%의 우리딘을 5-메톡시 우리딘(5moU)으로 대체하도록 변형된 mRNA를 트라이링크사로부터 구입하였다. 인간 1차 단핵구(도 8A) 및 인간 1차 대식세포(도 8B)에서 원형 mRNA 및 변형된 mRNA의 발현을 측정하였다. 60,000개 세포에 1㎍의 각각의 RNA 종을 전기천공한 지 24시간 후에 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 측정하였다. 24시간마다 배지를 변경하고 인간 1차 대식세포의 전기천공 4일 후에 발광을 또한 측정하였다(도 8C). 결과는 도 8에서 찾을 수 있다. 발광의 차이는 각 경우에 통계학적으로 유의하였다($p < 0.05$).

[1908] **실시예 12**

[1909] *1차 T 세포에서 IRES에 의한 발현 및 기능성 안정성.*

[1910] 아나베나 인트론/엑손 영역, 가우시아 루시퍼라제 발현 서열 및 이전에 시험된 IRES의 하위세트를 포함하는 작제물을 원형화하고, 반응 생성물을 크기 배제 HPLC로 정제시켰다. 150,000개의 1차 인간 CD3+ T 세포에 각각의 circRNA 1㎍을 전기천공하였다. 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 전기천공 24시간 후에 측정하였다(**도 9A**). 아이치 바이러스 및 CVB3 IRES 작제물은 24시간에 가장 많은 발현을 가졌다.

[1911] 각각의 작제물의 기능적 안정성을 비교하기 위해서 3일 동안 전기천공 후 24시간마다 발광을 또한 측정하였다(도 9B). 살리바이러스 A SZ1 IRES를 사용한 작제물이 가장 안정적이었다.

[1912] **실시예 13**

[1913] *1차 T 세포 및 PBMC에서 원형 RNA 및 선형 RNA의 발현 및 기능적 안정성.*

[1914] 아나베나 인트론/엑손 영역, 가우시아 루시퍼라제 발현 서열 및 살리바이러스 A SZ1 IRES 또는 살리바이러스 FHB IRES를 포함하는 작제물을 원형화하였다. 가우시아 루시퍼라제 발현 서열 및 약 150nt polyA 테일을 포함하고, 100%의 우리딘을 5-메톡시 우리딘(5moU)으로 대체하도록 변형된 mRNA를 트라이링크사로부터 구입하였다. 살리바이러스 A SZ1 IRES HPLC 정제된 원형 mRNA 및 변형된 mRNA의 발현을 인간 1차 CD3+ T 세포에서 측정하였다. 살리바이러스 FHB HPLC 정제된 원형, 비정제된 원형 및 변형된 mRNA의 발현을 인간 PBMC에서 측정하였다. 150,000개 세포에 1㎍의 각각의 RNA 종을 전기천공한 지 24시간 후에 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 측정하였다. 1차 인간 T 세포에 대한 데이터를 도 10A 및 도 10B에 도시하고, PBMC에 대한 데이터를 도 10C에 도시한다. 정제된 원형 RNA와 비정제된 않은 원형 RNA 또는 선형 RNA 사이의 발현 차이는 각각의 경우에 유의하였다($p<0.05$).

[1915] 작제물 기능적 안정성을 비교하기 위해서 3일에 걸쳐 전기천공 후 24시간마다 1차 T 세포 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 측정하였다. 데이터를 도 10B에 도시한다. 정제된 원형 RNA와 선형 RNA 사이의 제1일 측정으로부터 상대 발광의 차이는 1차 T 세포에 대해 제2일 및 제3일 둘 다에서 유의하였다.

[1916] **실시예 14**

[1917] *아나베나 인트론에서 순열 부위에 의한 원형화 효율.*

[1918] CVB3 IRES, 가우시아 루시퍼라제 발현 서열, 아나베나 인트론/엑손 영역, 스페이서, 내부 상동성 영역 및 상동

성 아암을 포함하는 RNA 작제물을 생산하였다. 전통적인 아나베나 인트론 순열 부위 및 P9 내의 5개의 연속 순열 부위를 사용하는 작제물의 원형화 효율을 HPLC로 측정하였다. P9 내의 5개 연속 순열 부위에 대한 HPLC 크로마토그램을 도 11A에 도시한다.

[1919] 다양한 순열 부위에서 원형화 효율을 측정하였다. 원형화 효율은 circRNA/(circRNA + 전구체 RNA) 각각에 대한 HPLC 크로마토그램 곡선 하 면적으로 정의된다. 각각의 순열 부위에서 원형화 효율의 순위 정량화를 도 11B에 도시한다. 3개의 순열 부위(도 11B에 제시됨)를 추가 조사를 위해 선택하였다.

[1920] 본 실시예에서 원형 RNA를 시험관내 전사(IVT)에 의해 원형화하고, 그 다음 스핀 칼럼을 통해 정제시켰다. Mg^{2+} 및 구아노신 뉴클레오타이드와의 추가 인큐베이션 단계가 포함되면 모든 작제물에 대한 원형화 효율이 더 높을 가능성이 있을 것이지만; 이 단계의 제거는 원형 RNA 작제물 간의 비교 및 이의 최적화를 가능하게 하였다. 이러한 최적화 수준은 키메라 항원 수용체를 암호화하는 것과 같은 큰 RNA 작제물로 높은 원형화 효율을 유지하는 데 특히 유용하다.

[1921] **실시예 15**

[1922] *대안 인트론의 원형화 효율.*

[1923] 가변 종 기원의 순열 그룹 1 인트론 또는 순열 부위 및 CVB3 IRES, 가우시아 루시퍼라제 발현 서열, 스페이서, 내부 상동성 영역 및 상동성 아암을 포함하는 몇몇 불변 요소를 함유하는 전구체 RNA를 생산하였다. 원형화 데이터는 도 12에서 찾을 수 있다. 도 12A는 전구체, CircRNA 및 인트론을 분할하는 크로마토그램을 도시한다. 도 12B는 인트론 작제물의 함수로서, 도 12A에 도시된 크로마토그램에 기초한 원형화 효율의 순위 정량화를 제공한다.

[1924] 본 실시예에서 원형 RNA를 시험관내 전사(IVT)에 의해 원형화하고, 그 다음 스핀 칼럼 정제시켰다. Mg^{2+} 및 구아노신 뉴클레오타이드와의 추가 인큐베이션 단계가 포함되면 모든 작제물에 대한 원형화 효율이 더 높을 가능성이 있을 것이지만; 이 단계의 제거는 원형 RNA 작제물 간의 비교 및 이의 최적화를 가능하게 하였다. 이러한 최적화 수준은 키메라 항원 수용체를 암호화하는 것과 같은 큰 RNA 작제물로 높은 원형화 효율을 유지하는 데 특히 유용하다.

[1925] **실시예 16**

[1926] *상동성 아암 존재 또는 길이에 의한 원형화 효율.*

[1927] CVB3 IRES, 가우시아 루시퍼라제 발현 서열, 아나베나 인트론/엑손 영역, 스페이서 및 내부 상동성 영역을 포함하는 RNA 작제물을 생산하였다. 30nt, 25% GC 상동성 아암을 갖거나 상동성 아암을 갖지 않는("NA") 3개의 아나베나 인트론 순열 부위를 나타내는 작제물을 시험하였다. 이들 작제물은 Mg^{2+} 인큐베이션 단계 없이 원형화되도록 하였다. 원형화 효율을 측정하고, 비교하였다. 데이터는 도 13A 및 도 13B에서 찾을 수 있다. 원형화 효율은 상동성 아암이 없는 각각의 작제물에 대해서 더 높았다. 도 13A는 원형화 효율의 순위 정량화를 제공하고; 도 13B는 전구체, circRNA 및 인트론을 분할하는 크로마토그램을 제공한다.

[1928] 3개의 순열 부위 각각에 대해, 10nt, 20nt 및 30nt의 아암 길이 및 25%, 50% 및 75%의 GC 함량으로 작제물을 생성시켰다. 이들 작제물의 스플라이싱 효율을 측정하고, 상동성 아암이 없는 작제물과 비교하였다(도 14). 스플라이싱 효율은 스플라이싱 반응에서 총 RNA에 대한 유리 인트론의 비율로 정의된다.

[1929] 도 15A(좌측)는 개선된 스플라이싱 효율에 대한 강한 상동성 아암의 기여를 나타내는 HPLC 크로마토그램을 도시한다. 상단 좌측: 75% GC 함량, 10nt 상동성 아암. 중앙 좌측: 75% GC 함량, 20nt 상동성 아암. 하단 좌측: 75% GC 함량, 30nt 상동성 아암.

[1930] 도 15A(우측)는 circRNA 피크에서 숄더로 나타나는 증가된 니킹과 쌍을 이루는 증가된 스플라이싱 효율을 나타낸 HPLC 크로마토그램을 나타낸다. 상단 우측: 75% GC 함량, 10nt 상동성 아암. 중앙 우측: 75% GC 함량, 20nt 상동성 아암. 하단 우측: 75% GC 함량, 30nt 상동성 아암.

[1931] 도 15B(좌측)는 개선된 원형화 효율을 입증하기 위해 가정된 순열 부위 및 상동성 아암의 선택된 조합을 도시한다.

[1932] 도 15B(우측)는 이. 콜라이(E. coli) polyA 중합효소로 처리된, 개선된 원형화 효율을 입증하기 위해 가정된 순열 부위 및 상동성 아암의 선택된 조합을 도시한다.

[1933] 본 실시예에서 원형 RNA를 시험관내 전사(IVT)에 의해 원형화하고, 그 다음 스핀 칼럼 정제시켰다. 구아노신 뉴클레오타이드와의 추가적인 Mg^{2+} 인큐베이션 단계가 포함되면 모든 작제물에 대한 원형화 효율이 더 높을 가능성이 있을 것이지만; 이 단계의 제거는 원형 RNA 작제물 간의 비교 및 이의 최적화를 가능하게 하였다. 이러한 최적화 수준은 키메라 항원 수용체를 암호화하는 것과 같은 큰 RNA 작제물로 높은 원형화 효율을 유지하는 데 특히 유용하다.

[1934] **실시예 17**

[1935] 키메라 항원 수용체를 암호화하는 *원형 RNA*

[1936] 아나베나 인트론/엑손 영역, Kymriah 키메라 항원 수용체(CAR) 발현 서열 및 CVB3 IRES를 포함하는 작제물을 원형화하였다. 100,000개의 인간 1차 CD3+ T 세포에 500ng의 circRNA를 전기천공시키고, GFP 및 파이어플라이 루시퍼라제를 안정적으로 발현하는 라지 세포와 24시간 동안 공배양하였다. 효과기 대 표적 비율(E:T 비율) 0.75:1. 100,000개의 인간 1차 CD3+ T 세포를 모의 전기천공시키고, 대조군으로서 공배양하였다(도 16).

[1937] 100,000개의 인간 1차 CD3+ T 세포의 세트를 모의 전기천공시키거나 1㎍의 circRNA로 전기천공시킨 다음, GFP 및 파이어플라이 루시퍼라제 E:T 비율 10:1을 안정적으로 발현하는 라지 세포와 48시간 동안 공배양하였다(도 17).

[1938] 라지 표적 세포의 특이적 용해의 정량화를 파이어플라이 발광의 검출에 의해 결정하였다(도 18). 모의 전기천공되거나 또는 상이한 CAR 서열을 암호화하는 circRNA로 전기천공된 100,000개의 인간 1차 CD3+ T 세포를 GFP 및 파이어플라이 루시퍼라제를 안정적으로 발현하는 라지 세포와 함께 48시간 동안 공배양하였다. 특이적 용해%를 1-[CAR 조건 발광]/[모의 조건 발광]으로 정의하였다. E:T 비율 10:1.

[1939] **실시예 18**

[1940] *주카트 세포 및 휴지기 인간 T 세포에서 원형 RNA 및 선형 RNA의 발현 및 기능적 안정성.*

[1941] 아나베나 인트론/엑손 영역, 가우시아 루시퍼라제 발현 서열 및 이전에 시험된 IRES의 하위세트를 포함하는 작제물을 원형화하고, 반응 생성물을 크기 배제 HPLC로 정제시켰다. 150,000개의 주카트 세포에 1㎍의 원형 RNA 또는 5moU-mRNA를 전기천공하였다. 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 전기천공 24시간 후에 측정하였다(도 19A 좌측). 150,000개의 휴지기 1차 인간 CD3+ T 세포(자극 후 10일)에 1㎍의 원형 RNA 또는 5moU-mRNA를 전기천공하였다. 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 전기천공 24시간 후에 측정하였다(도 19A 우측).

[1942] 전기천공하고, 그 다음 완전 배지 교체 후 24시간마다 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 측정하였다. 기능적 안정성 데이터를 도 19B에 도시한다. 원형 RNA는 각각의 경우에 선형 RNA보다 더 높은 기능적 안정성을 가졌고, 주카트 세포에서 더 확연한 차이가 있었다.

[1943] **실시예 19**

[1944] *선형 RNA 또는 다양한 원형 RNA 작제물로 전기천공된 세포의 IFN-β1, RIG-I, IL-2, IL-6, IFNɣ 및 TNFα 전사체 유도.*

[1945] 아나베나 인트론/엑손 영역, 가우시아 루시퍼라제 발현 서열 및 이전에 시험된 IRES의 하위세트를 포함하는 작제물을 원형화하고, 반응 생성물을 크기 배제 HPLC로 정제시켰다. 150,000개의 CD3+ 인간 T 세포에 1㎍의 원형 RNA, 5moU-mRNA 또는 면역자극성 양성 대조군 폴리 이노신:사이토신을 전기천공하였다. IFN-β1(도 20A), RIG-I(도 20B), IL-2(도 20C), IL-6(도 20D), IFNɣ(도 20E) 및 TNFα(도 20F) 전사체 유도를 전기천공 18시간 후에 측정하였다.

[1946] **실시예 20**

[1947] *상이한 양의 원형 RNA 또는 선형 RNA로 전기천공된 CAR 발현 세포에 의한 표적 세포의 특이적 용해 및 IFNɣ 전사체 유도; 상이한 E:T 비율에서 CAR 발현 세포에 의한 표적 및 비-표적 세포의 특이적 용해.*

[1948] 아나베나 인트론/엑손 영역, 항-CD19 CAR 발현 서열 및 CVB3 IRES를 포함하는 작제물을 원형화하고, 반응 생성물을 크기 배제 HPLC로 정제시켰다. 모의 전기천공되거나 또는 항-CD19 CAR 서열을 암호화하는 다른 양의 circRNA로 전기천공된 150,000개의 인간 1차 CD3+ T 세포를 2:1의 E:T 비율에서 GFP 및 파이어플라이 루시퍼라제를 안정적으로 발현하는 라지 세포와 함께 12시간 동안 공배양하였다. 라지 표적 세포의 특이적 용해를 파이

어플라이 발광의 검출에 의해 결정하였다(도 21A). 특이적 용해 %를 1-[CAR 조건 발광]/[모의 조건 발광]으로 정의하였다. IFNɣ 전사체 유도를 전기천공 24시간 후에 측정하였다(도 21B).

[1949] 150,000개의 인간 1차 CD3+ T 세포를 모의 전기천공시키거나 또는 항-CD19 CAR 서열을 암호화하는 500ng circRNA 또는 m1ψ-mRNA로 전기천공시킨 다음, 상이한 E:T 비율에서 파이어플라이 루시퍼라제를 안정적으로 발현하는 라지 세포와 함께 24시간 동안 공배양하였다. 라지 표적 세포의 특이적 용해 %를 파이어플라이 발광의 검출에 의해 결정하였다(도 22A). 특이적 용해%를 1-[CAR 조건 발광]/[모의 조건 발광]으로 정의하였다.

[1950] CAR 발현 T 세포를 또한 상이한 E:T 비율에서 파이어플라이 루시퍼라제를 안정적으로 발현하는 라지 또는 K562 세포와 함께 24시간 동안 공배양하였다. 라지 표적 세포 또는 K562 비-표적 세포의 특이적 용해를 파이어플라이 발광의 검출에 의해 결정하였다(도 22B). 특이적 용해%를 1-[CAR 조건 발광]/[모의 조건 발광]으로 정의한다.

[1951] **실시예 21**

[1952] *CAR을 암호화하는 원형 RNA 또는 선형 RNA로 전기천공된 T 세포에 의한 표적 세포의 특이적 용해.*

[1953] 아나베나 인트론/엑손 영역, 항-CD19 CAR 발현 서열 및 CVB3 IRES를 포함하는 작제물을 원형화하고, 반응 생성물을 크기 배제 HPLC로 정제시켰다. 인간 1차 CD3+ T 세포에 각각 CD19-표적화 CAR을 암호화하는 500ng의 원형 RNA 또는 등몰량의 m1ψ-mRNA를 전기천공하였다. 라지 세포를 10:1의 E:T 비율로 7일에 걸쳐 CAR-T 세포 배양물에 첨가하였다. 특이적 용해%를 제1일, 제3일, 제5 및 제7일에 두 작제물 모두에 대해서 측정하였다(도 23).

[1954] **실시예 22**

[1955] *항-CD19 CAR 또는 항-BCMA CAR을 발현하는 T 세포에 의한 라지 세포의 특이적 용해.*

[1956] 아나베나 인트론/엑손 영역, 항-CD19 또는 항-BCMA CAR 발현 서열 및 CVB3 IRES를 포함하는 작제물을 원형화하고, 반응 생성물을 크기 배제 HPLC로 정제시켰다. 150,000개의 1차 인간 CD3+ T 세포에 500ng의 circRNA를 전기천공시킨 다음, 2:1의 E:T 비율로 라지 세포와 공배양하였다. 전기천공 12시간 후 특이적 용해%를 측정하였다(도 24).

[1957] **실시예 23**

[1958] ***실시예 23A:*** *화합물의 합성*

[1959] 본 발명의 대표적인 이온화 가능한 지질의 합성은 PCT 출원 제 PCT/US2016/052352호, 제PCT/US2016/068300호, 제PCT/US2010/061058호, 제PCT/US2018/058555호, 제PCT/US2018/053569호, 제PCT/US2017/028981호, 제PCT/US2019/025246호, 제PCT/US2018/035419호, 제PCT/US2019/015913호 및 미국 출원 공보 제20190314524호, 제20190321489호 및 제20190314284호에 기재되어 있고, 이들 각각의 내용은 전문이 참조에 의해 본 명세서에 포함된다.

[1960] ***실시예 23B:*** *화합물의 합성*

[1961] 본 발명의 대표적인 이온화 가능한 지질의 합성은 내용 전문이 참조에 의해 본 명세서에 포함된 미국 출원 공보 제US20170210697A1호에 기재되어 있다.

[1962] **실시예 24**

[1963] *기관에 의한 단백질 발현*

[1964] FLuc를 암호화하는 원형 또는 선형 RNA를 생성시키고, 하기 제형을 갖는 전달 비히클에 로딩하였다: 50%의 표 10b의 이온화 가능한 지질 15, 10%의 DSPC, 1.5%의 PEG-DMG, 38.5%의 콜레스테롤. CD-1 마우스에게 0.2mg/kg을 투여하고, 6시간(라이브 IVIS) 및 24시간(라이브 IVIS 및 생체외 IVIS)에 발광을 측정하였다. 간, 비장, 신장, 폐 및 심장의 총 플럭스(관심 영역에 걸친 광자/초)를 측정하였다(도 25 및 도 26).

[1965] **실시예 25**

[1966] *비장에서의 발현의 분포*

[1967] GFP를 암호화하는 원형 또는 선형 RNA를 생성시키고, 하기 제형을 갖는 전달 비히클에 로딩한다: 50%의 표 10b의 이온화 가능한 지질 15, 10%의 DSPC, 1.5%의 PEG-DMG, 38.5%의 콜레스테롤. 제형을 CD-1 마우스에게 투여한다. 유세포 분석법을 비장 세포에서 실시하여 세포 유형에 따른 발현의 분포를 결정한다.

[1968] **실시예 26**

[1969] *나노입자 조성물의 생성*

[1970] 원형 RNA를 세포로 전달하는데 사용하기 위한 안전하고, 효과적인 나노입자 조성물을 조사하기 위해, 다양한 제형을 제조하고, 시험한다. 구체적으로, 나노입자 조성물의 지질 성분에서 특정 요소 및 이들의 비율을 최적화한다.

[1971] 나노입자는 1종 유체 스트림으로 또는 2개의 유체 스트림의 미세유체 및 T-접합 혼합과 같은 혼합 공정으로 제조될 수 있는데, 유체 스트림 중 하나가 원형 RNA를 함유하고 다른 하나는 지질 성분을 갖는다.

[1972] 지질 조성물은 이온화 가능한 지질, 선택적으로 헬퍼 지질(예컨대, DOPE, DSPC 또는 올레산, 미국 앨라배마주 알라바스터 소재의 아반티 폴라 리피즈사(Avanti Polar Lipids)로부터 입수 가능함), PEG 지질(예컨대, PEG-DMG라고도 알려져 있는 1,2-다이미리스토일-sn-글리세롤 메톡시폴리에틸렌 글리콜, 미국 앨라배마주 알라바스터 소재의 아반티 폴라 리피즈사로부터 입수 가능함) 및 용매, 예를 들어, 에탄올 중의 약, 예를 들어, 40 또는 50mM 농도의 콜레스테롤과 같은 구조 지질을 조합함으로써 제조된다. 용액은 예를 들어, -20℃에서 보관하기 위해 냉동되어야 한다. 지질을 조합하여 목적하는 몰비(예를 들어, 하기 표 31a 및 표 31b 참조)를 생성하고, 물 및 에탄올, 예를 들어, 약 5.5mM 내지 약 25mM의 최종 지질 농도로 희석시킨다.

[1973] [표 31a]

제형 번호	설명
1	C12-200, DOPE, Chol 및 DMG-PEG2K(40:30:25:5)의 에탄올 용액의 분취물 50mg/mL를 혼합하고, 3mL 최종 부피로 에탄올로 희석시킨다. 별개로, circRNA의 수성 완충 용액(10mM 시트르산염/150mM NaCl, pH 4.5)을 1mg/mL 스톡으로부터 제조한다. 지질 용액을 수성 circRNA 용액에 신속하고 주입하고, 진탕하여 20% 에탄올 중의 최종 현탁액을 산출한다. 생성된 나노입자 현탁액을 여과하고, 1×PBS(pH 7.4)로 정용여과하고, 농축하고, 2 내지 8° C에서 저장한다.
2	DODAP, DOPE, 콜레스테롤 및 DMG-PEG2K(18:56:20:6)의 에탄올 용액의 분취물 50mg/mL를 혼합하고, 3mL 최종 부피로 에탄올로 희석시킨다. 별개로, EPO circRNA의 수성 완충 용액(10mM 시트르산염/150mM NaCl, pH 4.5)을 1mg/mL 스톡으로부터 제조한다. 지질 용액을 수성 circRNA 용액에 신속하고 주입하고, 진탕하여 20% 에탄올 중의 최종 현탁액을 산출한다. 생성된 나노입자 현탁액을 여과하고, 1×PBS(pH 7.4)로 정용여과하고, 농축하고, 2 내지 8° C에서 저장한다. 최종 농도=1.35mg/mL EPO circRNA(캡슐화됨), Zave=75.9nm (Dv(50)=57.3nm; Dv(90)=92.1nm).
3	HGT4003, DOPE, 콜레스테롤 및 DMG-PEG2K(50:25:20:5)의 에탄올 용액의 분취물 50mg/mL를 혼합하고, 3mL 최종 부피로 에탄올로 희석시킨다. 별개로, circRNA의 수성 완충 용액(10mM 시트르산염/150mM NaCl, pH 4.5)을 1mg/mL 스톡으로부터 제조한다. 지질 용액을 수성 circRNA 용액에 신속하고 주입하고, 진탕하여 20% 에탄올 중의 최종 현탁액을 산출한다. 생성된 나노입자 현탁액을 여과하고, 1×PBS(pH 7.4)로 정용여과하고, 농축하고, 2 내지 8° C에서 저장한다.
4	ICE, DOPE 및 DMG-PEG2K(70:25:5)의 에탄올 용액의 분취물 50mg/mL를 혼합하고, 3mL 최종 부피로 에탄올로 희석시킨다. 별개로, circRNA의 수성 완충 용액(10mM 시트르산염/150mM NaCl, pH 4.5)을 1mg/mL 스톡으로부터 제조한다. 지질 용액을 수성 circRNA 용액에 신속하고 주입하고, 진탕하여 20% 에탄올 중의 최종 현탁액을 산출한다. 생성된 나노입자 현탁액을 여과하고, 1×PBS(pH 7.4)로 정용여과하고, 농축하고, 2 내지 8° C에서 저장한다.
5	HGT5000, DOPE, 콜레스테롤 및 DMG-PEG2K(40:20:35:5)의 에탄올 용액의 분취물 50mg/mL를 혼합하고, 3mL 최종 부피로 에탄올로 희석시킨다. 별개로, EPO circRNA의 수성 완충 용액(10mM 시트르산염/150mM NaCl, pH 4.5)을 1mg/mL 스톡으로부터 제조한다. 지질 용액을 수성 circRNA 용액에 신속하고 주입하고, 진탕하여

[1974]

	20% 에탄올 중의 최종 현탁액을 산출한다. 생성된 나노입자 현탁액을 여과하고, 1×PBS(pH 7.4)로 정용여과하고, 농축하고, 2 내지 8° C에서 저장한다. 최종 농도=1.82mg/mL EPO mRNA(캡슐화됨). Zave=105.6nm(Dv(50)=53.7nm; Dv(90)=157nm).
6	HGT5001, DOPE, 콜레스테롤 및 DMG-PEG2K(40:20:35:5)의 에탄올 용액의 분취물 50mg/mL를 혼합하고, 3mL 최종 부피로 에탄올로 희석시킨다. 별개로, EPO circRNA의 수성 완충 용액(10mM 시트르산염/150mM NaCl, pH 4.5)을 1mg/mL 스톡으로부터 제조한다. 지질 용액을 수성 circRNA 용액에 신속하고 주입하고, 진탕하여 20% 에탄올 중의 최종 현탁액을 산출한다. 생성된 나노입자 현탁액을 여과하고, 1×PBS(pH 7.4)로 정용여과하고, 농축하고, 2 내지 8° C에서 저장한다.

[1975]

[1976] 일부 실시형태에서, 전달 비히클은 표 31a에 기재된 바와 같은 제형을 갖는다.

[1977] [표 31b]

[1978]

조성(몰%)	성분
40:20:38.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
45:15:38.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
50:10:38.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
55:5:38.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG

[1979]

조성(몰%)	성분
60:5:33.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
45:20:33.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
50:20:28.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
55:20:23.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
60:20:18.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
40:15:43.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
50:15:33.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
55:15:28.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
60:15:23.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
40:10:48.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
45:10:43.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
55:10:33.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
60:10:28.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
40:5:53.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
45:5:48.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
50:5:43.5:1.5	화합물:인지질:피토스테롤*:PEG-DMG
40:20:40:0	화합물:인지질:피토스테롤*:PEG-DMG
45:20:35:0	화합물:인지질:피토스테롤*:PEG-DMG
50:20:30:0	화합물:인지질:피토스테롤*:PEG-DMG
55:20:25:0	화합물:인지질:피토스테롤*:PEG-DMG
60:20:20:0	화합물:인지질:피토스테롤*:PEG-DMG
40:15:45:0	화합물:인지질:피토스테롤*:PEG-DMG

[1980] 일부 실시형태에서, 전달 비히클은 표 31b에 기재된 바와 같은 제형을 갖는다.

[1981] circRNA를 포함하는 나노입자 조성물의 경우, 탈이온수 중의 0.1mg/㎖ 농도의 circRNA 용액을 완충액, 예를 들어, pH 3 내지 4에서 50mM 시트르산나트륨 완충액으로 희석시켜 스톡 용액을 형성한다. 대안적으로, 탈이온수 중의 0.15mg/㎖ 농도의 circRNA 용액을 완충액, 예를 들어, pH 3 내지 4.5에서 6.25mM 아세트산나트륨 완충액으로 희석시켜 스톡 용액을 형성한다.

[1982] 원형 RNA 및 지질 성분을 포함하는 나노입자 조성물은 약 5:1 내지 약 50:1의 지질 성분 대 circRNA wt:wt 비율로 원형 RNA를 포함하는 용액과 지질 용액을 조합함으로써 제조된다. 지질 용액을 예를 들어, NanoAssemblr 미세유체 기반 시스템을 사용하여 약 10㎖/분 내지 약 18㎖/분 또는 약 5㎖/분 내지 약 18㎖/분의 유량으로 circRNA 용액에 빠르게 주입하여, 약 1:1 내지 약 4:1의 물 대 에탄올 비율을 갖는 현탁액을 생성시킨다.

[1983] 나노입자 조성물은 에탄올을 제거하고, 완충액 교환을 달성하기 위해 투석에 의해 처리될 수 있다. 10kDa 또는 20kDa의 분자량 컷오프를 갖는 Slide-A-Lyzer 카세트(미국 일리노이주 락포드 소재의 써모 피셔 사이언티픽사(Thermo Fisher Scientific Inc.))를 사용하여 1차 생성물의 부피의 200배 부피로 인산염 완충 식염수(PBS), pH 7.4에 대해 제형을 2회 투석한다. 이어서, 제형을 4℃에서 밤새 투석한다. 생성된 나노입자 현탁액을 0.2㎛ 멸균 필터(독일 늄브레크트 소재의 사르스테트사(Sarstedt)를 통해 유리 바이알로 여과하고, 크림프 마개로 밀봉한다. 0.01mg/㎖ 내지 0.15mg/㎖의 나노입자 조성물 용액을 일반적으로 얻는다.

[1984] 상기에 기재된 방법은 나노-침전 및 입자 형성을 유도한다.

[1985] T-접합 및 직접 주입을 포함하지만 이에 제한되지 않는 대안적인 공정을 사용하여 동일한 나노-침전을 달성할 수 있다. B. 나노입자 조성물의 특징규명

[1986] Zetasizer Nano ZS(영국 워세스터셔 맬버른 소재의 맬버른 인스트루먼츠사(Malvern Instruments Ltd))를 사용하여 입자 크기, 다분산 지수(PDI) 및 입자 크기를 결정하는 데 있어서 1xPBS 그리고 제타 전위를 결정하는 데 있어서 15mM PBS에서 나노입자 조성물의 제타 전위를 결정할 수 있다.

[1987] 자외선-가시광선 분광법을 사용하여 나노입자 조성물에서 circRNA의 농도를 결정할 수 있다. 1xPBS에 희석된 제형 100㎕를 메탄올과 클로로폼의 4:1(v/v) 혼합물 900㎕에 첨가한다. 혼합 후, 용액의 흡광도 스펙트럼을 예를 들어, DU 800 분광 광도계(미국 캘리포니아주 브레아 소재의 베크만 쿨터사(Beckman Coulter, Inc.)의 베크만 쿨터)에서 230nm와 330nm 사이에서 기록한다. 나노입자 조성물 중의 circRNA의 농도를 조성물에 사용된 circRNA의 흡광 계수 및 예를 들어, 260nm의 파장에서의 흡광도와 예를 들어, 330nm의 파장에서의 기준 값 사이의 차이에 기초하여 계산할 수 있다.

[1988] QUANT-IT™ RIBOGREEN® RNA 검정(미국 캘리포니아주 칼스배드 소재의 인비트로젠사(Invitrogen Corporation))을 사용하여 나노입자 조성물에 의한 circRNA의 캡슐화를 평가할 수 있다. 샘플을 TE 완충 용액(10mM Tris-HCl, 1mM EDTA, pH 7.5)에서 대략 5㎍/㎖ 또는 1㎍/㎖의 농도로 희석시킨다. 50㎕의 희석된 샘플을 폴리스티렌 96웰 플레이트로 옮기고, 50㎕의 TE 완충액 또는 50㎕의 2 내지 4% Triton X-100 용액을 웰에 첨가한다. 플레이트를 37℃의 온도에서 15분 동안 인큐베이션시킨다. RIBOGREEN® 시약을 TE 완충액에 1:100 또는 1:200으로 희석시키고, 이 용액 100㎕를 각각의 웰에 첨가한다. 예를 들어, 약 480nm의 여기 파장 및 예를 들어, 약 520nm의 방출 파장에서 형광 플레이트 판독기(Wallac Victor 1420 Multilabel Counter; 미국 매사추세츠주 왈텀 소재의 퍼킨 엘머사(Perkin Elmer))를 사용하여 형광 강도를 측정할 수 있다. 각각의 샘플의 형광 값에서 시약 블랭크의 형광 값을 뺀 다음, 무손상 샘플(Triton X-100을 첨가하지 않은 경우)의 형광 강도를 파괴된 샘플(Triton X-100의 첨가로 인해 발생)의 형광 값으로 나누어 유리 circRNA의 백분율을 결정한다.

[1989] C. *생체내 제형 연구:*

[1990] 다양한 나노입자 조성물이 표적화 세포에 circRNA를 얼마나 효과적으로 전달하는지를 모니터링하기 위해서, circRNA를 포함하는 상이한 나노입자 조성물을 제조하고, 설치류 집단에 투여한다. 마우스에 지질 나노입자 제형을 갖는 나노입자 조성물을 포함하는 단일 용량을 정맥내, 근육내, 동맥내 또는 종양내 투여한다. 일부 경우에, 마우스가 용량을 흡입하도록 만들 수 있다. 용량 크기는 0.001mg/kg에서 10mg/kg 범위일 수 있으며, 여기서 10mg/kg은 마우스의 체질량 1kg마다 나노입자 조성물 중에 10mg의 circRNA를 포함하는 용량을 기재한다. PBS를 포함하는 대조군 조성물을 또한 사용할 수 있다.

[1991] 나노입자 조성물을 마우스에게 투여할 때, 특정 제형 및 이의 용량의 용량 전달 프로파일, 용량 반응 및 독성은 효소-결합 면역흡착 검정(ELISA), 생물발광 영상화 또는 기타 방법에 의해 측정될 수 있다. 단백질 발현의 시간

경과를 또한 평가할 수 있다. 평가를 위해 설치류로부터 수집한 샘플은 혈액 및 조직(예를 들어, 근육 주사 부위로부터의 근육 조직 및 내부 조직)을 포함할 수 있고; 샘플 수집은 동물 희생을 포함할 수 있다.

[1992] circRNA를 포함하는 조성물의 투여에 의해 유도된 더 높은 수준의 단백질 발현은 더 높은 circRNA 번역 및/또는 나노입자 조성물 circRNA 전달 효율을 나타낼 것이다. 비-RNA 성분이 번역 머시너리 자체에 영향을 미치는 것으로 생각되지 않기 때문에, 더 높은 수준의 단백질 발현은 다른 나노입자 조성물 또는 이의 부재에 비해 주어진 나노입자 조성물에 의한 circRNA의 더 높은 전달 효율을 나타낼 가능성이 있다.

[1993] **실시예 27**

[1994] *나노입자 조성물의 특징규명*

[1995] Zetasizer Nano ZS(영국 워세스터셔 맬버른 소재의 맬버른 인스트루먼츠사)를 사용하여 입자 크기, 다분산 지수(PDI) 및 입자 크기를 결정하는 데 있어서 1xPBS 그리고 제타 전위를 결정하는 데 있어서 15mM PBS에서 전달 히비클 조성물의 제타 전위를 결정할 수 있다.

[1996] 자외선-가시광선 분광법은 사용하여 전달 비히클 조성물에서 치료 및/또는 예방(예를 들어, RNA)의 농도를 결정할 수 있다. 1xPBS에 희석된 제형 100㎕를 메탄올과 클로로폼의 4:1(v/v) 혼합물 900㎕에 첨가한다. 혼합 후, 용액의 흡광도 스펙트럼을 예를 들어, DU 800 분광 광도계(미국 캘리포니아주 브레아 소재의 베크만 쿨터사(Beckman Coulter, Inc.)의 베크만 쿨터)에서 230㎚와 330㎚ 사이에서 기록한다. 전달 비히클 조성물에서 치료 및/또는 예방의 농도를 조성물에 사용된 치료 및/또는 예방의 흡광 계수 및 예를 들어, 260㎚의 파장에서의 흡광도와 예를 들어, 330㎚의 파장에서의 기준 값 사이의 차이에 기초하여 계산할 수 있다.

[1997] RNA를 포함하는 전달 비히클 조성물의 경우, QUANT-IT™ RIBOGREEN® RNA 검정(미국 캘리포니아주 칼스배드 소재의 인비트로젠사(Invitrogen Corporation))을 사용하여 전달 비히클에 의한RNA의 캡슐화를 평가할 수 있다. 샘플을 TE 완충 용액(10mM Tris-HCl, 1mM EDTA, pH 7.5)에서 대략 5㎍/㎖ 또는 1㎍/㎖의 농도로 희석시킨다. 50㎕의 희석된 샘플을 폴리스티렌 96웰 플레이트로 옮기고, 50㎕의 TE 완충액 또는 50㎕의 2 내지 4% Triton X-100 용액을 웰에 첨가한다. 플레이트를 37℃의 온도에서 15분 동안 인큐베이션시킨다. RIBOGREEN® 시약을 TE 완충액에 1:100 또는 1:200으로 희석시키고, 이 용액 100㎕를 각각의 웰에 첨가한다. 예를 들어, 약 480㎚의 여기 파장 및 예를 들어, 약 520㎚의 방출 파장에서 형광 플레이트 판독기(Wallac Victor 1420 Multilabel Counter; 미국 매사추세츠주 왈텀 소재의 퍼킨 엘머사(Perkin Elmer))를 사용하여 형광 강도를 측정할 수 있다. 각각의 샘플의 형광 값에서 시약 블랭크의 형광 값을 뺀 다음, 무손상 샘플(Triton X-100을 첨가하지 않은 경우)의 형광 강도를 파괴된 샘플(Triton X-100의 첨가로 인해 발생)의 형광 값으로 나누어 유리 RNA의 백분율을 결정한다.

[1998] **실시예 28**

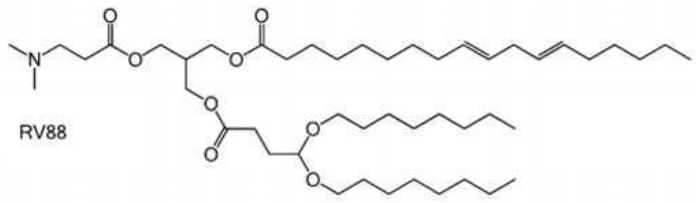
[1999] *T 세포 표적화*

[2000] 전달 비히클을 T-세포에 표적화하기 위해, T 세포 항원 결합제, 예를 들어, 항-CD8 항체를 전달 비히클의 표면에 커플링시킨다. 항-T 세포 항원 항체를 PBS 중의 EDTA의 존재 하에서 과량의 DTT로 약간 환원시켜 유리 힌지 영역 티올을 노출시킨다. DTT를 제거하기 위해서, 항체를 탈염 칼럼에 통과시킨다. 이종이작용성 가교제 SM(PEG)24를 사용하여 circRNA-로딩된 전달 비히클의 표면에 항체를 고정시킨다(아민기가 PEG 지질의 헤드기에 존재하고, DTT, 아민과 티올기 간의 SM(PEG)24 가교에 의해서 항체 상의 유리 티올기가 생성되었다). 전달 비히클을 먼저 과량의 SM(PEG)24와 함께 인큐베이션시키고, 원심분리시켜 미반응 가교제를 제거한다. 이어서 활성화된 전달 비히클을 과량의 감소된 항-T 세포 항원 항체와 함께 인큐베이션시킨다. 결합되지 않은 항체를 원심 여과 장치를 사용하여 제거한다.

[2001] **실시예 29**

[2002] *RV88을 사용한 RNA 함유 전달 비히클.*

[2003] 본 실시예에서, RNA 함유 전달 비히클을 circRNA의 전달을 위해 양이온성 지질 RV88과 함께 2-D 보텍스 미세유체 칩을 사용하여 합성한다.



[2004]

[2005] [표 32a]

물질 및 기기	공급원	카탈로그 번호
1M Tris-HCl, pH 8.0, 멸균	테크노바사	T1080
5M 염화나트륨 용액	테크노바사	S0250
QB 시트르산염 완충제, pH 6.0 (100 mM)	테크노바사	Q2446
뉴클레아제-무함유 수	암비온사	AM9937
Triton X-100	시그마-알드리치사	T8787-100ML
RV88	GVK 바이오사	
DSPC	리포이드사	556500
콜레스테롤	시그마사	C3045-5G
PEG2K	아반티 폴라 리피즈사	880150
에탄올	아크로스 오가닉사	615090010
5 mL보로실리케이트 유리 바이알	쎄모 사이언티픽사	ST5-20
PD MiniTrap G-25 탈염 칼럼	지이 헬쓰케어사	VWR Cat. #95055-984
Quant-iT RiboGreen RNA 검정 키트	몰레큘러 프로브즈사/라이프 테크놀로지즈사	R11490
흑색 96-웰 마이크로플레이트	그레이어사	655900

[2006]

[2007] RV88, DSPC 및 콜레스테롤을 모두 보로실리카 바이알에서 10㎎/㎖ 농도의 에탄올에서 제조한다. 지질 14:0-PEG2K PE를 보로실리카 유리 바이알에서도 4㎎/㎖의 농도로 제조한다. 스톡 농도에서 지질의 용해를 2분 동안 에탄올에서 지질의 초음파처리에 의해 달성한다. 그런 다음 용액을 10분 동안 37℃에서 170rpm으로 설정된 궤도 틸팅 진탕기에서 가열한다. 그런 다음 바이알을 최소 45분 동안 26℃에서 평형화한다. 이어서, 표 32b에 나타낸 바와 같은 부피의 스톡 지질을 첨가함으로써 지질을 혼합한다. 이어서, 용액을 최종 지질 농도가 7.92㎎/㎖가 되도록 에탄올로 조정한다.

[2008] [표 32b]

조성	MW	%	nmol	mg	스톡 (mg/ml)	ul	에탄올 (ul)
RV88	794.2	40%	7200	5.72	10	571.8	155.3
DSPC	790.15	10%	1800	1.42	10	142.2	
콜레스테롤	386.67	48%	8640	3.34	10	334.1	
PEG2K	2693.3	2%	360	0.97	4	242.4	

[2009]

[2010] pH 6.0의 75mM 시트르산염 완충액 및 1.250㎎/㎖의 RNA 농도를 갖는 스톡 용액으로서 RNA를 제조한다. 그 다음, RNA의 농도를 26℃로 평형화된 pH 6.0에서 75mM 시트르산염 완충액을 사용하여 0.1037㎎/㎖로 조정한다. 그 다음, 용액을 26℃에서 최소 25분 동안 인큐베이션시킨다.

[2011] 미세유체 챔버를 에탄올로 세정하고, RNA 용액을 주사기에 로딩하고, 에탄올 지질을 다른 주사기에 로딩하여 neMYSIS 주사기 펌프를 준비한다. 두 주사기 모두를 로딩하고, neMESYS 소프트웨어의 제어 하에 둔다. 그 다음, 용액을 2의 수성상 대 유기의 비율 및 22㎖/분(RNA의 경우 14.67㎖/분, 지질 용액의 경우 7.33㎖/분)의 총 유량으로 혼합 칩에 적용한다. 두 펌프를 동시에 시작한다. 미세유체 칩에서 흘러나온 혼합액을 4×1㎖ 분획으로 수집하고, 첫 번째 분획은 폐기물로 버린다. RNA-리포솜을 함유하는 나머지 용액을 G-25 mini 탈염 칼럼을 사용함으로써 10mM Tris-HCl, 1mM EDTA, pH 7.5로 교환한다. 완충액 교환 후, 물질을 각각 DLS 분석 및 Ribogreen 검정을 통해 크기 및 RNA 포획에 대해 특징규명한다.

[2012] **실시예 30**

[2013] *RV94을 사용한 RNA 함유 전달 비히클.*

[2014] 본 실시예에서, RNA 함유 리포솜을 circRNA의 전달을 위해 양이온성 지질 RV94과 함께 2-D 보텍스 미세유체 칩을 사용하여 합성한다.

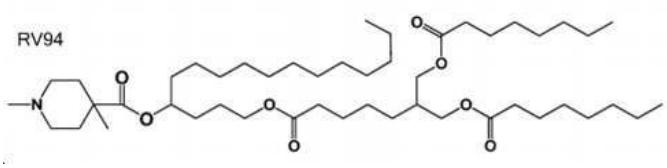
[2015] 

표 33

[2016]

물질 및 기기	공급원	카탈로그 번호
1M Tris-HCl, pH 8.0, 멸균	테크노바사	T1080
5M 염화나트륨 용액	테크노바사	S0250
QB 시트르산염 완충제, pH 6.0 (100 mM)	테크노바사	Q2446
뉴클레아제-무함유 수	암비온사	AM9937
Triton X-100	시그마-알드리치사	T8787-100ML
RV94	GVK 바이오사	
DSPC	리포이드사	556500
콜레스테롤	시그마사	C3045-5G
PEG2K	아반티 폴라 리피즈사	880150
에탄올	아크로스 오가닉사	615090010
5 mL 보로실리케이트 유리 바이알	써모 사이언티픽사	ST5-20
PD MiniTrap G-25 탈염 칼럼	지이 헬쓰케어사	VWR Cat. #95055-984
Quant-iT RiboGreen RNA 검정 키트	몰레큘러 프로브즈사/라이프 테크놀로지즈사	R11490
흑색 96-웰 마이크로플레이트	그레니어사	655900

[2017] 지질을 표 34에 제시된 물질 양을 사용하여 실시예 29에서와 같이 7.92㎎/㎖의 최종 지질 농도로 제조하였다.

표 34

[2018]

조성	MW	%	nmol	mg	스톡 (mg/ml)	ul	에탄올 (ul)
RV94	808.22	40%	2880	2.33	10	232.8	155.3
DSPC	790.15	10%	720	0.57	10	56.9	
콜레스테롤	386.67	48%	3456	1.34	10	133.6	
PEG2K	2693.3	2%	144	0.39	4	97.0	

[2019] circRNA의 수용액을 pH 6.0의 75mM 시트레이트 완충액과 1.250㎎/㎖의 circRNA를 포함하는 스톡 용액으로 제조한다. 그 다음, RNA의 농도를 26℃로 평형화된 pH 6.0에서 75mM 시트르산염 완충액을 사용하여 0.1037㎎/㎖로 조정한다. 그 다음, 용액을 26℃에서 최소 25분 동안 인큐베이션시킨다.

[2020] 미세유체 챔버를 에탄올로 세정하고, RNA 용액을 주사기에 로딩하고, 에탄올 지질을 다른 주사기에 로딩하여 neMYSIS 주사기 펌프를 준비한다. 두 주사기 모두를 로딩하고, neMESYS 소프트웨어의 제어 하에 둔다. 그 다음,

용액을 2의 수성상 대 유기의 비율 및 22㎖/분(RNA의 경우 14.67㎖/분, 지질 용액의 경우 7.33㎖/분)의 총 유량으로 혼합 칩에 적용한다. 두 펌프를 동시에 시작한다. 미세유체 칩에서 흘러나온 혼합액을 4×1㎖ 분획으로 수집하고, 첫 번째 분획은 폐기물로 버린다. circRNA-전달 비히클을 함유하는 나머지 용액을 상기에 기재된 바와 같이 G-25 mini 탈염 칼럼을 사용함으로써 10mM Tris-HCl, 1mM EDTA, pH 7.5로 교환한다. 완충액 교환 후, 물질을 각각 DLS 분석 및 Ribogreen 검정을 통해 크기 및 RNA 포획에 대해 특징규명한다. 리포솜의 생물물리학적 분석을 표 35에 제시한다.

표 35

샘플 명칭	N:P 비율	TFR	비율	RNA 캡슐화 양	RNA 캡슐화 수율	크기	
		ml/분	(수성/유기상)	(μg/ml)	%	d.nm	PDI
SAM-RV94	8	22	2	31,46	86,9	113,1	0,12

[2021]

[2022] **실시예 31**

[2023] *인라인 혼합을 위한 일반적인 프로토콜.*

[2024] 하나는 지질을 함유하고 다른 하나는 circRNA를 함유하는 개별 및 별개의 스톡 용액을 제조한다. 목적하는 지질 또는 지질 혼합물, DSPC, 콜레스테롤 및 PEG 지질을 함유하는 지질 스톡을 90% 에탄올에 용해시킴으로써 준비한다. 나머지 10%는 낮은 pH 시트르산염 완충액이다. 지질 스톡의 농도는 4mg/㎖이다. 이 시트르산염 완충액의 pH는 사용되는 지질의 유형에 따라 pH 3내지 pH 5 범위일 수 있다. circRNA를 또한 4mg/㎖ 농도의 시트르산염 완충액에 용해시킨다. 5㎖의 각각의 스톡 용액을 준비한다.

[2025] 스톡 용액은 완전히 투명하고, 지질은 circRNA와 조합하기 전에 완전히 가용화되도록 보장된다. 스톡 용액을 가열하여 지질을 완전히 가용화시킬 수 있다. 이러한 공정에서 사용되는 circRNA는 비변형된 올리고뉴클레오타이드 또는 변형된 올리고뉴클레오타이드일 수 있으며, 콜레스테롤과 같은 친유성 모이어티와 접합될 수 있다.

[2026] 각각의 용액을 T-접합부로 펌핑하여 개별 스톡을 배합한다. 이중-헤드 Watson-Marlow 펌프를 사용하여 2개의 스트림의 시작 및 중지를 동시에 제어한다. 1.6mm 폴리프로필렌 튜빙을 선형 유량을 증가시키기 위해 0.8mm 튜빙으로 추가로 축소시킨다. 폴리프로필렌 라인(ID=0.8mm)을 T-접합부의 양쪽에 부착한다. 폴리프로필렌 T는 1.6mm의 선형 모서리를 갖고, 결과적인 부피는 4.1mm^3이다. 폴리프로필렌 라인의 큰 단부(1.6mm) 각각을 가용화된 지질 스톡 또는 가용화된 circRNA를 함유하는 시험 튜브에 배치한다. T-접합부 이후에, 배합된 스트림이 나가는 곳에 단일 튜브를 배치한다. 그 다음 튜빙을 2배 부피의 PBS가 있는 용기로 연장하고, 이것을 신속하게 교반한다. 펌프의 유량은 300rpm 또는 110㎖/분의 설정이다. 에탄올을 제거하고, 투석에 의해서 PBS에 대해 교환한다. 그 다음 지질 제형을 원심분리 또는 정용여과를 사용하여 적절한 작업 농도로 농축시킨다.

[2027] C57BL/6 마우스(미국 매사추세츠주 소재의 찰스 리버 랩스(Charles River Labs))에게 꼬리 정맥 주사를 통해 식염수 또는 제형화된 circRNA를 제공한다. 투여 후 다양한 시점에서, 혈청 샘플을 안와 후 출혈에 의해 수집한다. 인자 VII 단백질의 혈청 수준을 발색 검정(Biophen FVTI, 미국 오하이오주 애니아라사(Aniara Corporation))을 사용하여 샘플에서 결정한다. 인자 VII의 간 RNA 수준을 결정하기 위해서, 동물을 희생시키고, 간을 수거하고, 액체 질소에서 급속 냉동한다. 조직 용해물을 냉동 조직으로부터 준비하고, 인자 VII의 간 RNA 수준을 분지화 DNA 검정(QuantiGene Assay, 미국 캘리포니아주 소재의 파노믹스사(Panomics))을 사용하여 정량한다.

[2028] C57BL/6 마우스에서 정맥내(볼러스) 주사 후 48시간에 FVTI siRNA 처리된 동물에서 FVII 활성을 평가한다. 마이크로플레이트 규모에서 제조자의 지시에 따라, 혈청 또는 조직에서 단백질 수준을 결정하기 위한 상업적으로 입수 가능한 키트를 사용하여 FVII를 측정한다. 미처리 대조군 마우스에 대해 FVII 감소를 결정하고, 그 결과를 잔류 FVII%로 표현한다. 2가지 용량 수준(0.05 및 0.005mg/kg FVII siRNA)을 각각의 신규 리포솜 조성물의 스크린에 사용한다.

[2029] **실시예 32**

[2030] *사전형성된 소포를 사용한 circRNA 제형*

[2031] 사전형성된 비히클을 함유하는 양이온성 지질을 사전 형성된 소포 방법을 사용하여 제조한다. 양이온성 지질, DSPC, 콜레스테롤 및 PEG-지질을 각각 40/10/40/10의 몰비로 에탄올에 용해시킨다. 최종 에탄올 및 지질 농도가 각각 30%(vol/vol) 및 6.1mg/mℓ가 되도록 혼합하면서 수성 완충액(50mM 시트르산염, pH 4)에 지질 혼합물을 첨가하고, 실온에서 2분 동안 평형을 유지하도록 한 다음 압출한다. Nicomp 분석에 의해 결정된 바와 같이 70 내지 90nm의 소포 직경이 얻어질 때까지, Lipex 압출기(캐나다 BC주 밴쿠버 소재의 노던 리피즈사(Northern Lipids)를 사용하여 22℃에서 2개의 적층된 80nm 기공 크기 필터(Nuclepore)를 통해 수화된 지질을 압출한다. 작은 소포를 형성하지 않는 양이온성 지질 혼합물의 경우, DSPC 헤드기 상의 포스페이트기를 양성자화하기 위해서 더 낮은 pH 완충액(50mM 시트르산염, pH 3)를 사용하여 지질 혼합물을 수화시키는 것이 안정적인 70 내지 90 nm 소포를 형성하는 데 도움이 된다.

[2032] FVII circRNA(50mM 시트르산염, 30% 에탄올을 함유하는 pH 4 수용액에 가용화됨)를 혼합하면서 약 5mℓ/분의 속도로 35℃로 미리 평형화된 소포에 첨가한다. 0.06(wt wt)의 최종 표적 circRNA/지질 비율이 달성된 후, 혼합물을 35℃에서 추가로 30분 동안 인큐베이션시켜 FVII RNA의 소포 재조직화 및 캡슐화를 허용한다. 그 다음, 에탄올을 제거하고, 외부 완충액을 투석 또는 접선 유동 정용여과에 의해 PBS(155mM NaCl, 3mM Na2HPO4, ImM KH2PO4, pH 7.5)로 교체한다. 크기 배제 스핀 칼럼 또는 이온 교환 스핀 칼럼을 사용하여 캡슐화되지 않은 RNA를 제거한 후 최종 캡슐화된 circRNA-대-지질 비율을 결정한다.

[2033] **실시예 33**

[2034] *조작된 원형 RNA로부터의 삼중특이적 항원 결합 단백질의 발현*

[2035] 원형 RNA를 하기를 포함하도록 설계한다: (1) 3' 포스트 스플라이싱 그룹 I 인트론 단편; (2) 내부 리보솜 유입 부위(IRES); (3) 삼중특이적 항원-결합 단백질 암호 영역; 및 (4) 3' 상동성 영역. 삼중특이적 항원-결합 단백질 영역을 작제하여 표적 항원, 예를 들어, GPC3에 결합할 예시적인 삼중특이적 항원-결합 단백질을 생성시킨다.

[2036] *scFv CD3 결합 도메인의 생성*

[2037] 인간 CD3엡실론 쇄 정규 서열은 유니프로트 수탁 번호 P07766이다. 인간 CD3감마 쇄 정규 서열은 유니프로트 수탁 번호 P09693이다. 인간 CD3델타쇄 정규 서열은 유니프로트 수탁 번호 P043234이다. CD3엡실론, CD3감마 또는 CD3델타에 대한 항체를 공지된 기술, 예컨대, 친화도 성숙을 통해 생성시킨다. 뮤린 항-CD3 항체가 출발 물질로 사용되는 경우, 뮤린 항-CD3 항체의 인간화가 임상 환경에서 바람직한데, 여기서 마우스-특이적 잔기는 본 명세서에 기재된 삼중특이적 항원-결합 단백질의 치료를 제공받는 대상체에서 인간-항-마우스 항원(HAMA) 반응을 유도할 수 있다. 인간화는 뮤린 항-CD3 항체로부터의 CDR 영역을, 선택적으로 CDR 및/또는 프레임워크 영역에 대한 다른 변형을 포함하는 적절한 인간 생식계열 억셉터 프레임워크에 이식함으로써 달성된다.

[2038] 따라서 인간 또는 인간화된 항-CD3 항체를 사용하여 삼중특이적 항원-결합 단백질의 CD3 결합 도메인에 대한 scFv 서열을 생성시킨다. 인간 또는 인간화된 VL 및 VH 도메인을 암호화하는 DNA 서열을 얻고, 작제물에 대한 코돈을 선택적으로 호모 사피엔스 세포에서의 발현을 위해 최적화한다. VL 및 VH 도메인이 scFv에 나타나는 순서는 다양하며(즉, VL-VH 또는 VH-VL 배향), "G4S" 또는 "G_4S" 소단위$(G_4S)_3$의 3개 카피가 가변 도메인을 연결하여 scFv 도메인을 생성한다. 항-CD3 scFv 플라스미드 작제물은 선택적인 Flag, His 또는 기타 친화성 태그를 가질 수 있으며, 이를 HEK293 또는 기타 적합한 인간 또는 포유동물 세포주 내에 전기천공하고, 정제시킨다. 검증 분석은 FACS에 의한 결합 분석, Proteon을 사용한 운동 분석 및 CD3 발현 세포의 염색을 포함한다.

[2039] *scFv 글리피칸-3(GPC3) 결합 도메인의 생성*

[2040] 글리피칸-3(GPC3)은 간세포 암종에 존재하지만 건강한 정상 간 조직에는 존재하지 않는 세포 표면 단백질 중 하나이다. 이것은 간세포 암종에서 흔히 증가하는 것으로 관찰되며 HCC 환자의 나쁜 예후와 관련이 있다. 그것은 Wnt 신호를 활성화한다고 알려져 있다. MDX-1414, HN3, GC33 및 YP7을 포함하는 GPC3 항체가 생성되었다.

[2041] CD3에 대한 scFv 결합 도메인의 생성을 위한 상기 방법과 유사하게 GPC-3 또는 또 다른 표적 항원에 대한 scFv 결합을 생성시킨다.

[2042] *시험관내에서 삼중특이적 항원-결합 단백질의 발현*

[2043] CHO-K1 중국 햄스터 난소 세포(ATCC, CCL-61)의 유도체인 CHO 세포 발현 시스템(Flp-In®; 라이프 테크놀로지즈사)(문헌[Kao and Puck, Proc. Natl. Acad Sci USA 1968; 60(4):1275-81])을 사용한다. 부착 세포를 라이프 테크놀로지즈사에서 제공하는 표준 세포 배양 프로토콜에 따라 계대배양한다.

[2044] 현탁액에서의 성장에 대한 적응을 위해, 세포를 조직 배양 플라스크에서 탈착시키고, 무혈청 배지에 두었다. 현탁액 적응 세포를 10% DMSO가 포함된 배지에서 동결 보존한다.

[2045] 분비된 삼중특이적 항원-결합 단백질을 안정적으로 발현하는 재조합 CHO 세포주를 현탁-적응 세포의 형질주입에 의해 생성시킨다. 항생제 하이그로마이신 B로의 선택 동안, 생존 세포 밀도를 주 2회 측정하고, 세포를 원심분리시키고, 0.1 x 10^6개 생존 세포/㎖의 최대 밀도로 새로운 선택 배지에 재현탁시킨다. 삼중특이적 항원-결합 단백질을 안정적으로 발현하는 세포 풀을 선택 2 내지 3주 후에 회수하고, 그 시점에 세포를 진탕 플라스크에서 표준 배양 배지로 옮긴다. 재조합 분비 단백질의 발현을 단백질 겔 전기영동 또는 유세포 분석을 수행하여 확인한다. 안정적인 세포 풀을 DMSO 함유 배지에서 동결 보존한다.

[2046] 삼중특이적 항원-결합 단백질을 세포 배양 상청액으로의 분비에 의해 안정적으로 형질주입된 CHO 세포주의 10일 유가식 배양으로 생산한다. 세포 배양 상청액을 전형적으로 75% 초과의 배양 생존율에서 10일 후에 수거한다. 격일로 생산 배양액에서 샘플을 수집하고, 세포 밀도 및 생존력을 평가한다. 수거일에, 세포 배양 상청액을 추가 사용 전에 원심분리 및 진공 여과에 의해 정화시킨다.

[2047] 세포 배양 상청액에서 단백질 발현 역가 및 생성물 온전성을 SDS-PAGE에 의해 분석한다.

[2048] *삼중특이적 항원-결합 단백질의 정제*

[2049] 삼중특이적 항원-결합 단백질을 2-단계 절차로 CHO 세포 배양 상청액으로부터 정제시킨다. 제1 단계에서 작제물을 친화성 크로마토그래피시키고, 제2 단계에서 Superdex 200 상에서 분취용 크기 배제 크로마토그래피(SEC)시킨다. 샘플을 완충제 교환시키고, 한외여과에 의해 1㎎/㎖ 초과의 전형적인 농도로 농축시키고, 각각 환원 및 비환원 조건 하에서 SDS PAGE한 후 항-(반감기 연장 도메인) 또는 항 이디오타입 항체를 사용하여 면역블로팅함으로써 그리고 분석용 SEC를 사용함으로써 최종 샘플의 순도 및 균질성(전형적으로 90% 초과)을 평가한다. 정제된 단백질을 사용할 때까지 -80℃에서 분취물로 저장한다.

[2050] **실시예 34**

[2051] *반감기 연장 도메인을 갖는 조작된 원형 RNA의 발현은 반감기 연장 도메인을 갖지 않는 것보다 개선된 약동학적 매개변수를 갖는다.*

[2052] 실시예 23의 circRNA 분자 상에 암호화된 삼중특이적 항원-결합 단백질을 근육내로 0.5㎎/㎏ 볼러스 주사로서 시노몰거스 원숭이에게 투여한다. 또 다른 시노몰거스 원숭이 군에게 CD3 및 GPC-3에 대한 결합 도메인을 갖지만 반감기 연장 도메인이 없는 크기의 circRNA 분자에 암호화된 대등한 단백질을 제공한다. 제3 및 제4 군에게 각각 CD3 및 반감기 연장 도메인 결합 도메인을 갖는 circRNA 분자에 암호화된 단백질 및 GPC-3 및 반감기 연장 도메인을 갖는 단백질을 제공한다. circRNA에 의해 암호화된 두 단백질은 삼중특이적 항원-결합 단백질과 크기가 대등하다. 각각의 시험군은 5마리의 원숭이로 이루어진다. 혈청 샘플을 제시된 시점에서 취하고, 연속적으로 희석시키고, CD3 및/또는 GPC-3에 대한 결합 ELISA를 사용하여 단백질의 농도를 결정한다.

[2053] 시험 물품 혈장 농도를 사용하여 약동학 분석을 수행한다. 각각의 시험 항목에 대한 군 평균 혈장 데이터는 투여 후 시간에 대해 플로팅하는 경우 다중-지수 프로파일을 따른다. 이러한 데이터는 분포 및 제거 단계에 대한 1차 속도 상수 및 볼러스 입을 갖는 표준 2-구획 모델에 적합하다. i.v에 대한 데이터의 최적 적합에 대한 일반 수학식은 다음과 같다: $c(t)=Ae^{-at}+Be^{-pt}$, 식 중, c(t)는 시간 t에서의 혈장 농도이고, A 및 B는 Y축의 절편이고, a 및 β는 각각 분포 및 제거 단계에 대한 겉보기 1차 속도 상수이다. a 단계는 제거의 초기 단계이며, 동물의 모든 세포외 유체로의 단백질 분포를 반영하는 반면, 디케이 곡선의 두 번째 또는 β 단계 부분은 진정한 혈장 제거를 나타낸다. 이러한 수학식에 적합한 방법은 당업계에 공지되어 잇다. 예를 들어, A=D/V(a-k21)/(a-p), B=D/V(p-k21)/(a-p), a 및 β(α>β의 경우)는 이차 방정식의 근이다: r2+(k12+k21+k10)r+k21k10=0, V=분포 부피, k10=제거 속도, k12=구획 1에서 구획 2로의 전달 속도 및 k21=구획 2에서 구획 1로의 전달 속도 및 D = 투여된 용량의 추정 매개변수 사용

[2054] 데이터 분석: 농도 대 시간 프로파일의 그래프를 KaleidaGraph(KaleidaGraph™ V. 3.09 Copyright 1986-1997. 시너지 소프트웨어사(Synergy Software.), 미국 팬실베니아주 리딩 소재)를 사용하여 작성한다. 보고 가능 한

값보다 작은 것으로 보고된 값(LTR)은 PK 분석에 포함하지 않고, 그래프에 나타내지 않는다. 약동학적 매개변수는 WinNonlin 소프트웨어(WinNonlin® Professional V. 3.1 WinNonlin™ Copyright 1998-1999. 파사이트사(Pharsight Corporation.), 미국 캘리포니아주 마운틴 뷰 소재)를 사용한 구획 분석에 의해 결정한다. 약동학적 매개변수는 문헌[Ritschel W A and Kearns G L, 1999, EST: Handbook Of Basic Pharmacokinetics Including Clinical Applications, 5th edition, American Pharmaceutical Assoc., Washington, D C]에 기재된 바와 같이 계산한다.

[2055] 실시예 23의 circRNA 분자에 암호화된 삼중특이적 항원-결합 단백질은 반감기 연장 도메인이 결여된 단백질과 비교하여 제거 반감기의 증가와 같은 개선된 약동학적 파라미터를 가질 것으로 예상된다.

[2056] **실시예 35**

[2057] *삼중특이적 항원-결합 단백질의 세포독성*

[2058] 실시예 23의 circRNA 분자에 암호화된 삼중특이적 항원-결합 단백질을 GPC-3+ 표적 세포에 대한 T 세포 의존성 세포독성의 매개에 대해 시험관내에서 평가한다.

[2059] 형광 표지된 GPC3 표적 세포를 실시예 23의 삼중특이적 항원-결합 단백질의 존재 하에서 효과기 세포로서 무작위 공여자 또는 T-세포의 단리된 PBMC와 함께 인큐베이션시킨다. 37℃에서 4시간 동안 인큐베이션시킨 후, 가습 인큐베이터에서, 표적 세포로부터의 상청액으로 형광 염료의 방출을 분광형광계에서 측정한다. 실시예 23의 삼중특이적 항원-결합 단백질 없이 인큐베이션된 표적 세포 및 인큐베이션 종료 시 사포닌 첨가에 의해 완전히 용해된 표적 세포를 각각 음성 및 양성 대조군으로 제공한다.

[2060] 측정된 잔류하는 생존 표적 세포에 기초하여, 하기 방정식에 따라 특정 세포 용해의 백분율을 계산한다: [1-(생존 표적(샘플)의 수/생존 표적의 수(자발)의 수)] x 100%. S자형 용량 반응 곡선 및 EC50 값을 GraphPad 소프트웨어를 사용하여 비선형 회귀/4-매개변수 로지스틱 피팅에 의해 계산한다. 주어진 항체 농도에 대해 얻은 용해 값을 사용하여 Prism 소프트웨어를 사용하여 4개 매개변수 로지스틱 피트 분석에 의해 S자형 용량-반응 곡선을 계산한다.

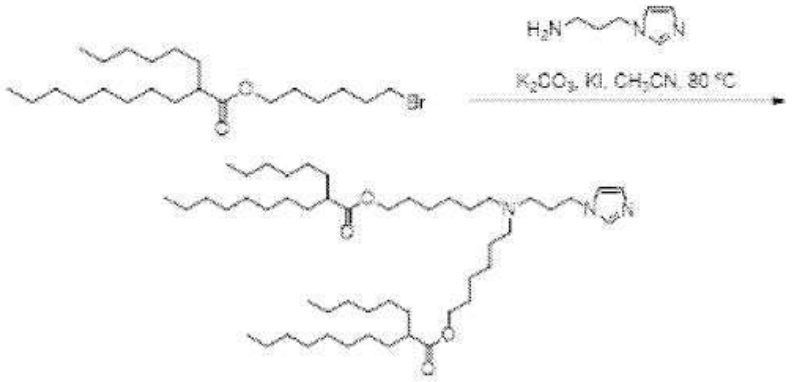
[2061] **실시예 36**

[2062] *이온화 가능한 지질의 합성*

[2063] ***38.1*** *((3-(2-메틸-1H-이미다졸-1-일)프로필)아잔다이일)비스(헥산-6,1-다이일) 비스(2-헥실데칸오에이트)(**지질 27, 표 10a**) 및 ((3-(1H-이미다졸-1-일)프로필)아잔다이일)비스(헥산-6,1-다이일) 비스(2-헥실데칸오에이트))(**지질 26, 표 10a**)의 합성*

[2064] 응축기가 연결된 100㎖ 둥근 바닥 플라스크에, 3-(1*H*-이미다졸-1-일)프로판-1-아민(100㎎, 0.799㎜ol) 또는 3-(2-메틸-1H-이미다졸-1-일)프로판-1-아민(0.799㎜ol), 6-브로모헥실 2-헥실데칸오에이트(737.2㎎, 1.757㎜ol), 탄산칼륨(485㎎, 3.515㎜ol) 및 아이오딘화칼륨(13㎎, 0.08㎜ol)을 아세토니트릴(30㎖)에서 혼합하고, 반응 혼합물을 48시간 동안 80℃까지 가열시켰다. 혼합물을 실온으로 냉각시키고, 셀라이트 패드로 여과하였다. 여과액을 에틸 아세테이트로 희석시켰다. 물, 염수로 세척한 후, 무수 황산나트륨 상에서 건조시켰다. 용매를 증발시키고, 조 잔류물을 플래시 크로마토그래피(SiO_2: CH_2Cl_2= CH_2Cl_2 중의 100%에서 10%의 메탄올)로 정제시키고, 무색 오일 생성물을 얻었다(92㎎, 15%). *((3-(1H-이미다졸-1-일)프로필)아잔다이일)비스(헥산-6,1-다이일) 비스(2-헥실데칸오에이트))의* 분식은 $C_{50}H_{95}N_3O_4$ 이고, 분자량(M_w)은 801.7이다.

[2065] *((3-(1H-이미다졸-1-일)프로필)아잔다이일)비스(헥산-6,1-다이일) 비스(2-헥실데칸오에이트))(지질 26, 표 10a)*의 합성에 대한 반응식.



[2066]

[2067] 지질 26의 특징규명을 LC-MS로 수행하였다. 도 27A 내지 도 27C는 지질 26의 특징규명을 도시한다. 도 27A는 지질 26에 대해서 관찰된 양성자 NMR을 도시한다. 도 27B는 지질 26에 대한 대표적인 LC/MS 트레이스이며, 총 이온 및 UV 크로마토그램이 도시되어 있다.

[2068] ***38.2*** *지질 22-S14의 합성*

[2069] ***38.2.1*** *2-(테트라데실티오)에탄-1-올의 합성*

[2070] 25℃에서 아세토니트릴(200㎖) 중의 2-설파닐에탄올(5.40g, 69.11mmol, 4.82㎖, 0.871*eq*)의 혼합물에 1-브로모테트라데칸(22g, 79.34mmol, 23.66㎖, 1*eq*) 및 탄산칼륨(17.55g, 126.95mmol, 1.6*eq*)을 첨가하였다. 반응 혼합물을 40℃까지 가온시키고, 12시간 동안 교반하였다. TLC(에틸 아세테이트/석유 에터 = 25/1, R_f = 0.3, I_2로 염색)는 출발 물질이 완전히 소모되었고, 새로운 주요 스팟이 생성되었다는 것을 나타내었다. 반응 혼합물을 여과하고, 여과 케이크를 아세토니트릴(50㎖)로 세척하고, 그 다음 여과액을 진공 하에서 농축시켜 잔류물을 얻었고, 이것을 실리카겔 상의 칼럼(에틸 아세테이트/석유 에터 = 1/100에서 1/25)으로 정제시켜 2-(테트라데실티오)에탄-1-올(14g, 수율 64.28%)을 백색 고체로서 제공하였다.

[2071] ^{1}H NMR (ET36387-45-P1A, 400 MHz, 클로로폼-*d*) δ 0.87 - 0.91 (m, 3 H) 1.27 (s, 20 H) 1.35 - 1.43 (m, 2 H) 1.53 - 1.64 (m, 2 H) 2.16 (br s, 1 H) 2.49 - 2.56 (m, 2 H) 2.74 (t, *J* = 5.93 Hz, 2 H) 3.72 (br d, *J* = 4.89 Hz, 2 H). 도 28은 상응하는 핵 자기 공명(NMR) 스펙트럼을 나타낸다.

[2072] ***38.2.2*** *2-(테트라데실티오)에틸 아크릴레이트의 합성*

[2073] 다이클로로메탄(240㎖) 중의 2-(테트라데실티오)에탄-1-올(14g, 51.00mmol, 1eq)의 용액에 트라이에틸아민(7.74g, 76.50mmol, 10.65㎖, 1.5eq) 및 프로프-2-엔오일 클로라이드(5.54g, 61.20mmol, 4.99㎖, 1.2eq)를 0℃에서 질소 하에서 적가하였다. 반응 혼합물을 25℃까지 가온시키고, 12시간 동안 교반하였다. TLC(에틸 아세테이트/석유 에터 = 25/1, R_f = 0.5, I_2로 염색)는 출발 물질이 완전히 소모되었고, 새로운 주요 스팟이 생성되었다는 것을 나타내었다. 반응 용액을 감압 하에서 농축시켜 조물질을 얻었고, 이것을 실리카겔 상의 칼럼(에틸 아세테이트/석유 에터 = 1/100 to 1/25)으로 정제시켜 2-(테트라데실티오)에틸 아크릴레이트(12g, 수율 71.61%)를 무색 오일로서 제공하였다.

[2074] ^{1}H NMR (ET36387-49-P1A, 400 MHz, 클로로폼-*d*) δ 0.85 - 0.93 (m, 3 H) 1.26 (s, 19 H) 1.35 - 1.43 (m, 2 H) 1.53 - 1.65 (m, 2 H) 2.53 - 2.62 (m, 2 H) 2.79 (t, *J* = 7.03 Hz, 2 H) 4.32 (t, *J* = 7.03 Hz, 2 H) 5.86 (dd, *J* = 10.39, 1.47 Hz, 1 H) 6.09 - 6.19 (m, 1 H) 6.43 (dd, *J* = 17.30, 1.41 Hz, 1 H). 도 29는 상응하는 핵 자기 공명(NMR) 스펙트럼을 나타낸다.

[2075] ***38.2.3*** *비스(2-(테트라데실티오)에틸) 3,3'-((3-(2-메틸-1H-이미다졸-1-일)프로필)아잔다이일)다이프로피오네이트(**지질 22-S14**)의 합성*

[2076] 플라스크에 3-(2-메틸-1H-이미다졸-1-일)프로판-1-아민(300㎎, 2.16mmol) 및 2-(테트라데실티오)에틸 아크릴레이트 (1.70g, 5.17mmol)를 넣었다. 니트 반응 혼합물을 80℃까지 가열시키고, 48시간 동안 교반하였다. TLC(에틸 아세테이트, R_f = 0.3, I_2로 염색, 수산화암모늄 1방울 첨가)는 출발 물질이 완전히 소모되었고, 새로운 주요 스팟이 형성되었다는 것을 나타내었다. 반응 혼합물을 다이클로로메탄(4㎖)으로 희석시키고, 실리카겔 상의 칼럼(석유 에터/에틸 아세테이트 = 3/1에서 0/1, 0.1% 수산화암모늄 첨가)으로 정제시켜 비스(2-(테트라데실티오)에

틸) 3,3'-((3-(2-메틸-1H-이미다졸-1-일)프로필)아잔다이일)다이프로피오네이트 (501㎎, 수율 29.1%)를 무색 오일로서 얻었다.

[2077] ^{1}H NMR (ET36387-51-P1A, 400 MHz, 클로로폼-*d*) δ 0.87 (t, *J* = 6.73 Hz, 6 H) 1.25 (s, 40 H) 1.33 - 1.40 (m, 4 H) 1.52 - 1.61 (m, 4 H) 1.81 - 1.90 (m, 2 H) 2.36 (s, 3 H) 2.39 - 2.46 (m, 6 H) 2.53 (t, *J* = 7.39 Hz, 4 H) 2.70 - 2.78 (m, 8 H) 3.84 (t, *J* = 7.17 Hz, 2 H) 4.21 (t, *J* = 6.95 Hz, 4 H) 6.85 (s, 1 H) 6.89 (s, 1 H). 도 30은 상응하는 핵 자기 공명(NMR) 스펙트럼을 나타낸다.

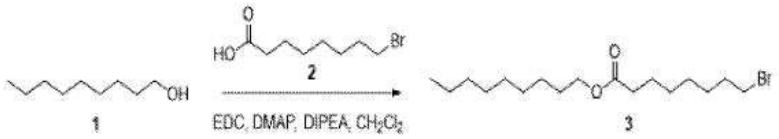
[2078] ***38.3*** *비스(2-(테트라데실티오)에틸) 3,3'-((3-(1H-이미다졸-1-일)프로필)아잔다이일)다이프로피오네이트(**지질 93-S14**)의 합성*

[2079] 플라스크에 3-(1H-이미다졸-1-일)프로판-1-아민(300㎎, 2.40㎜ol, 1eq) 및 2-(테트라데실티오)에틸 아크릴레이트(1.89g, 5.75㎜ol, 2.4eq)를 넣었다. 니트 반응 혼합물을 80℃까지 가열시키고, 48시간 동안 교반하였다. TLC (에틸 아세테이트, R_f = 0.3, I_2로 염색, 수산화암모늄 1방울 첨가)는 출발 물질이 완전히 소모되었고, 새로운 주요 스팟이 형성되었다는 것을 나타내었다. 반응 혼합물을 다이클로로메탄(4㎖)으로 희석시키고, 실리카겔 상의 칼럼(석유 에터/에틸 아세테이트 = 1/20에서 0/100, 0.1% 수산화암모늄 첨가)으로 정제시켜 비스(2-(테트라데실티오)에틸) 3,3'-((3-(1H-이미다졸-1-일)프로필)아잔다이일)다이프로피오네이트(512㎎, 수율 27.22%)를 무색 오일로서 얻었다.

[2080] ^{1}H NMR (ET36387-54-P1A, 400 MHz, 클로로폼-*d*) δ 0.89 (t, *J* = 6.84 Hz, 6 H) 1.26 (s, 40 H) 1.34 - 1.41 (m, 4 H) 1.58 (br t, *J* = 7.50 Hz, 4 H) 1.92 (t, *J* = 6.62 Hz, 2 H) 2.36 - 2.46 (m, 6 H) 2.55 (t, *J* = 7.50 Hz, 4 H) 2.75 (q, *J* = 6.84 Hz, 8 H) 3.97 (t, *J* = 6.95 Hz, 2 H) 4.23 (t, *J* = 6.95 Hz, 4 H) 6.95 (s, 1 H) 7.06 (s, 1 H) 7.51 (s, 1 H). 도 31은 상응하는 핵 자기 공명(NMR) 스펙트럼을 나타낸다.

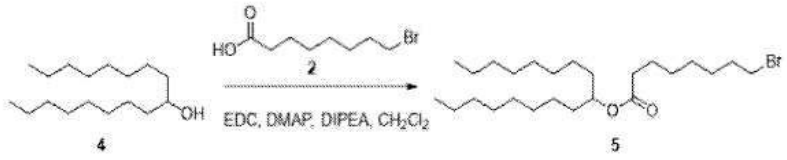
[2081] ***38.4*** *헵타데칸-9-일 8-((3-(2-메틸-1H-이미다졸-1-일)프로필)(8-(노닐옥시)-8-옥소옥틸)아미노)옥탄오에이트(지질 54, 표 10a)의 합성*

[2082] ***38.4.1*** *노닐 8-브로모옥탄오에이트(**3**)의 합성*

[2083] 

[2084] CH_2Cl_2(500㎖) 중의 8-브로모옥탄산(**2**)(18.6g, 83.18㎜ol) 및 노난-1-올(**1**)(10g, 69.32㎜ol)의 혼합물에 DMAP(1.7g, 13.86㎜ol), DIPEA(48㎖, 277.3㎜ol) 및 EDC(16g, 83.18㎜ol)를 첨가하였다. 반응을 실온에서 밤새 교반하였다. 반응 혼합물을 농축시킨 후, 조 잔류물을 에틸 아세테이트(500㎖)에 용해시키고, 1N HCl, 포화 $NaHCO_3$, 물 및 염수로 세척하였다. 유기층을 무수 Na_2SO_4 상에서 건조시켰다. 용매를 증발시키고, 조 잔류물을 플래시 크로마토그래피(SiO_2: 헥산 = 헥산 중의 100%에서 30%의 EtOAc)로 정제시키고, 무색 오일 생성물 **3**을 얻었다(9g, 37%).

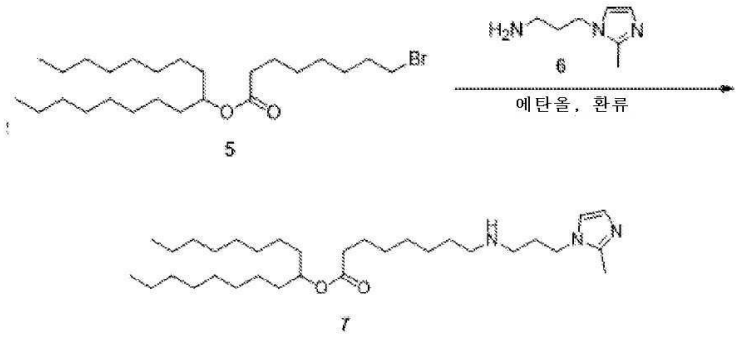
[2085] ***38.4.2*** *헵타데칸-9-일 8-브로모옥탄오에이트(**5**)의 합성*

[2086] 

[2087] CH_2Cl_2(300㎖) 중의 8-브로모옥탄산(**2**)(10g, 44.82㎜ol) 및 헵타데칸-9-올(**4**)(9.6g, 37.35㎜ol)의 혼합물에 DMAP(900㎎, 7.48㎜ol), DIPEA(26㎖, 149.7㎜ol) 및 EDC(10.7g, 56.03㎜ol)를 첨가하였다. 반응을 실온에서 밤새 교반하였다. 반응 혼합물을 농축시킨 후, 조 잔류물을 에틸 아세테이트(300㎖)에 용해시키고, 1N HCl, 포화 $NaHCO_3$, 물 및 염수로 세척하였다. 유기층을 무수 Na_2SO_4 상에서 건조시켰다. 용매를 증발시키고, 조 잔류물을

플래시 크로마토그래피(SiO_2: 헥산 = 헥산 중의 100%에서 30%의 EtOAc)로 정제시키고, 무색 오일 생성물 **5**를 얻었다(5g, 29%).

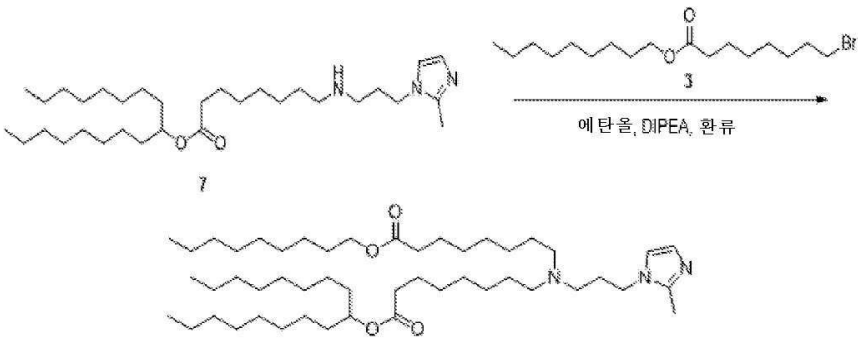
[2088] ***38.4.3*** *헵타데칸-9-일 8-((3-(2-메틸-1H-이미다졸-1-일)프로필)아미노)옥탄오에이트(**7**)의 합성*



[2089]

[2090] 응축기가 연결된 100 ㎖ 둥근 바닥 플라스크에서, 헵타데칸-9-일 8-브로모옥탄오에이트(**5**)(860㎎, 1.868㎜ol) 및 3-(2-메틸-1H-이미다졸-1-일)프로판-1-아민(**6**)(1.3g, 9.339㎜ol)을 에탄올(10㎖)에서 혼합하였다. 반응 혼합물을 밤새 환류까지 가열시켰다. MS(APCI)가 예측된 생성물을 나타내었다. 혼합물을 실온으로 냉각시키고, 농축시켰다. 조 잔류물을 플래시 크로마토그래피(SiO_2: CH_2Cl_2= CH_2Cl_2 중의 100%에서 10%의 메탄올+1%NH_4OH)로 정제시키고, 무색 오일 생성물 **7**을 얻었다(665㎎, 69%).

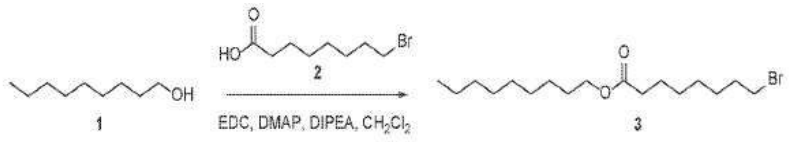
[2091] ***38.4.4****의 합성 헵타데칸-9-일 8-((3-(2-메틸-1H-이미다졸-1-일)프로필)(8-(노닐옥시)-8-옥소옥틸)아미노)옥탄오에이트(지질 54, 표 10a)*



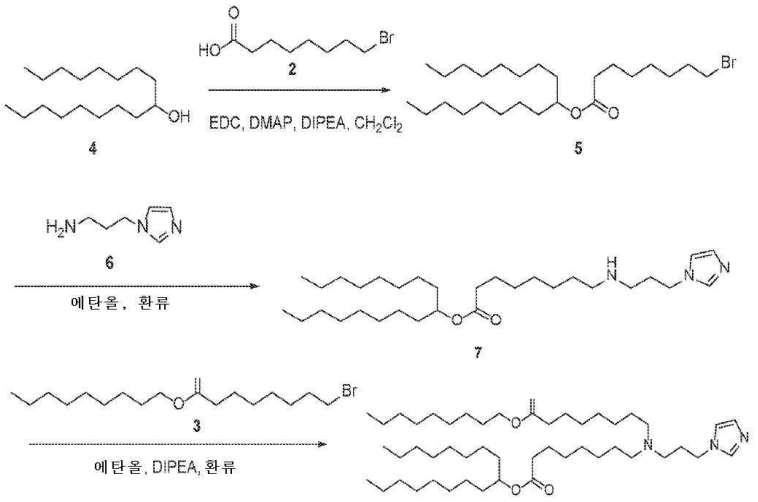
[2092]

[2093] 응축기가 연결된 100㎖ 둥근 바닥 플라스크에서, 헵타데칸-9-일 8-((3-(2-메틸-1*H*-이미다졸-1-일)프로필)아미노)옥탄오에이트(**7**)(665㎎, 1.279㎜ol) 및 노닐 8-브로모옥탄오에이트(**3**)(536㎎, 1.535㎜ol)를 에탄올(10㎖) 중에서 혼합하고, 그 다음 DIPEA(0.55㎖, 3.198㎜ol 첨가하였다. 반응 혼합물을 밤새 환류까지 가열시켰다. MS(APCI) 및 TLC(CH_2Cl_2 중의 10%MeOH+1%NH_4OH) 둘 다는 생성물 및 일부 미반응 출발 물질을 나타내었다. 혼합물을 실온으로 냉각시키고, 농축시켰다. 조 잔류물을 플래시 크로마토그래피(SiO_2: CH_2Cl_2= CH_2Cl_2 중의 100%에서 10%의 메탄올+1%NH_4OH)로 정제시키고, 무색 오일을 얻었다(170㎎, 17%).

[2094] ***38.5*** *헵타데칸-9-일 8-((3-(1H-이미다졸-1-일)프로필)(8-(노닐옥시)-8-옥소옥틸)아미노)옥탄오에이트(지질 53, 표 10a)의 합성*



[2095]



[2096]

[2097] 표 10a로부터의 지질 53을 상기 반응식에 따라서 합성한다. 반응 조건은 이미다졸 아민으로서 3-(1H-이미다졸-1-일)프로판-1-아민을 제외하고는 지질 54와 동일하다.

[2098] **실시예 37**

[2099] *원형 RNA를 갖는 지질 나노입자 제형*

[2100] 'NextGen' 혼합 챔버와 함께 Precision Nanosystems Ignite 장비를 사용하여 지질 나노입자(LNP)를 형성하였다. 표 10a로부터의 이온화 가능한 지질 26, DSPC, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사)을 16:1:4:1의 중량비 또는 62:4:33:1 몰비로 함유한 에탄올 상을 pH 5.2에서 원형 RNA 및 25mM 아세트산나트륨을 함유하는 수성상과 조합하였다. 3:1의 수성 대 에탄올 혼합 비율을 사용하였다. 그 다음 제형화된 LNP를 1L의 물에서 투석시키고, 18시간에 걸쳐 2회 교환하였다. 투석된 LNP를 0.2㎛ 필터를 사용하여 여과하였다. 생체내 투여 전에, LNP를 PBS에 희석시켰다. LNP 크기를 동적 광 산란에 의해 결정하였다. PBS(pH 7.4) 중의 20㎍/㎖ LNP 1㎖를 갖는 큐벳을 Malvern Panalytical Zetasizer Pro를 사용하여 Z-평균에 대해서 측정하였다. Z-평균 및 다분산도 지수를 기록하였다.

[2101] ***39.1** 표 10a로부터의 지질 26 및 27의 제형*

[2102] 'NextGen' 혼합 챔버와 함께 Precision Nanosystems Ignite 장비를 사용하여 지질 나노입자(LNP)를 형성하였다. 표 10a로부터의 이온화 가능한 지질 26 지질 27, DOPE, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사)을 16:1:4:1의 중량비 또는 62:4:33:1 몰비로 함유한 에탄올 상을 pH 5.2에서 원형 RNA 및 25mM 아세트산나트륨을 함유하는 수성상과 조합하였다. 3:1의 수성 대 에탄올 혼합 비율을 사용하였다. 그 다음 제형화된 LNP를 1L의 물에서 투석시키고, 18시간에 걸쳐 2회 교환하였다. 투석된 LNP를 0.2㎛ 필터를 사용하여 여과하였다. 생체내 투여 전에, LNP를 PBS에 희석시켰다. LNP 크기를 동적 광 산란에 의해 결정하였다. PBS(pH 7.4) 중의 20㎍/㎖ LNP 1㎖를 갖는 큐벳을 Malvern Panalytical Zetasizer Pro를 사용하여 Z-평균에 대해서 측정하였다. Z-평균 및 다분산도 지수를 기록하였다.

[2103] **39.2** *표 10a로부터의 지질 53 및 54의 제형*

[2104] 'NextGen' 혼합 챔버와 함께 Precision Nanosystems Ignite 장비를 사용하여 지질 나노입자(LNP)를 형성하였다. 표 10a로부터의 이온화 가능한 지질 53 또는 54, DOPE, 콜레스테롤 및 DSPE-PEG 2000(아반티 폴라 리피즈사)을 50:10:38.5:1.5의 몰비로 함유한 에탄올 상을 pH 5.2에서 원형 RNA 및 25mM 아세트산나트륨을 함유하는 수성상과 조합하였다. 3:1의 수성 대 에탄올 혼합 비율을 사용하였다. 그 다음 제형화된 LNP를 1L의 1x PBS에서 투석시키고, 18시간에 걸쳐 2회 교환하였다. 투석된 LNP를 0.2㎛ 필터를 사용하여 여과하였다. 생체내 투여 전에, LNP를 PBS에 희석시켰다. LNP 크기를 동적 광 산란에 의해 결정하였다. PBS(pH 7.4) 중의 20㎍/㎖ LNP 1㎖를 갖는 큐벳을 Malvern Panalytical Zetasizer Pro를 사용하여 Z-평균에 대해서 측정하였다. Z-평균 및 다분산도 지수를 기록하였다.

[2105] Malvern Panalytical Zetasizer Pro를 사용하여 LNP 제타 전위를 측정하였다. 물 중의 입자 용액 200㎕ 및 최종 입자 농도가 400㎍/㎖인 RNAse-무함유 증류수 800㎕를 함유하는 혼합물을 분석을 위해 zetasizer 모세관 셀에 로딩하였다.

[2106] RNA 캡슐화를 Ribogreen 검정을 사용하여 결정하였다. 나노입자 용액을 2㎍/㎖의 이론적인 oRNA 농도에서 트리스-에틸렌다이아민테트라아세트산(TE) 완충액에 희석시켰다. TE 완충액에 희석된 표준 oRNA 용액을 2㎍/㎖ 내지 0.125㎍/㎖ 범위로 만들었다. 입자 및 표준품을 모든 웰에 첨가하고, 두 번째 인큐베이션을 수행하였다(3분 동안 350rpm에서 37℃). SPECTRAmax® GEMINI XS 마이크로플레이트 분광형광계를 사용하여 형광을 측정하였다. 각각의 입자 용액에서 원형 RNA의 농도는 표준 곡선을 사용하여 계산하였다. 용해된 입자와 용해되지 않은 입자 사이에서 검출된 oRNA의 비율로부터 캡슐화 효율을 계산하였다.

[2107] [표 36a]

LNP의 특징규명

이온화 가능한 지질	크기 (nm)	PDI	캡슐화 효율 (%)	제타 데이터 전위 (mV)
22-S14	88	0.09	96	3.968
93-S14	119	0.02	96	-6.071
지질 26, 표 10a	86	0.08	92	-15.24

[2108]

[2109] [표 36b]

LNP의 특징규명

이온화 가능한 지질	Z-평균(nm)	PDI	RNA 포획(%)
22-S14	64	0.05	97
93-S14	74	0.04	95
지질 26, 표 10a	84	0.04	96

[2110]

[2111] **실시예 38**

[2112] *시험관내 분석*

[2113] 22 내지 25g 범위의 암컷 CD-1 또는 암컷 c57BL/6J 마우스에게 0.5mg/kg RNA를 정맥내 투여하였다. 주사 6시간 후, 마우스에게 15㎎/㎖ 농도의 D-루시페린 200㎕를 복강내 주사하였다. 주사 5분 후, 마우스를 아이소플루란을 사용하여 마취시키고, IVIS Spectrum In Vivo Imaging System(퍼킨 엘머사(Perkin Elmer)) 내부에 등쪽이 위로 향하게 하여 배치하였다. 지질 22-S14, 93-S14, 지질 26(표 10a)의 전신 총 IVIS 플럭스를 도 32A에 제시한다. 10분 주사 후, 마우스를 발광에 대해 스캔하였다. 마우스를 안락사시키고, 간, 비장, 신장, 폐 및 심장에서 발광을 스캔하기 위해 루시페린 주사 25분 이내에 장기를 적출하였다. 리빙 이미지(퍼킨 엘머사) 소프트웨어를 사용하여 영상(도 33A 및 도 33B, 도 34A 및 도 34B, 도 35A 및 도 35B)를 분석하였다. 관심 영역을 취하여 플럭스 및 평균 광도를 얻었고, 단백질 발현의 생체 분포에 대해 분석하였다(도 32A 및 도 32B).

[2114] 도 32A는 지질 22-S14 및 93-S14로 제조된 LNP와 비교하여 지질 26(표 10a) LNP를 갖는 루시퍼라제 oRNA로부터 관찰된 증가된 전신 총 플럭스를 도시한다. 도 32B는 발현을 개선하도록 설계된 상당한 구조적 변화에도 불구하고 지질 22-S14 및 93-S14에서 관찰된 원하는 비장 대 간 비율을 유지하면서 지질 26(표 10a)으로 증가된 전체 발현을 추가로 강조하는 조직의 생체외 IVIS 분석을 도시한다. 이러한 데이터는 이전에 보고된 지질과 비교하여 지질 26(표 10a)에 의해 제공되는 개선을 강조한다.

[2115] 표 10b의 지질 15 또는 표 10a의 지질 53 또는 54로 형성된 LNP에 캡슐화된 oRNA를 사용하여 상기에서 설명한 것과 유사한 분석을 수행하였다. 도 36A 내지 도 36C는 조직의 생체외 IVIS 분석을 나타내는데, 이는 각각 발현을 개선하도록 설계된 상당한 구조적 변화에도 불구하고 원하는 비장 대 간 비율을 유지하면서 지질 15, 53 및 54를 사용한 전체 발현을 강조한다. 도 36D는 PBS 대조군에 대한 결과를 나타낸다. 이들 데이터는 93-S14 및 22-S14와 같은 이전에 보고된 지질과 비교하여 표 10a의 지질 15, 53 및 54에 의해 제공되는 개선을 입증한다.

[2116] **실시예 39**

[2117] *루시퍼라제의 전달*

[2118] 인간 말초 혈액 단핵 세포(PBMC)(스템셀 테크놀로지즈사)에 파이어플라이 루시퍼라제*(f.luc)* 원형 RNA를 캡슐화하는 지질 나노입자(LNP)를 형질주입시키고, 루시퍼라제 발현에 대해 조사하였다. 2명의 상이한 공여자로부터의

PBMC를 파이어플라이 루시퍼라제를 암호화하는 원형 RNA를 포함하는 5개의 상이한 LNP 조성물(200ng)과 함께 시험관 내에서 37℃에서 RPMI, 2% 인간 혈청, IL-2(10ng/㎖) 및 50μM BME에서 인큐베이션시켰다. LNP 없이 배양된 PBMC를 음성 대조군으로 사용하였다. 24시간 후, 세포를 용해시키고 생물발광(프로메가 브라이트글로사(Promega BrightGlo))에 기초한 파이어플라이 루시퍼라제 발현에 대해 분석하였다.

[2119] 대표적인 데이터를 도 37A 및 도 37B에 제시하는데, 이것은 시험된 LNP가 원형 RNA를 1차 인간 면역 세포에 전달할 수 있어 단백질 발현이 발생함을 나타낸다.

[2120] **실시예 40**

[2121] *녹색 형광 단백질(GFP) 또는 키메라 항원 수용체(CAR)의 시험관내 전달*

[2122] 인간 PBMC(스템셀 테크놀로지즈사(스템셀 테크놀로지즈사))에 GFP를 캡슐화하는 LNP를 형질주입시키고, 유세포 분석법에 의해 조사하였다. 5명의 다른 공여자로부터의 PBMC(PBMC A 내지 E)를 GFP 또는 CD19-CAR(200ng)을 암호화하는 원형 RNA를 함유하는 하나의 LNP 조성물과 함께 시험관내에서 37℃에서, RPMI, 2% 인간 혈청, IL-2(10ng/㎖) 및 50μM BME에서 인큐베이션시켰다. LNP 없이 배양된 PBMC를 음성 대조군으로 사용하였다. LNP 인큐베이션 후 24, 48, 또는 72시간 후, 세포를 CD3, CD19, CD56, CD14, CD11b, CD45, fixable live dead 및 페이로드(GFP 또는 CD19-CAR)에 대해 분석하였다.

[2123] 대표적인 데이터를 도 38A 및 도 38B에 제시하는데, 이것은 시험된 LNP가 원형 RNA를 1차 인간 면역 세포에 전달할 수 있어 단백질 발현이 발생함을 나타낸다.

[2124] **실시예 41**

[2125] *다중 IRES 변이체는 시험관내에서 뮤린 CD19 CAR의 발현을 매개할 수 있다.*

[2126] 항-뮤린 CD19 CAR을 암호화하는 다중 원형 RNA 작제물은 고유한 IRES 서열을 함유하고, 이것에 1C1C7 세포를 지방형질주입시켰다. 지방형질주입 전에 1C1C7 세포를 완전한 RPMI에서 수 일 동안 확장시켰다. 세포가 적절한 수로 확장되면, 1C1C7 세포를 4개의 상이한 원형 RNA 작제물로 지방형질주입시켰다(Invitrogen RNAiMAX). 24시간 후, 1C1C7 세포를 His-태깅된 재조합 뮤린 CD19(시노 바이올로지컬사(Sino Biological)) 단백질과 함께 인큐베이션시킨 다음, 이차 항-His 항체로 염색하였다. 그 후, 세포를 유세포 분석법을 통해 분석하였다.

[2127] 대표적인 데이터를 도 39에 제시하는데, 이것은 제시된 바이러스(야생 등줄쥐(Apodemus agrarius) 피코르나바이러스, 카프린 코부바이러스, 파라보바이러스 및 살리바이러스)에서 유래한 IRES가 뮤린 T 세포에서 항-마우스 CD19 CAR의 발현을 유도할 수 있음을 나타낸다.

[2128] **실시예 42**

[2129] *뮤린 CD19 CAR은 시험관내에서 세포 사멸을 매개한다.*

[2130] 항-마우스 CD19 CAR을 암호화하는 원형 RNA를 뮤린 T 세포에 전기천공하여 CAR-매개 세포독성을 평가하였다. 전기천공을 위해, 써모피셔사의 Neon Transfection System을 사용하여 항-마우스 CD19 CAR을 암호화하는 원형 RNA로 T 세포를 전기천공한 다음 밤새 정치시켰다. 세포독성 검정을 위해 전기천공된 T 세포를 10% FBS, IL-2(10ng/㎖) 및 50uM BME를 함유하는 완전 RPMI에서 1:1 비율로 Fluc+ 표적 및 비표적 세포와 공배양하고, 37℃에서 밤새 인큐베이션시켰다. Fluc+ 표적 및 비-표적 세포의 용해를 검출하기 위해 공배양 24시간 후 루시퍼라제 검정 시스템(Promega Brightglo Luciferase System)을 사용하여 세포독성을 측정하였다. 표시된 값은 형질주입되지 않은 모의 신호를 기준으로 계산된다.

[2131] 대표적인 데이터를 도 40에 제시하는데, 이는 원형 RNA로부터 발현된 항-마우스 CD19 CAR이 시험관내에서 뮤린 T 세포에서 기능적임을 나타낸다.

[2132] **실시예 43**

[2133] *뮤린 CD19 CAR을 암호화하는 원형 RNA를 캡슐화한 지질을 사용한 B 세포의 기능적 결실*

[2134] 항-뮤린 CD19 CAR을 암호화하는 원형 RNA를 캡슐화하는 표 10b의 지질 15로 형성된 LNP를 C57BL/6J 마우스에 주사하였다. 대조군으로서, 파이어플라이 루시퍼라제(*f.Luc*)를 암호화하는 원형 RNA를 캡슐화하는 표 10b의 지질 15를 상이한 군의 마우스에게 주사하였다. 20 내지 25g 범위의 암컷 C57BL.6J에게 0.5mg/kg의 LNP를 격일로 5회 정맥 주사하였다. 주사 사이에, 혈액 채취물을 fixable live/dead, CD45, TCRvb, B220, CD11b 및 항-뮤린 CAR에 대한 유세포 분석을 통해 분석하였다. 마지막 주사 2일 후, 비장을 수거하고, 유세포 분석을 위해 가공하였

다. 비장 세포를 fixable live/dead, CD45, TCRvb, B220, CD11b, NK1.1, F4/80, CD11c 및 항-뮤린 CAR로 염색하였다. 항-뮤린 CD19 CAR LNP를 주사한 마우스의 데이터를 *f.Luc* LNP를 제공받은 마우스에 대해 정규화하였다.

[2135] 대표적인 데이터를 도 41A, 도 41B 및 도 41C에 제시하는데, 이것은 LNP와 함께 생체내 전달된 원형 oRNA로부터 발현된 항-마우스 CD19 CAR이 생체내 뮤린 T 세포에서 기능적임을 나타낸다.

[2136] **실시예 44**

[2137] *원형 RNA로부터 발현된 CD19 CAR은 mRNA로부터 발현된 것에 비해서 더 높은 수율 및 큰 세포독성 효과를 갖는다.*

[2138] N-말단에서부터 C-말단으로, FMC63-유래 scFv, CD8 막관통 도메인, 4-1BB 공자극 도메인, 및 CD3ζ 세포내 도메인을 포함하는 항-CD19 키메라 항원 수용체를 암호화하는 원형 RNA를, 표면 발현 및 CAR-매개 세포독성을 평가하기 위해 인간 말초 T 세포에 전기천공하였다. 비교를 위해, 본 실험에서 원형 RNA-전기천공된 T 세포를 mRNA-전기천공된 T 세포와 비교하였다. 전기천공을 위해, 공여자 인간 PBMC로부터의 상업적으로 입수 가능한 T 세포 단리 키트(밀테니이 바이오테크사(Miltenyi Biotec))를 사용하여 인간 PBMC로부터 CD3+ T 세포를 단리시켰다. 단리 후, T 세포를 항-CD3/항-CD28(스템셀 테크놀로지즈사(Stemcell Technologies))로 자극하고, 10% FBS, IL-2(10ng/㎖) 및 50uM BME를 함유하는 완전 RPMI에서 37℃에서 5일 동안 확장시켰다. 자극 5일 후, 써모피셔사의 Neon Transfection System을 사용하여 항-인간 CD19 CAR을 암호화하는 원형 RNA로 T 세포를 전기천공한 다음 밤새 정치시켰다. 세포독성 검정을 위해, 전기천공된 T 세포를 10% FBS, IL-2(10ng/㎖) 및 50uM BME를 함유하는 완전 RPMI에서 1:1 비율로 Fluc+ 표적 및 비표적 세포와 공배양하고, 37℃에서 밤새 인큐베이션시켰다. 공배양 24시간 후 루시퍼라제 검정 시스템(Promega Brightglo Luciferase System)을 사용하여 세포독성을 측정하여 Fluc+ 표적 및 비-표적 세포의 용해를 검출하였다 또한, 전기천공된 T 세포의 분취물을 취하고, 분석 당일 live dead fixable, CD3, CD45 및 키메라 항원 수용체(FMC63)에 대해 염색하였다.

[2139] 대표적인 데이터를 도 42 및 도 43에 제시한다. 도 42A 및 42B는 원형 RNA로부터 발현된 항-인간 CD19 CAR이 선형 mRNA로부터 발현된 항-인간 CD19 CAR보다 더 높은 수준으로 더 길게 발현된다는 것을 나타낸다. 도 43A 및 43B는 원형 RNA로부터 발현된 항-인간 CD19 CAR이 선형 mRNA로부터 발현된 항-인간 CD19 CAR에 비해 더 큰 세포독성 효과를 발휘한다는 것을 나타낸다.

[2140] **실시예 45**

[2141] *단일 원형 RNA로부터의 2개의 CAR의 기능적 발현*

[2142] 키메라 항원 수용체를 암호화하는 원형 RNA를 인간 말초 T 세포 내에 전기천공하여 표면 발현 및 CAR-매개 세포독성을 평가하였다. 본 연구의 목적은 2개의 CAR을 암호화하는 원형 RNA가 2A(P2A) 또는 IRES 서열로 확률적으로 발현될 수 있는지 평가하는 것이다. 전기천공을, 위해 CD3+ T 세포를 상업적으로 구입하고(셀레로사(Cellero)), 항-CD3/항-CD28(스템셀 테크놀로지즈사)로 자극하고, 10% FBS, IL-2(10ng/㎖) 및 50uM BME를 함유하는 완전 RPMI에서 37℃에서 5일 동안 확장시켰다. 자극 4일 후, 써모피셔사의 Neon Transfection System을 사용하여 항-인간 CD19 CAR, 항-인간 CD19 CAR-2A-항-인간 BCMA CAR 및 항-인간 CD19 CAR-IRES-항-인간 BCMA CAR을 암호화하는 원형 RNA로 T 세포를 전기천공한 다음, 밤새 정치시켰다. 세포독성 검정을 위해, 전기천공된 T 세포를 인간 CD19 또는 BCMA 항원을 발현하는 Fluc+ K562 세포와 함께 10% FBS, IL-2(10ng/㎖) 및 50uM BME를 함유하는 완전 RPMI에서 1:1 비율로 공배양하고, 37℃에서 밤새 인큐베이션시켰다. 공배양 24시간 후 루시퍼라제 검정 시스템(Promega BrightGlo Luciferase System)을 사용하여 세포독성을 측정하여 Fluc+ 표적 세포의 용해를 검출하였다

[2143] 대표적인 데이터를 도 44에 제시하는데, 이것은 2개의 CAR이 동일한 원형 RNA 작제물로부터 기능적으로 발현될 수 있고 세포독성 효과기 기능을 발휘할 수 있음을 나타낸다.

[2144] **실시예 46**

[2145] *Cre 리포터 마우스를 사용한 생체내 원형 RNA 형질주입*

[2146] Cre 재조합효소(Cre)를 암호화하는 원형 RNA를 이전에 기재된 바와 같이 지질 나노입자로 캡슐화한다. 6 내지 8주령의 암컷 B6.Cg-Gt(ROSA)26Sortm9(CAG-tdTomato)Hze/J(Ai9) 마우스에게 0.5mg/kg RNA의 지질 나노입자를 정맥내 투여하였다. 형광성 tdTomato 단백질은 Cre 재조합 시 Ai9 마우스에서 전사 및 번역되었는데, 이는 원형 RNA가 tdTomato+ 세포로 전달되고 번역되었음을 의미한다. 48시간 후, 마우스를 안락사시키고, 비장을 수거하고, 단일 세포 현탁액으로 처리하고, 유세포 분석법을 통한 면역표현형 분석을 위해 다양한 형광단-접합

항체로 염색하였다.

[2147] 도 45A는 표 10a의 지질 27 또는 지질 26 또는 표 10b의 지질 15로 지질로 형성된 LNP로 처리한 후 총 골수 세포(CD11b+), B 세포(CD19+) 및 T 세포(TCR-B+)를 비롯한 다양한 비장 면역 세포(CD45+, 생) 하위세트에서 tdTomato 발현 빈도와 함께 대표적인 FACS 플롯을 나타낸다. PBS가 주사된 Ai9 마우스는 배경 tdTomato 형광을 나타내었다. 도 45B는 tdTomato를 발현하는 골수 세포, B 세포 및 T 세포의 비율을 정량하는데(평균 + 표준 편차, n = 3), 이는 Cre 원형 RNA가 성공적으로 형질주입된 각각의 세포 집단의 비율과 동등하다. 표 10a의 지질 27 및 지질 26으로 제조된 LNP는 지질 93-S14에 비해서 상당히 더 높은 골수 및 T 세포 형질주입을 나타내는데, 이는 지질 구조 변형에 의해 부여된 개선을 강조한다.

[2148] 도 45C 표 10a로의 지질 27 및 지질 26을 갖는 tdTomato를 발현하는 추가 비장 면역 세포 집단의 비율(평균 + 표준편차, n = 3)을 도시하는데, 이것은 또한 NK 세포(NKp46+, TCR-B-), 전통적인 단핵구(CD11b+, Ly-6G-, Ly-6C_hi), 비전통적인 단핵구(CD11b+, Ly-6G-, Ly-6C_lo), 호중구(CD11b+, Ly-6G+) 및 수지상 세포(CD11c+, MHC-II+)를 포함한다. 이러한 실험은 표 10a의 지질 27 및 지질 26 및 표 10b의 지질 15로 제조된 LNP가 원형 RNA를 마우스의 많은 비장 면역 세포 하위세트에 전달하고, 그 세포에서 원형 RNA로부터 성공적인 단백질 발현을 유도하는 데 효과적이라는 것을 입증한다.

[2149] **실시예 47**

[2150] ***실시예 47A**: 면역-침묵 원형 RNA를 생산하기 위한 빌트-인 polyA 서열 및 친화성-정제*

[2151] PolyA 서열(20 내지 30nt)을 RNA 작제물(인트론 내에 빌트-인 polyA 서열을 갖는 전구체 RNA)의 5' 및 3' 단부에 삽입하였다. 대안적으로, 예를 들어, 추가 효소의 사용이 필요한 이. 콜라이. polyA 중합효소 또는 효모 polyA 중합효소를 사용하여 전구체 RNA 및 인트론을 전사 후에 폴리아데닐화할 수 있다.

[2152] 본 실시예에서 원형 RNA를 시험관내 전사(IVT)에 의해 원형화하였고, 스플라이싱 반응으로부터 polyA-태킹된 서열(유리 인트론 및 전구체 RNA 포함)을 선택적으로 제거하기 위해 상업적으로 입수 가능한 올리고-dT 수지 상에서 세척함으로써 친화성-정제시켰다. 상이한 비율(GMP:GTP = 8, 12.5 또는 13.75)의 구아노신 모노포스페이트(GMP) 및 구아노신 트라이포스페이트(GTP)를 함유하는 상업적인 IVT 키트(뉴 잉글랜드 바이오랩스사(New England Biolabs)) 또는 항문, 항문관형 IVT 믹스(오르나 쎄라퓨틱스사(Orna Therapeutics))로 IVT를 수행하였다. 일부 실시형태에서, 높은 GMP:GTP 비율의 GMP가 제1 뉴클레오타이드로서 우선적으로 포함될 수 있는데, 이것은 대부분의 모노포스페이트-캡핑된 전구체 RNA를 생성한다. 비교로서, 원형 RNA 생성물을 Xrn1, Rnase R 및 Dnase I로의 처리(효소 정제)에 의해 대안적으로 정제시켰다.

[2153] 이어서, 친화성 정제 또는 효소 정제 공정을 사용하여 제조된 원형 RNA의 면역원성을 평가하였다. 간략하면, 제조된 원형 RNA를 A549 세포에 형질주입시켰다. 24시간 후, 세포를 용해시키고 모의 샘플과 비교하여 인터페론 베타-1 유도를 qPCR에 의해 측정하였다. 삼인산화된 RNA인 3p-hpRNA를 양성 대조군으로 사용하였다.

[2154] 도 46B 및 도 46C는 polyA 서열이 스플라이싱 동안 제거되고 스플라이싱되지 않은 전구체 분자에 존재하는 요소에 포함될 때 스플라이싱 반응으로부터 비-원형 생성물을 제거하는 음성 선택 친화성 정제를 나타낸다. 도 46D는 시험된 IVT 조건 및 정제 방법으로 제조된 원형 RNA가 모두 면역정지임을 나타낸다. 이러한 결과는 음성 선택 친화성 정제가 원형 RNA 정제를 위한 효소 정제와 동등하거나 더 우수하며, 맞춤형 원형 RNA 합성 조건(IVT 조건)이 최대 면역정지를 달성하기 위해 GMP 과잉에 대한 의존도를 줄일 수 있음을 시사한다.

[2155] ***실시예 47B**: 원형 RNA 생산을 위한 전용 결합 부위 및 친화성-정제*

[2156] polyA 태그 대신 특별하게 설계 서열(DBS, 전용 결합 부위)를 포함할 수 있다.

[2157] polyA 태그 대신, 수지에 결합할 수 있는 특별히 설계된 상보적 올리고뉴클레오타이드와 같은 전용 결합 부위(DBS)를 사용하여 전구체 RNA 및 유리 인트론을 선택적으로 고갈시킬 수 있다. 본 실시예에서, DBS 서열(30nt)을 전구체 RNA의 5' 및 3' 단부에 삽입하였다. RNA를 전사시키고, 전사된 산물을 수지에 연결된 맞춤형 상보적 올리고뉴클레오타이드 상에서 세척하였다.

[2158] 도 47B 및 도 47C는 스플라이싱 동안 제거되는 요소에 설계된 DBS 서열을 포함하는 것이 음성 친화성 정제를 통해 스플라이싱 반응에서 스플라이싱되지 않은 전구체 RNA 및 유리 인트론 성분의 제거를 가능하게 한다는 것을 입증한다.

[2159] ***실시예 47C**: 디스트로핀을 암호화하는 원형 RNA의 생산*

[2160] RNA 전구체의 시험관내 전사에 이어 Xrn1, DNase 1 및 RNase R의 혼합물을 사용하여 효소 정제시켜 잔류하는 선형 성분을 분해시킴으로써 디스트로핀을 암호화하는 12kb 12,000nt 원형 RNA를 생산하였다. 도 48은 디스트로핀을 암호화하는 원형 RNA가 성공적으로 생산되었음을 나타낸다.

[2161] **실시예 48**

[2162] *3' 인트론 단편과 IRES 사이의 5' 스페이서는 원형 RNA 발현을 개선시킨다.*

[2163] 주카트 세포에서 3' 인트론 단편과 IRES 사이에 상이한 5' 스페이서를 갖는 정제된 circRNA의 발현 수준을 비교하였다. 간략하면, 60,000개 세포에 250ng의 각각의 RNA를 전기천공한 지 24시간 후에 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 측정하였다.

[2164] 추가로 주카트 세포에서 3' 인트론 단편과 IRES 사이에 상이한 5' 스페이서를 갖는 정제된 circRNA의 안정성을 비교하였다. 간략하면, 60,000개 세포에 250ng의 각각의 RNA를 전기천공한 지 2일 후에 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 측정하고, 제1일 발현에 정규화하였다.

[2165] 그 결과를 도 49A 및 도 49B에 제시하는데, 이는 스페이서를 추가하면 IRES 기능 및 추가된 스페이서의 서열 동일성 및 길이의 중요성을 향상시킬 수 있음을 나타낸다. 잠재적인 설명은 스페이서가 IRES 바로 앞에 추가되어 IRES가 인트론 단편과 같은 다른 구조화된 요소로부터 단리되어 접힐 수 있도록 함으로써 기능할 가능성이 있다는 것이다.

[2166] **실시예 49**

[2167] 본 실시예는 염소 코바이러스 IRES의 5' 또는 3' 단부로부터의 결실 스캐닝을 설명한다. IRES 경계는 일반적으로 양호하게 특징규명되어 있지 않아서, 경험적 분석이 필요한데, 본 실시예는 번역을 유도하는 데 필요한 핵심 기능 서열을 찾는 위해서 사용될 수 있다. 간략하면, 가우시아 루시퍼라제 암호 서열에 작동 가능하게 연결된 절단된 IRES 요소를 갖는 원형 RNA 작제물을 생성시켰다. 절단된 IRES 요소는 5' 또는 3' 단부로부터 제거된 제시된 길이의 뉴클레오타이드 서열을 가졌다. 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 RNA로의 1차 인간 T 세포의 전기천공 후 24시간 및 48시간 후에 측정하였다. 24시간 시점에서의 발현 수준에 대한 48시간 시점에서의 발현 수준의 비율로 발현의 안정성을 계산하였다.

[2168] 도 **50**에 도시된 바와 같이, IRES의 5' 단부로부터의 40개 초과의 뉴클레오타이드의 결실은 발현을 감소시키고, IRES 기능을 손상시켰다. 발현의 안정성은 IRES 요소의 절단에 의해 상대적으로 영향을 받지 않았지만, 발현 수준은 IRES의 3' 단부로부터의 141개 뉴클레오타이드의 결실에 의해 실질적으로 감소한 반면, 3' 단부로부터의 57 또는 122개 뉴클레오타이드의 결실은 IRES에 긍정적인 영향을 가졌다.

[2169] 또한 6-뉴클레오타이드 시작 전 서열의 결실이 루시퍼라제 리포터의 발현 수준을 감소시키는 것을 관찰하였다. 6-뉴클레오타이드 서열을 전통적인 코작 서열(GCCACC)로 교체하는 것은 유의미한 영향을 미치지 않았지만 적어도 발현은 유지시켰다.

[2170] **실시예 50**

[2171] 본 실시예는 염소 코브바이러스(CKV) IRES, 파라보바이러스 IRES, 등줄쥐 피코르나바이러스(AP) IRES, 코브바이러스 SZAL6 IRES, 크로히바이러스 B(CrVB) IRES, CVB3 IRES 및 SAFV IRES를 비롯한, 선택된 IRES 서열의 변형(예를 들어, 절단)을 설명한다. IRES 요소의 서열은 서열번호 348 내지 389에 제공된다. 간략하면, 가우시아 루시퍼라제 암호 서열에 작동 가능하게 연결된 절단된 IRES 요소를 갖는 원형 RNA 작제물을 생성시켰다. HepG2 세포에 원형 RNA를 형질주입시켰다. 상청액의 발광을 형질주입 24시간 및 48시간 후에 평가하였다. 24시간 시점에서의 발현 수준에 대한 48시간 시점에서의 발현 수준의 비율로 발현의 안정성을 계산하였다.

[2172] 도 **51**에 도시된 바와 같이, 절단은 IRES의 정체성에 따라 다양한 효과를 가졌으며, 이는 IRES들 간에 종종 상이한 번역에 사용되는 단백질 인자 및 개시 기전에 좌우될 수 있다. 5' 및 3' 결실은 예를 들어, CKV IRES의 맥락에서 효과적으로 조합될 수 있다. 일부 경우에 정규 코작 서열을 추가하면 발현이 상당히 개선되거나(SAFV에서와 같이, Full 대 Full+K) 발현이 감소한다(CKV에서와 같이, 5d40/3d122 대 5d40/3d122+K).

[2173] **실시예 51**

[2174] 본 실시예는 돌연변이 대체 번역 개시 부위를 비롯한, CK-739, AP-748 및 PV-743 IRES 서열의 변형을 설명한다. 간략하면, 가우시아 루시퍼라제 암호 서열에 작동 가능하게 연결된 변형된 IRES 요소를 갖는 원형 RNA 작제물을 생성시켰다. 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 RNA로의 1C1C7 세포의 전기천공 후 24시간

및 48시간 후에 측정하였다.

[2175] CUG는 가장 일반적으로 발견되는 대체 시작 부위이지만 다른 많은 부위도 특징규명되어 있다. 이러한 삼중항은 시작 코돈 이전에 IRES 스캐닝 트랙에 존재할 수 있으며 올바른 폴리펩타이드의 번역에 영향을 미칠 수 있다. 서열번호 378 내지 380에 제공된 IRES 서열을 사용하여 4개의 대체 시작 부위 돌연변이를 생성하였다. 도 52에 도시된 바와 같이, CK-739 IRES에서 대체 번역 개시 부위의 돌연변이는 올바른 폴리펩타이드의 번역에 영향을 미치며, 일부 경우에는 긍정적으로, 다른 경우에는 부정적으로 영향을 미쳤다. 모든 대체 번역 개시 부위의 돌연변이는 번역 수준을 감소시켰다.

[2176] 시작 코돈 이전의 6개 뉴클레오타이드인 대체 코작 서열도 발현 수준에 영향을 미칠 수 있다. 시작 코돈의 상류의 6개 염기 서열은 각각 CK-739 IRES에서 gTcacG, aaagtc, gTcacG, gtcatg, gcaaac, acaacc였으며 "6nt Pre-Start" 군에서는 샘플 번호 1 내지 5였다. 도 52에 도시된 바와 같이, 시작 코돈 이전의 특정 6-뉴클레오타이드 서열의 치환이 번역에 영향을 미쳤다.

[2177] 또한 AP-748 및 PV-743 IRES 서열에서 5' 및 3' 말단 결실이 발현을 감소시키는 것을 관찰하였다. 그러나, 긴 스캐닝 트랙을 갖는 CK-739 IRES에서, 번역은 스캐닝 트랙 내의 결실에 의해 상대적으로 영향을 받지 않았다.

[2178] **실시예 52**

[2179] 본 실시예는 5' 및/또는 3' 미번역 영역(UTR)을 삽입하고 IRES 혼성체를 생성시킴으로써 선택된 IRES 서열의 변형을 설명한다. 간략하면, 가우시아 루시퍼라제 암호 서열에 작동 가능하게 연결된 변형된 IRES 요소를 갖는 원형 RNA 작제물을 생성시켰다. 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 RNA로의 HepG2 세포의 전기천공 후 24시간 및 48시간 후에 측정하였다.

[2180] UTR이 삽입된 IRES 서열을 서열번호 390 내지 401에 제공한다. 도 53에 도시된 바와 같이, IRES의 3' 단부 바로 뒤와 시작 코돈 앞에 5' UTR을 삽입하면 염소 코브바이러스(CK) IRES에서 번역이 약간 증가했지만, 일부 경우에 살리바이러스 SZ1 IRES에서 번역이 중단되었다. 정지 카세트 바로 뒤에 3' UTR을 삽입하면 두 IRES 서열에 영향을 미치지 않았다.

[2181] 혼성 CK IRES 서열을 서열번호 390 내지 401에 제공한다. CK IRES를 염기로 사용하고, CK IRES의 특정 영역을 다른 IRES 서열, 예를 들어, SZ1 및 AV(아이치바이러스)에서 유사하게 보이는 구조로 대체하였다. 도 53에 도시된 바와 같이, 특정 혼성 합성 IRES 서열은 기능적이었는데, 이는 혼성 IRES가 유사한 예측 구조를 나타내는 구별되는 IRES 서열의 부분을 사용하여 작제될 수 있지만 이러한 구조를 결실시키면 IRES 기능을 완전히 제거할 수 있음을 나타낸다.

[2182] **실시예 53**

[2183] 본 실시예는 정지 코돈 또는 카세트 변이체를 도입함으로써 원형 RNA를 변형시키는 방법을 설명한다. 간략하면, 가우시아 루시퍼라제 암호 서열에 작동 가능하게 연결된 IRES 요소에 이어 가변 정지 코돈 카세트를 사용하여 원형 RNA 작제물을 생성시켰는데, 이는 각각의 프레임에 정지 코돈을 포함하고, 가우시아 루시퍼라제 암호 서열의 판독 프레임에 2개의 정지 코돈을 포함하였다. 1C1C7 세포에 원형 RNA를 형질주입시켰다. 상청액의 발광을 형질주입 24시간 및 48시간 후에 평가하였다.

[2184] 정지 코돈 카세트의 서열을 서열번호 406 내지 412에 제시한다. 도 54에 도시된 바와 같이, 특정 정지 코돈 카세트는 발현 안정성에 거의 영향을 미치지 않았지만 발현 수준을 개선시켰다. 특히, 2개의 프레임 1(가우시아 루시퍼라제 암호 서열의 판독 프레임) 정지 코돈을 갖는 정지 카세트(첫 번째는 TAA이고 뒤이어 프레임 2 정지 코돈 및 프레임 3 정지 코돈)는 기능적 번역을 촉진하는 데 효과적이다.

[2185] **실시예 54**

[2186] 본 실시예는 5' UTR 변이체를 삽입하여 원형 RNA의 변형을 설명한다. 간략하면, IRES의 3' 단부와 시작 코돈 사이에 삽입된 5' UTR 변이체가 있는 IRES 요소로 원형 RNA 작제물을 생성시켰고, IRES는 가우시아 루시퍼라제 암호 서열에 작동 가능하게 연결되어 있다. 1C1C7 세포에 원형 RNA를 형질주입시켰다. 상청액의 발광을 형질주입 24시간 및 48시간 후에 평가하였다.

[2187] 5' UTR 변이체의 서열을 서열번호 402 내지 405에 제시한다. 도 55에 도시된 바와 같이, 정규 코작 서열(UTR4)을 갖는 CK IRES는 36개 뉴클레오타이드의 비구조화/낮은 GC 스페이서 서열이 추가될 때(UTR2) 보다 효과적이었는데, 이는 GC-풍부 코작 서열이 코어 IRES 폴딩을 방해할 수 있음을 시사한다. 코작 서열과 함께 더 높은 GC/

구조화된 스페이서를 사용하는 것은 동일한 이점(UTR3)을 나타내지 않았는데, 이는 가능하게는 스페이서 자체에 의한 IRES 접힘의 간섭 때문일 수 있다. 코작 서열을 gTcacG(UTR1)로 돌연변이시키는 것은 스페이서에 대한 필요성 없이 코작+스페이서 대안과 동일한 수준으로 번역을 향상시켰다.

[2188] **실시예 55**

[2189] 본 실시예는 발현 수준에 대한 원형 RNA의 miRNA 표적 부위의 영향을 설명한다. 간략하면, 인간 에리트로포이에틴(hEPO) 암호 서열에 작동 가능하게 연결된 IRES 요소를 사용하여 원형 RNA 작제물을 생성시켰고, 여기서 2개의 직렬 miR-122 표적 부위를 작제물에 삽입하였다. miR-122-발현 Huh7 세포에 원형 RNA를 형질주입시켰다. 상청액에서 hEPO 발현을 샌드위치 ELISA에 의한 형질주입 24시간 및 48시간 후에 평가하였다.

[2190] 도 56에 도시된 바와 같이, miR-122 표적 부위가 원형 RNA에 삽입된 hEPO 발현 수준은 제거되었다. 이 결과는 원형 RNA의 발현이 miRNA에 의해 조절될 수 있다는 것을 입증한다. 이와 같이, 재조합 단백질의 발현이 바람직하지 않은 세포 유형에서 발현되는 miRNA의 표적 부위를 혼입함으로써 세포 유형 또는 조직 특이적 발현을 달성할 수 있다.

[2191] **실시예 56**

[2192] ***실시예 56A**: 2A 부위를 함유하는 하나의 circRNA로부터의 2개의 단백질의 발현 및 기능적 RNA 안정성.*

[2193] 아나베나 인트론/엑손 영역, 다양한 IRES, 가우시아 루시퍼라제를 암호화하는 제1 발현 서열, 2A 자가-절단 펩타이드, 및 GFP 또는 EGFP를 암호화하는 제2 발현 서열을 포함하는 작제물을 원형화하였다. 20,000 내지 40,000개의 293 세포에 40ng의 circRNA를 형질주입한다. 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 각각의 시점에서 완전 배지 교체로 형질주입 후 24시간마다 측정한다. 발현된 GFP 또는 EGFP의 형광을 형질주입 24시간 후에 측정한다.

[2194] ***실시예 56B**: 2개의 IRES 서열을 함유하는 하나의 circRNA로부터의 2개의 단백질 발현 및 기능적 RNA 안정성.*

[2195] 아나베나 인트론/엑손 영역, 제1 IRES, 가우시아 루시퍼라제를 암호화하는 제1 발현 서열, 선택적으로 스페이서, 제2 IRES, 및 GFP 또는 EGFP를 암호화하는 제2 발현 서열을 포함하는 작제물(제1 및/또는 제2 IRES는 작제물에 따라 다름)을 원형화하였다. 293개의 세포에 1㎍의 각각의 원형화 반응을 전기천공한다. 상청액에서 분비된 가우시아 루시퍼라제 및 GFP 또는 EGFP로부터의 발광/형광을 전기천공 후 24시간에 측정한다. 전기천공된 293 세포의 각각의 라운드에서 IRES 작제물의 기능적 안정성을 3일에 걸쳐 측정하였다. 세포에 1㎍의 각각의 원형화 반응을 전기천공시키고, 그 다음 완전 배지 교체 후 상청액에서 분비된 가우시아 루시퍼라제 및 GFP 또는 EGFP로부터의 발광/형광을 24시간마다 측정한다.

[2196] **실시예 57**

[2197] *절단 부위에 의해 분리된 2개의 발현 서열을 함유하는 원형 및 선형 RNA의 발현 및 기능적 안정성.*

[2198] 아나베나 인트론/엑손 영역, 가우시아 루시퍼라제를 암호화하는 제1 발현 서열, 2A 자가-절단 펩타이드, GFP 또는 EGFP를 암호화하는 발현 서열을 포함하는 작제물을 원형화한다. 가우시아 루시퍼라제를 암호화하는 제1 발현 서열, 2A 자가-절단 펩타이드, EGFP를 암호화하는 제2 발현 서열, 약 150nt polyA 테일을 암호화하고, 100%의 우리딘을 1-메틸 슈도우리딘으로 대체하도록 변형된 mRNA를 생성시킨다. 원형 및 변형된 mRNA의 발현을 293개 세포에서 측정한다. 세포에 1㎍의 각각의 RNA 종을 전기천공시킨 후 24시간 후에 상청액에서 분비된 가우시아 루시퍼라제 및 GFP 또는 EGFP로부터의 발광/형광을 측정한다. 작제물 기능적 안정성을 비교하기 위해서, 상청액에서 분비된 가우시아 루시퍼라제 및 GFP 또는 EGFP로부터의 발광/형광을 3일에 걸쳐 전기천공시킨 후 24시간마다 측정한다.

[2199] **실시예 58**

[2200] ***실시예 58A**: 단일 발현 작제물을 사용한 발현과 이중 발현의 기능적 안정성의 비교*

[2201] 아나베나 인트론/엑손 영역, IRES, 가우시아 루시퍼라제 발현 서열, 2A 자가-절단 펩타이드 및 GFP 또는 EGFP 발현 서열을 포함하는 작제물을 원형화한다. 단일 발현 서열(가우시아 루시퍼라제 또는 GFP 또는 EGFP를 암호화함)을 함유하는 작제물을 또한 원형화한다. 단일 발현 서열(가우시아 루시퍼라제 또는 GFP 또는 EGFP를 암호화함)을 함유하는 선형 작제물을 또한 생성시킨다. 293 세포에 1㎍의 각각의 원형화 반응 또는 등가량의 선형 RNA를 전기천공시킨다. 상청액에서 분비된 가우시아 루시퍼라제 및 GFP 또는 EGFP 양성성으로부터의 형광을 전기천공 후 24시간에 측정한다.

[2202] 전기천공된 293 세포의 각각의 라운드에서 작제물의 기능적 안정성을 3일에 걸쳐 측정하였다. 세포에 1㎍의 각각의 원형화 반응 또는 등가의 mRNA를 전기천공시키고, 그 다음 완전 배지 교체 후 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 24시간마다 측정한다.

[2203] ***실시예 58B**: 단일 발현 작제물을 사용한 이중 발현의 발현 및 기능적 안정성 비교*

[2204] 아나베나 인트론/엑손 영역, 제 IRES, 가우시아 루시퍼라제 발현 서열, 제2 IRES 및 GFP 또는 EGFP 발현 서열을 포함하는 작제물을 원형화한다. 단일 발현 서열(가우시아 루시퍼라제 또는 GFP 또는 EGFP를 암호화함)을 함유하는 작제물을 또한 원형화한다. 단일 발현 서열(가우시아 루시퍼라제 또는 GFP 또는 EGFP를 암호화함)을 함유하는 선형 작제물을 또한 생성시킨다. 293 세포에 1㎍의 각각의 원형화 반응 또는 등가량의 선형 RNA를 전기천공시킨다. 상청액에서 분비된 가우시아 루시퍼라제 및 GFP 또는 EGFP 양성성으로부터의 형광을 전기천공 후 24시간에 측정한다.

[2205] 전기천공된 293 세포의 각각의 라운드에서 작제물의 기능적 안정성을 3일에 걸쳐 측정하였다. 세포에 1㎍의 각각의 원형화 반응 또는 등가의 mRNA를 전기천공시키고, 그 다음 완전 배지 교체 후 상청액에서 분비된 가우시아 루시퍼라제로부터의 발광을 24시간마다 측정한다.

[2206] **실시예 59**

[2207] ***실시예 59A**: IRES 및 절단 부위를 사용하여 발현된 TCR 페이로드 간의 기능적 상호작용.*

[2208] 아나베나 인트론/엑손 영역, IRES, TCR 알파쇄 발현 서열, 2A 자가 절단 펩타이드, 및 TCR 베타쇄 발현 서열을 포함하는 작제물을 원형화시킨다.

[2209] 인간 1차 CD3+ T 세포에 circRNA를 전기천공시키고, TCR 표적, GFP 및 파이어플라이 루시퍼라제를 안정적으로 발현하는 세포와 함께 24시간 동안 공배양한다. 인간 1차 CD3+ T 세포를 모의 전기천공시키고, 대조군으로서 공배양한다. 표적 세포의 특이적 용해의 정량화를 파이어플라이 발광의 검출에 의해 결정한다. 특이적 용해 %는 1-[TCR 조건 발광]/[모의 조건 발광]으로 정의된다.

[2210] ***실시예 59B**: 2개의 IRES를 사용하여 발현된 TCR 페이로드 간의 기능적 상호 작용.*

[2211] 아나베나 인트론/엑손 영역, 제1 IRES, TCR 알파쇄 발현 서열, 선택적으로 스페이서, 제2 IRES, 및 TCR 베타쇄 발현 서열을 포함하는 작제물을 원형화한다.

[2212] 인간 1차 CD3+ T 세포에 circRNA를 전기천공시키고, TCR 표적, GFP 및 파이어플라이 루시퍼라제를 안정적으로 발현하는 세포와 함께 24시간 동안 공배양한다. 인간 1차 CD3+ T 세포를 모의 전기천공시키고, 대조군으로서 공배양한다. 표적 세포의 특이적 용해의 정량화를 파이어플라이 발광의 검출에 의해 결정한다. 특이적 용해 %는 1-[TCR 조건 발광]/[모의 조건 발광]으로 정의된다.

[2213] **실시예 60**

[2214] ***실시예 60A**: IRES 및 절단 부위를 사용하여 발현된 CAR 페이로드.*

[2215] 아나베나 인트론/엑손 영역, IRES, 항-CD19 CAR 발현 서열, 2A 자가 절단 펩타이드, 및 항-C20 CAR 발현 서열을 포함하는 작제물을 원형화시킨다.

[2216] 인간 1차 CD3+ T 세포에 circRNA를 전기천공시키고, 두 CAR 표적, GFP 및 파이어플라이 루시퍼라제를 안정적으로 발현하는 세포와 함께 24시간 동안 공배양한다. 인간 1차 CD3+ T 세포를 모의 전기천공시키고, 대조군으로서 공배양한다. 표적 세포의 특이적 용해의 정량화를 파이어플라이 발광의 검출에 의해 결정한다. 특이적 용해 %는 1-[TCR 조건 발광]/[모의 조건 발광]으로 정의된다.

[2217] ***실시예 60B**: 2개의 IRES를 사용하여 발현된 TCR 페이로드*

[2218] 아나베나 인트론/엑손 영역, 제1 IRES, 항-CD19 CAR 알파쇄 발현 서열, 선택적으로 스페이서, 제2 IRES, 및 항-CD20 CAR 베타쇄 발현 서열을 포함하는 작제물을 원형시킨다.

[2219] 인간 1차 CD3+ T 세포에 circRNA를 전기천공시키고, 두 CAR 표적, GFP 및 파이어플라이 루시퍼라제를 안정적으로 발현하는 세포와 함께 24시간 동안 공배양한다. 인간 1차 CD3+ T 세포를 모의 전기천공시키고, 대조군으로서 공배양한다. 표적 세포의 특이적 용해의 정량화를 파이어플라이 발광의 검출에 의해 결정한다. 특이적 용해 %는 1-[TCR 조건 발광]/[모의 조건 발광]으로 정의된다.

[2220] **실시예 61**

[2221] *항-CD19 CAR을 함유하는 LNP 및 원형 RNA 작제물은 생체내에서 혈액 및 비장의 B 세포를 감소시킨다.*

[2222] 항-CD19 CAR 발현을 암호화하는 원형 RNA 작제물을 상기 기재된 바와 같이 지질 나노입자 내에 캡슐화시켰다. 비교를 위해, 루시퍼라제 발현을 암호화하는 원형 RNA를 별도의 지질 나노입자 내에 캡슐화시켰다.

[2223] 6주 내지 8주령의 C57BL/6 마우스에 LNP 용액 중 하나를 격일로 주사하여 각각의 마우스 내에서 총 4회의 LNP를 주사하였다. 마지막 LNP 주사 24시간 후, 마우스의 비장 및 혈액을 수거하고, 염색하고, 유세포 분석을 통해 분석하였다. **도 57A** 및 **도 57B**에 도시된 바와 같이, 항-CD19 CAR을 암호화하는 LNP-원형 RNA 작제물을 함유하는 마우스는 루시퍼라제를 암호화하는 LNP-원형 RNA로 처리된 마우스와 비교하여 말초 혈액 및 비장에서 CD19+ B 세포의 통계적으로 유의한 감소를 초래하였다.

[2224] **실시예 62**

[2225] *CAR을 암호화하는 원형 RNA 내에 함유된 IRES 서열은 CAR 발현 및 T-세포의 세포독성을 개선시킨다.*

[2226] 활성화된 뮤린 T-세포에 고유한 IRES 및 뮤린 항-CD19 1D3ζ CAR 발현 서열을 함유하는 200ng의 원형 RNA 작제물을 전기천공시켰다. 이러한 작제물에 함유된 IRES는 염소 코브바이러스, 등줄쥐 피코르나바이러스, 파라보바이러스 또는 살리바이러스로부터 전체 또는 부분적으로 유래되었다. 염소 코브바이러스 유래 IRES를 추가로 코돈 최적화하였다. 대조군으로서, IRES가 없는 야생형 제타 마우스 CAR을 함유하는 원형 RNA를 비교에 사용하였다. 표면 발현을 평가하기 위해 전기천공 24시간 후 T-세포를 CD-19 CAR에 대해 염색한 다음 A20 Fluc 표적 세포와 공배양하였다. 이어서, T-세포를 표적 세포와 공배양한 후 24시간 후에 Fluc+ A20 세포의 세포독성 사멸에 대해 검정을 평가하였다.

[2227] **도 58A**, **도 58B**, **도 58C** 및 **도 59**에 도시된 바와 같이, 고유한 IRES는 T-세포가 CAR 단백질을 발현하는 빈도 및 세포 표면 상에서의 CAR 발현 수준을 증가시킬 수 있었다. CAR 단백질의 발현 빈도가 증가하고 세포 표면 상에서 CAR 발현 수준이 증가하면 항종양 반응이 개선된다.

[2228] **실시예 63**

[2229] *1차 인간 T 세포에서 원형 RNA 작제물로부터 발현된 사이토솔 및 표면 단백질.*

[2230] 원형 RNA 작제물은 형광 사이토솔 리포터 또는 표면 항원 리포터를 암호화하는 서열을 함유하였다. 형광 리포터는 녹색 형광성 단백질, mCitrine, mWasabi, Tsapphire를 포함하였다. 표면 리포터는 CD52 및 $Thy1.1^{bio}$를 포함하였다. 1차 인간 T-세포를 항-CD3/항-CD28 항체로 활성화하고, 리포터 서열을 함유하는 원형 RNA의 활성화 6일 후에 전기천공하였다. T-세포를 수거하고, 전기천공 24시간 후에 유세포 분석을 통해 분석하였다. 표면 항원을 상업적으로 입수 가능한 항체(예를 들어, 바이오레전드사, 밀테니이사 및 비디사)로 염색하였다.

[2231] **도 60A** 및 **도 60B**에 도시된 바와 같이, 사이토솔 및 표면 단백질은 1차 인간 T-세포에서 단백질을 암호화하는 원형 RNA로부터 발현될 수 있다.

[2232] **실시예 64**

[2233] *고유한 IRES 서열을 함유하는 원형 RNA는 선형 mRNA보다 번역 발현이 개선되었다.*

[2234] 원형 RNA 작제물은 파이어플라이 루시퍼라제(FLuc)에 대한 발현 서열과 함께 고유한 IRES를 함유하였다.

[2235] 2명의 공여자로부터의 인간 T-세포를 농축시키고, 항-CD3/항-CD28 항체로 자극하였다. 수 일의 증식 후, 활성화된 T 세포를 수거하고, FLuc 페이로드를 발현하는 mRNA 또는 원형 RNA의 동일한 몰로 전기천공하였다. 염소 코브바이러스, 등줄쥐 피코르나바이러스 및 파라보바이러스로부터 유래된 것을 비롯한 다양한 IRES 서열을 연구하여 7일 동안 페이로드 발현의 발현 수준 및 지속성을 평가하였다. 7일에 걸쳐, T-세포를 Promega Brightglo로 용해시켜 생물발광을 평가하였다.

[2236] **도 61C, 도 61D, 도 61E, 도 61F 및 도 61G**에 도시된 바와 같이 원형 RNA 내에 IRES가 존재하면 사이토솔 단백질의 번역 및 발현이 수십 배 증가할 수 있고, 이는 선형 mRNA에 비해 발현이 향상될 수 있다. 이것은 여러 인간 T 세포 공여자에서 일관되게 발견되었다.

[2237] **실시예 65**

[2238] ***실시예 65A:*** *항-CD19를 암호화하는 LNP-원형 RNA는 K562 세포의 인간 T-세포 사멸을 매개한다.*

[2239] 원형 RNA 작제물은 항-CD19 항체를 암호화하는 서열을 함유하였다. 그 다음 원형 RNA 작제물을 지질 나노입자(LNP) 내에 캡슐화시켰다.

[2240] 인간 T-세포를 항-CD3/항-CD28로 자극하였고, 최대 6까지 증식되도록 두었다. 제6일에, LNP-원형 RNA 및 ApoE3(1㎍/㎖)를 T-세포와 공배양하여 형질주입을 매개하였다. 24시간 후, Fluc+ K562 세포에 항-CD19 항체를 암호화하는 원형 RNA 200ng를 전기천공시키고, 그 후에 제7일에 공배양하였다. 공배양 후 48시간에, Fluc 발현을 통한 K562 세포의 세포독성 사멸 및 CAR 발현에 검정을 평가하였다.

[2241] **도 62A 및 도 62B**에 도시된 바와 같이, 시험관내에서 T-세포로의 CAR의 LNP-매개 전달로부터의 항-CD19 CAR의 T 세포 발현 및 조작된 K562 세포에서 특이적 항원 의존 방식으로 종양 세포를 용해하는 능력이 존재한다.

[2242] ***실시예 65B:*** *항-BCMA 항체를 암호화하는 LNP-원형 RNA는 K562 세포의 인간 T-세포 사멸을 매개한다.*

[2243] 원형 RNA 작제물은 항-BCMA 항체를 암호화하는 서열을 함유하였다. 그런 다음 원형 RNA 작제물을 지질 나노입자(LNP) 내에 캡슐화시켰다.

[2244] 인간 T-세포를 항-CD3/항-CD28로 자극하였고, 최대 6까지 증식되도록 두었다. 제6일에, LNP-원형 RNA 및 ApoE3(1㎍/㎖)를 T-세포와 공배양하여 형질주입을 매개하였다. 24시간 후, Fluc+ K562 세포에 항-BCMA 항체를 암호화하는 원형 RNA 200ng를 전기천공시키고, 그 후에 제7일에 공배양하였다. 공배양 후 48시간에, Fluc 발현을 통한 K562 세포의 세포독성 사멸 및 CAR 발현에 검정을 평가하였다.

[2245] **도 62B**에 도시된 바와 같이, 시험관내에서 T-세포로의 CAR의 LNP-매개 전달로부터의 BCMA CAR의 T 세포 발현 및 조작된 K562 세포에서 특이적 항원 의존 방식으로 종양 세포를 용해하는 능력이 존재한다.

[2246] **실시예 66**

[2247] *항-CD19 CAR T-세포는 시험관내에서 항종양 활성을 나타낸다.*

[2248] 인간 T-세포를 항-CD3/항-CD28로 활성화하고, 200ng의 항-CD19 CAR-발현 원형 RNA로 1회 전기천공하였다. 전기천공된 T-세포를 FLuc+ Nalm6 표적 세포 및 비-표적 Fluc+K562 세포와 공배양하여 CAR-매개 사멸을 평가하였다. 공배양 후 24시간 후, T-세포를 용해시키고, 표적 및 비-표적 세포에 의한 잔류 FLuc 발현에 대해 조사하여 총 8일에 걸친 발현 및 발현의 안정성을 평가하였다.

[2249] **도 63A** 및 **도 63B**에 나타낸 바와 같이, T-세포는 특이적 항원-의존적 방식으로 원형 RNA CAR 작제물을 발현한다. 결과는 또한 CAR을 암호화하는 선형 mRNA와 비교하여 CAR을 암호화하는 원형 RNA의 개선된 세포독성 및 기능적 표면 수용체의 전달을 나타낸다.

[2250] **실시예 67**

[2251] *ApoE3로 매개된 원형 RNA의 효과적인 LNP 형질주입*

[2252] 인간 T-세포를 항-CD3/항-CD28로 자극하고 최대 6일까지 증식되도록 두었다. 제6일에, 지질 나노입자(LNP) 및 ApoE3(1㎍/㎖)가 있거나 없이 녹색 형광 단백질 용액을 발현하는 원형 RNA를 T-세포와 공배양하였다. 24시간 후, T-세포를 생/사 T-세포에 대해 염색하고, 생 T-세포를 유세포 분석기에서 GFP 발현에 대해 분석하였다.

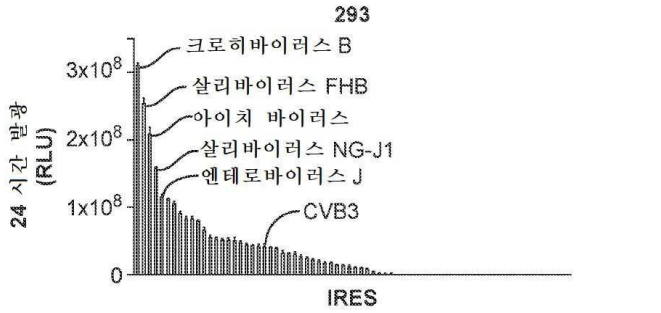
[2253] **도 64A, 도 64B, 도 61C, 도 61D** 및 **도 64E**에 도시된 바와 같이, 효율적인 LNP 형질주입은 활성화된 T-세포의 ApoE3에 의해 매개될 수 있으며, 이어서 상당한 페이로드 발현이 이어진다. 이 결과는 다수의 공여자에 걸쳐서 나타났다.

[2254] **참조에 의한 포함**

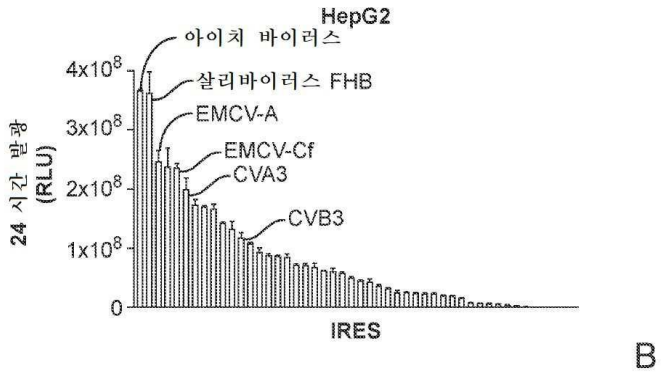
[2255] 본 명세서에 언급된 모든 간행물, 특허 및 특허 출원은 각각의 개별 간행물, 특허 또는 특허 출원이 본 명세서에 참조에 의해 포함되는 것으로 구체적이고 개별적으로 표시된 것과 동일한 정도로 참조에 의해 본 명세서에 포함된다.

도면

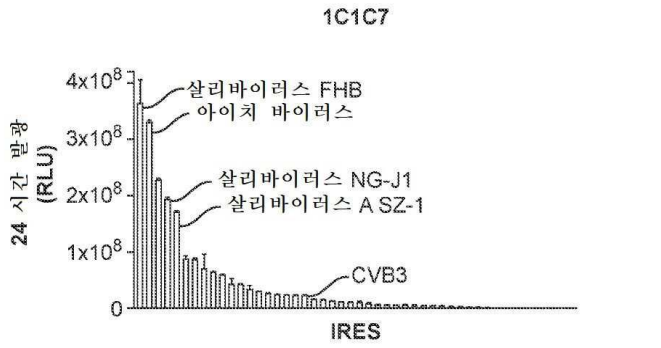
도면1ac


293
크로히바이러스 B
살리바이러스 FHB
아이치 바이러스
살리바이러스 NG-J1
엔테로바이러스 J
CVB3
3x10^8
2x10^8
1x10^8
0
24 시간 발광 (RLU)
IRES

A

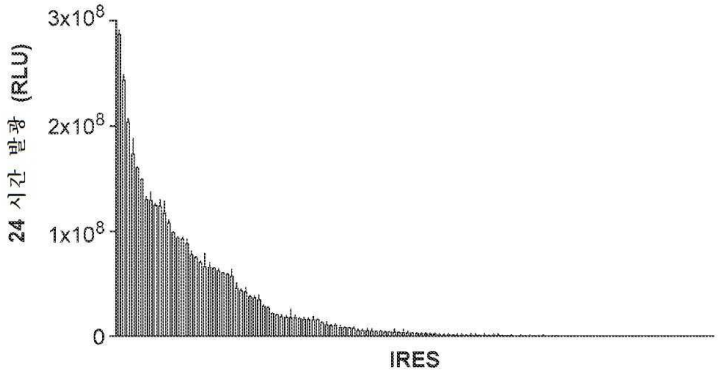

HepG2
아이치 바이러스
살리바이러스 FHB
EMCV-A
EMCV-Cf
CVA3
CVB3
4x10^8
3x10^8
2x10^8
1x10^8
0
24 시간 발광 (RLU)
IRES

B

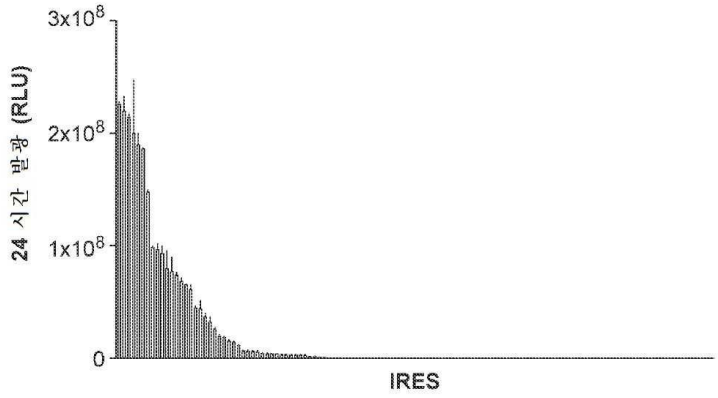

1C1C7
살리바이러스 FHB
아이치 바이러스
살리바이러스 NG-J1
살리바이러스 A SZ-1
CVB3
4x10^8
3x10^8
2x10^8
1x10^8
0
24 시간 발광 (RLU)
IRES

C

도면1de

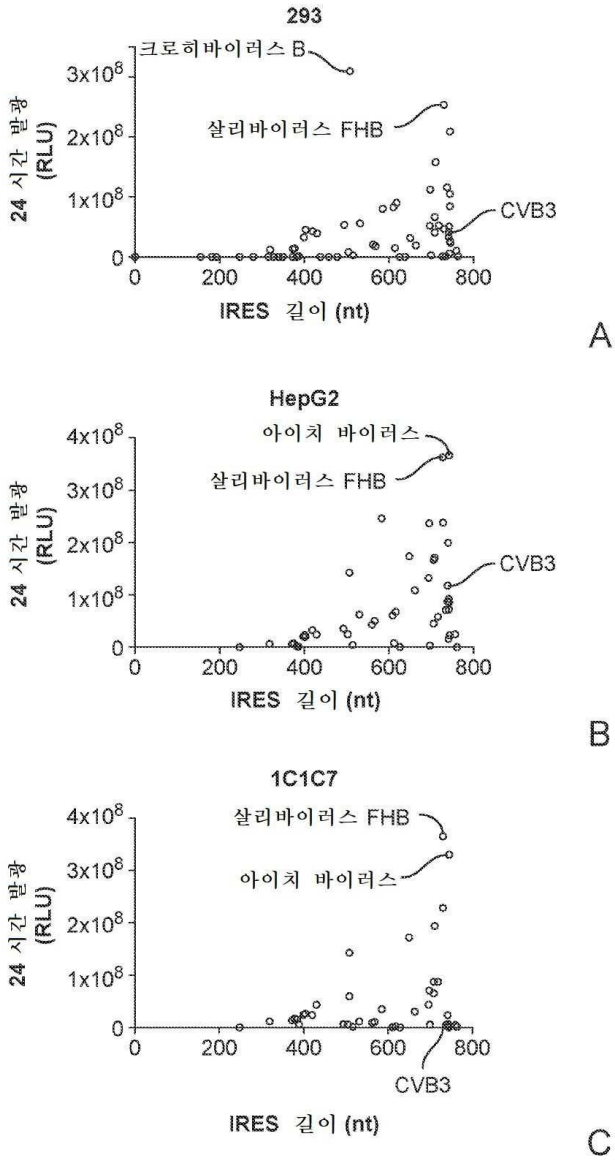

3x10^8
2x10^8
1x10^8
0
24 시간 발광 (RLU)
IRES

D

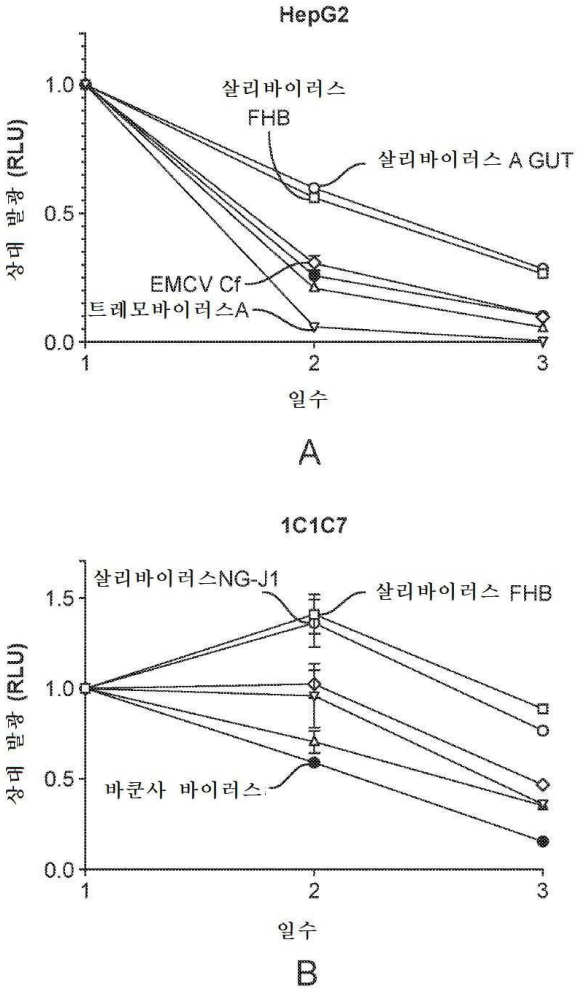

3x10^8
2x10^8
1x10^8
0
24 시간 발광 (RLU)
IRES

E

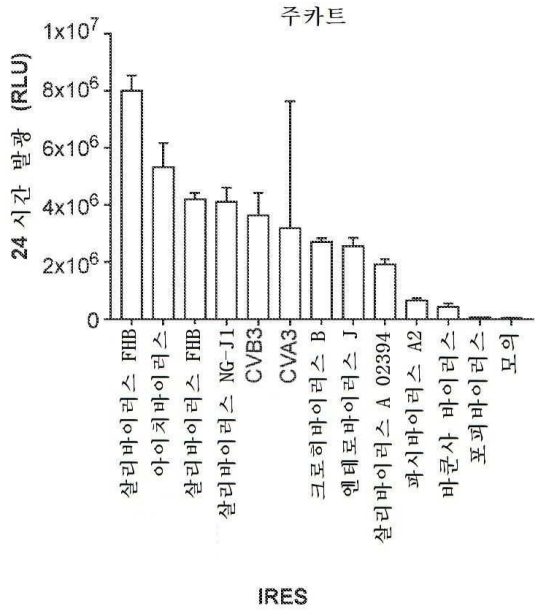
도면2


293
크로히바이러스 B
살리바이러스 FHB
CVB3
24 시간 발광 (RLU)
3x10^8
2x10^8
1x10^8
0
0
200
400
600
800
IRES 길이 (nt)
A
HepG2
아이치 바이러스
살리바이러스 FHB
CVB3
24 시간 발광 (RLU)
4x10^8
3x10^8
2x10^8
1x10^8
0
0
200
400
600
800
IRES 길이 (nt)
B
1C1C7
살리바이러스 FHB
아이치 바이러스
CVB3
24 시간 발광 (RLU)
4x10^8
3x10^8
2x10^8
1x10^8
0
0
200
400
600
800
IRES 길이 (nt)
C

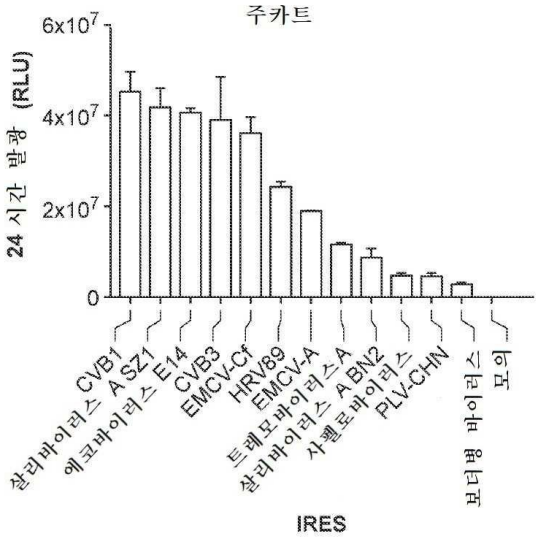
도면3


HepG2
살리바이러스
FHB
살리바이러스 A GUT
EMCV Cf
트레모바이러스A
상대 발광 (RLU)
1.0
0.5
0.0
1
2
3
일수
A
1C1C7
살리바이러스NG-J1
살리바이러스 FHB
바쿤사 바이러스
상대 발광 (RLU)
1.5
1.0
0.5
0.0
1
2
3
일수
B

도면4

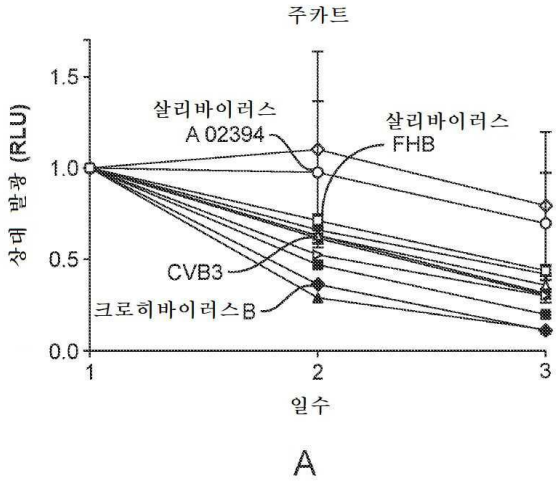

주카트
24 시간 발광 (RLU)
1x10^7
8x10^6
6x10^6
4x10^6
2x10^6
0
살리바이러스 FHB
아이치바이러스
살리바이러스 FHB
살리바이러스 NG-J1
CVB3
CVA3
크로히바이러스 B
엔테로바이러스 J
살리바이러스 A 02394
파시바이러스 A2
바쿤사 바이러스
포퍼바이러스
모의
IRES

A

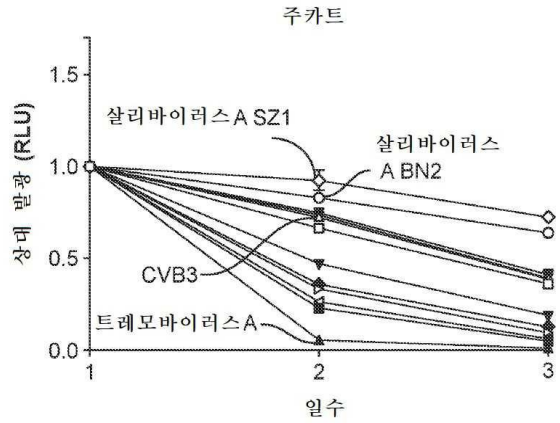

주카트
24 시간 발광 (RLU)
6x10^7
4x10^7
2x10^7
0
CVB1
살리바이러스 A SZ1
에코바이러스 E14
CVB3
EMCV-Cf
HRV89
EMCV-A
트레모바이러스 A
살리바이러스 A BN2
사펠로바이러스
PLV-CHN
보더병 바이러스
모의
IRES

B

도면5

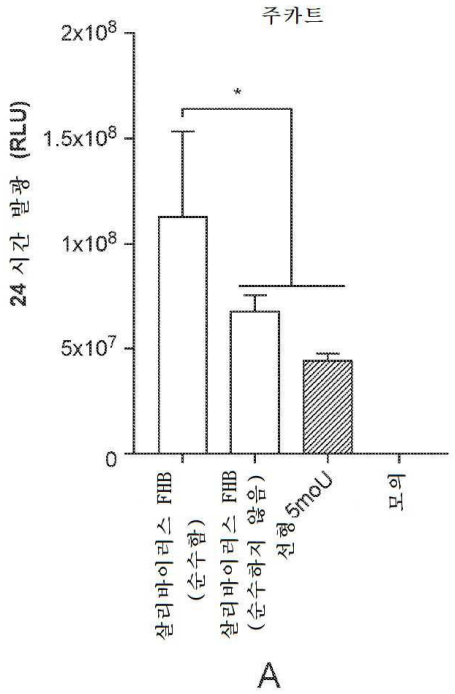

주카트
상대 발광 (RLU)
1.5
1.0
0.5
0.0
살리바이러스
A 02394
살리바이러스
FHB
CVB3
크로히바이러스B
1
2
3
일수

A

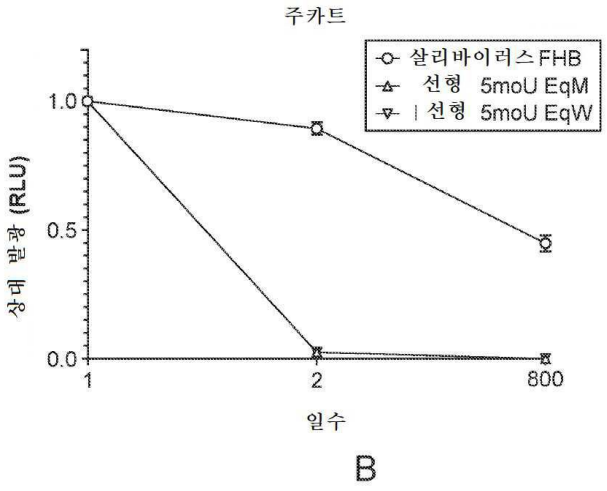

주카트
상대 발광 (RLU)
1.5
1.0
0.5
0.0
살리바이러스A SZ1
살리바이러스
A BN2
CVB3
트레모바이러스A
1
2
3
일수

B

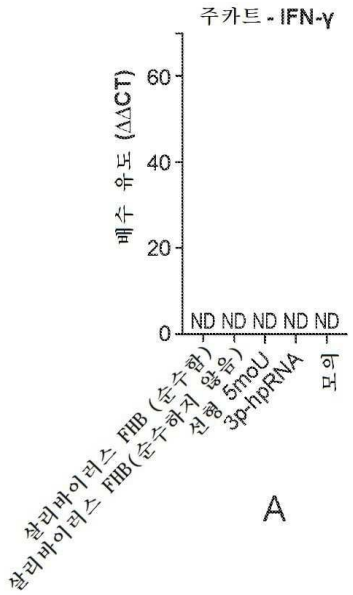
도면6a

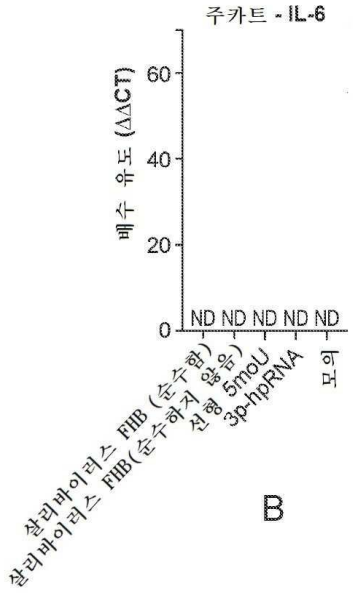

주카트
2x10^8
1.5x10^8
1x10^8
5x10^7
0
24 시간 발광 (RLU)
*
살리바이러스 FHB (순수함)
살리바이러스 FHB (순수하지 않음)
선형 5moU
모의
A

도면6b

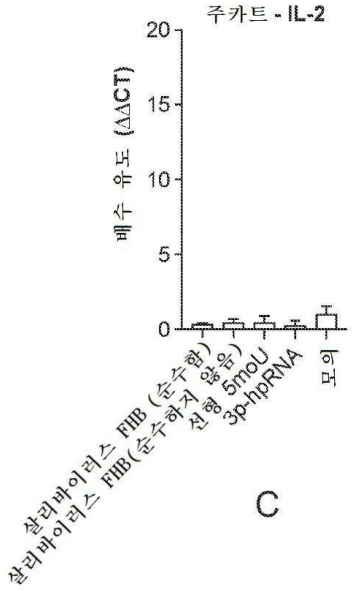

주카트
살리바이러스 FHB
선형 5moU EqM
선형 5moU EqW
1.0
0.5
0.0
상대 발광 (RLU)
1
2
800
일수
B

도면7ab

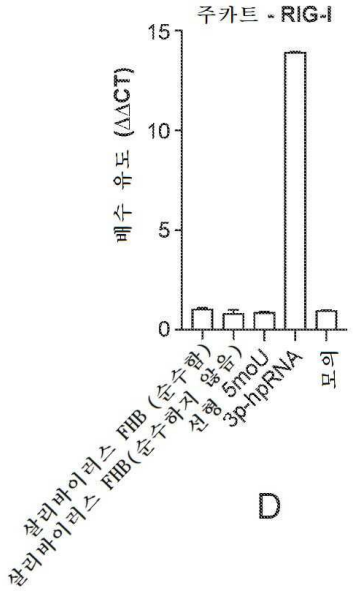

주카트 - IFN-γ
배수 유도 (ΔΔCT)
60
40
20
0
ND ND ND ND ND
살리바이러스 FHB (순수함)
살리바이러스 FHB(순수하지 않음)
신생 5moU
3p-hpRNA
모의
A


주카트 - IL-6
배수 유도 (ΔΔCT)
60
40
20
0
ND ND ND ND ND
살리바이러스 FHB (순수함)
살리바이러스 FHB(순수하지 않음)
신생 5moU
3p-hpRNA
모의
B

도면7cd

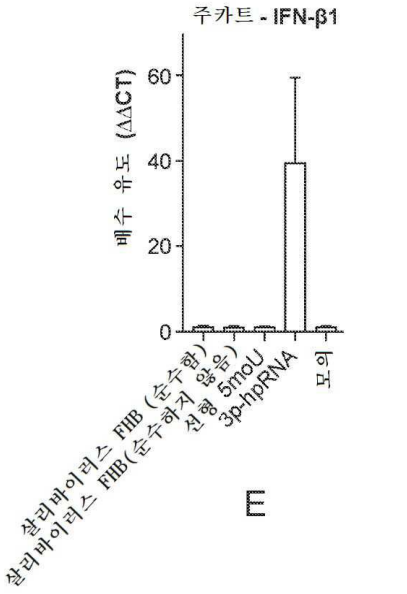

주카트 - IL-2
배수 유도 (ΔΔCT)
20
15
10
5
0
살리바이러스 FIIB (순수함)
살리바이러스 FIIB(순수하지 않음)
신형 5moU
3p-hpRNA
모의

C

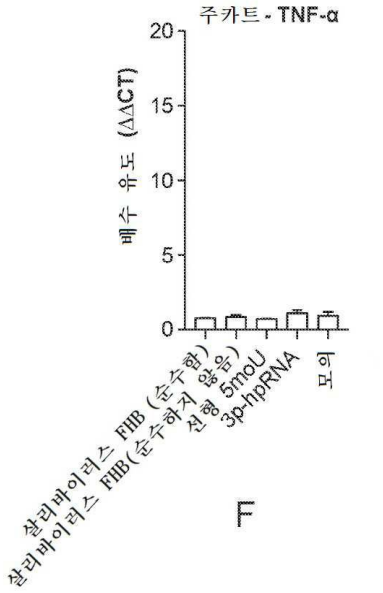

주카트 - RIG-I
배수 유도 (ΔΔCT)
15
10
5
0
살리바이러스 FIIB (순수함)
살리바이러스 FIIB(순수하지 않음)
신형 5moU
3p-hpRNA
모의

D

도면7ef

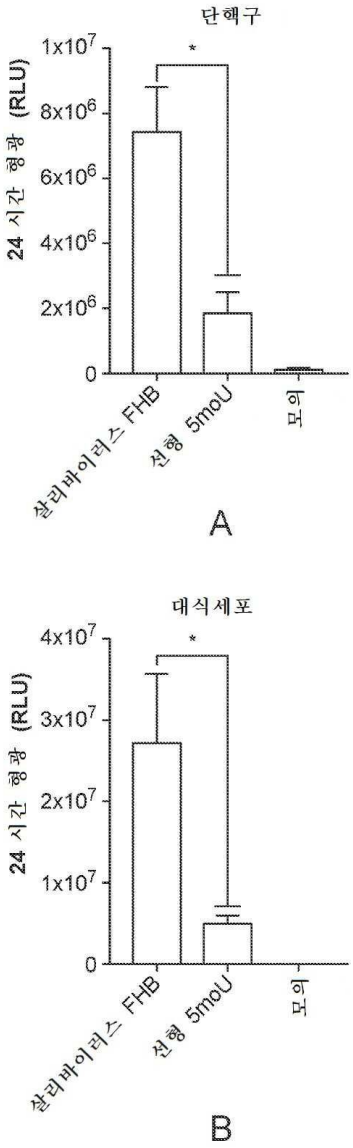

주카트 - IFN-β1
배수 유도 (ΔΔCT)
60
40
20
0
살리바이러스 FIB (순수함)
살리바이러스 FIB(순수하지 않음)
선형 5moU
3p-hpRNA
모의

E

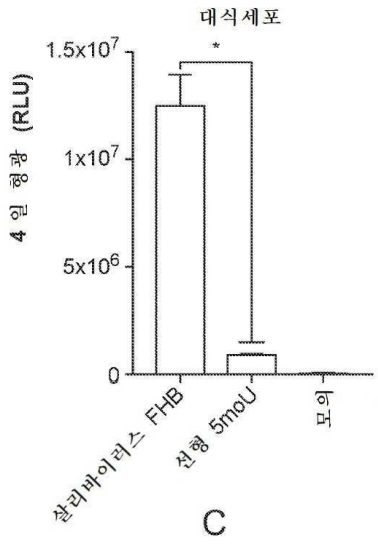

주카트 - TNF-α
배수 유도 (ΔΔCT)
20
15
10
5
0
살리바이러스 FIB (순수함)
살리바이러스 FIB(순수하지 않음)
선형 5moU
3p-hpRNA
모의

F

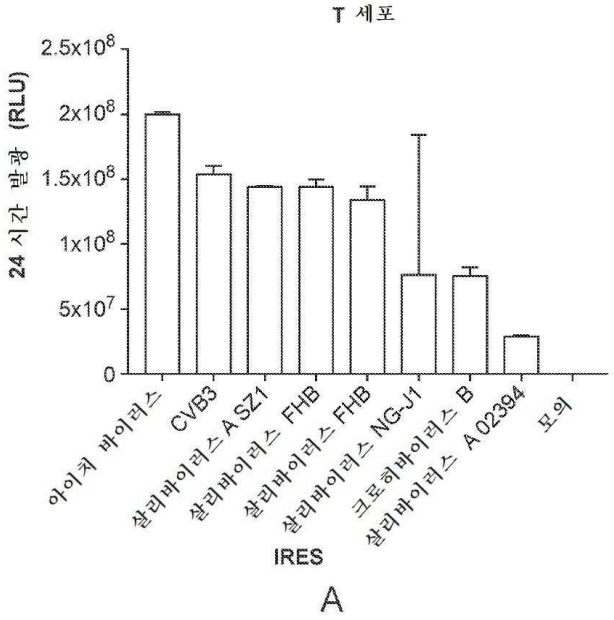
도면8ab


단핵구
24 시간 형광 (RLU)
1x10^7
8x10^6
6x10^6
4x10^6
2x10^6
0
*
살리바이러스 FHB
선형 5moU
모의
A
대식세포
24 시간 형광 (RLU)
4x10^7
3x10^7
2x10^7
1x10^7
0
*
살리바이러스 FHB
선형 5moU
모의
B

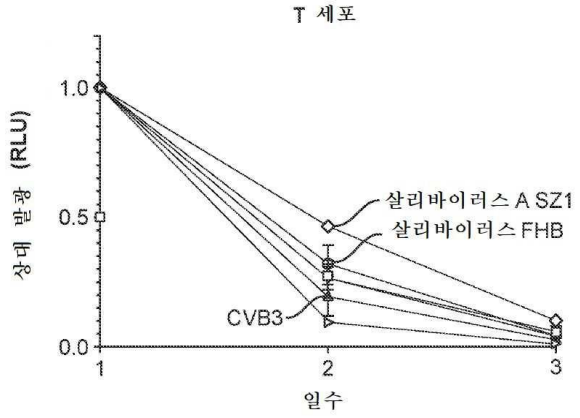
도면8c



도면9

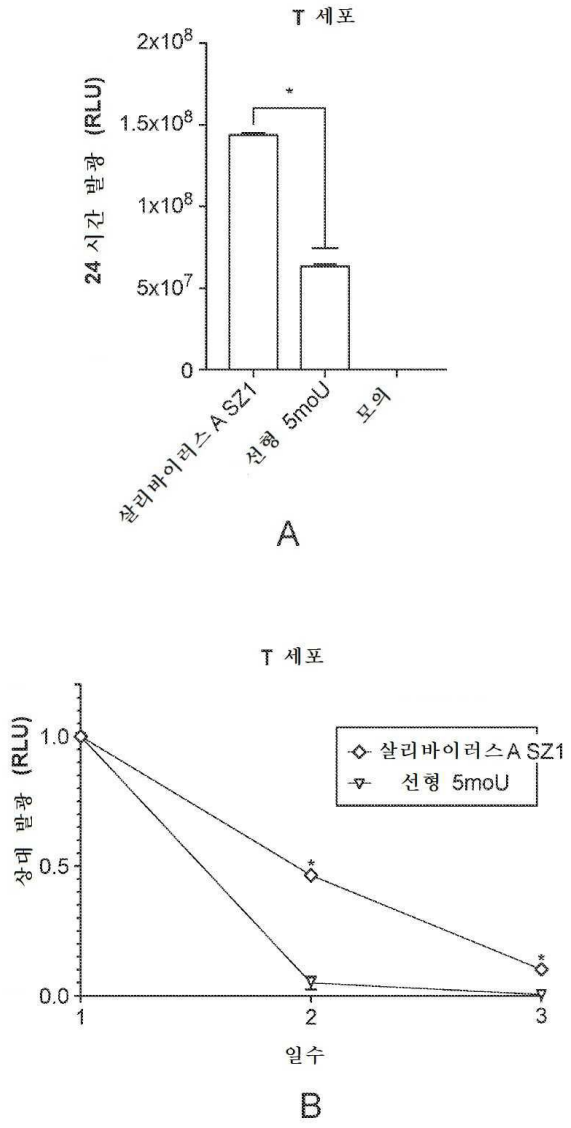

T 세포
24 시간 발광 (RLU)
2.5x10^8
2x10^8
1.5x10^8
1x10^8
5x10^7
0
아이치 바이러스
CVB3
살리바이러스 A SZ1
살리바이러스 FHB
살리바이러스 FHB
살리바이러스 NG-J1
크로히바이러스 B
살리바이러스 A 02394
모의
IRES

A

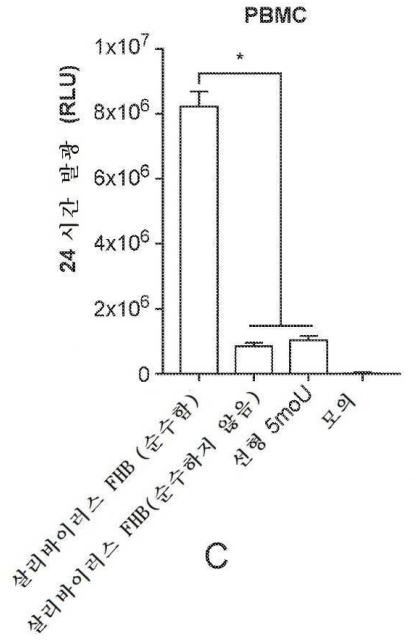

T 세포
상대 발광 (RLU)
1.0
0.5
0.0
살리바이러스 A SZ1
살리바이러스 FHB
CVB3
1
2
3
일수

B

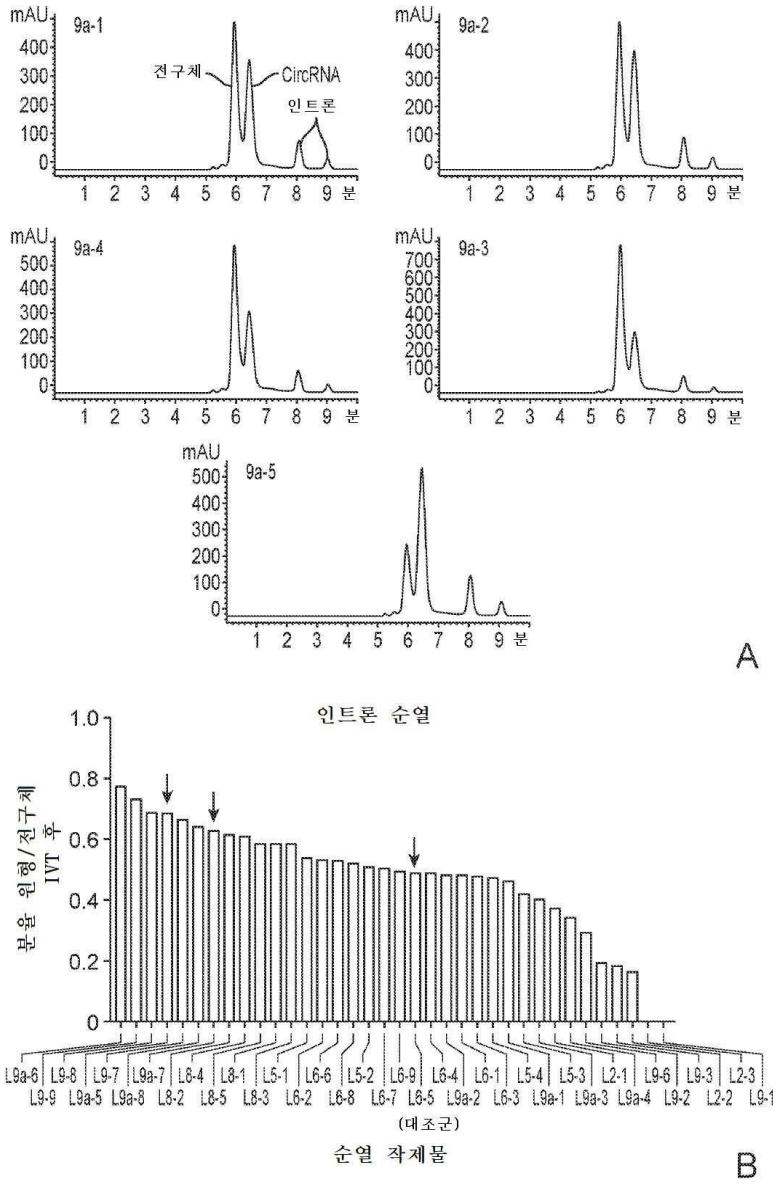
도면10ab


T 세포
24 시간 발광 (RLU)
2x10^8
1.5x10^8
1x10^8
5x10^7
0
*
살리바이러스 A SZ1
선형 5moU
모의
A
T 세포
상대 발광 (RLU)
1.0
0.5
0.0
살리바이러스 A SZ1
선형 5moU
*
*
1
2
3
일수
B

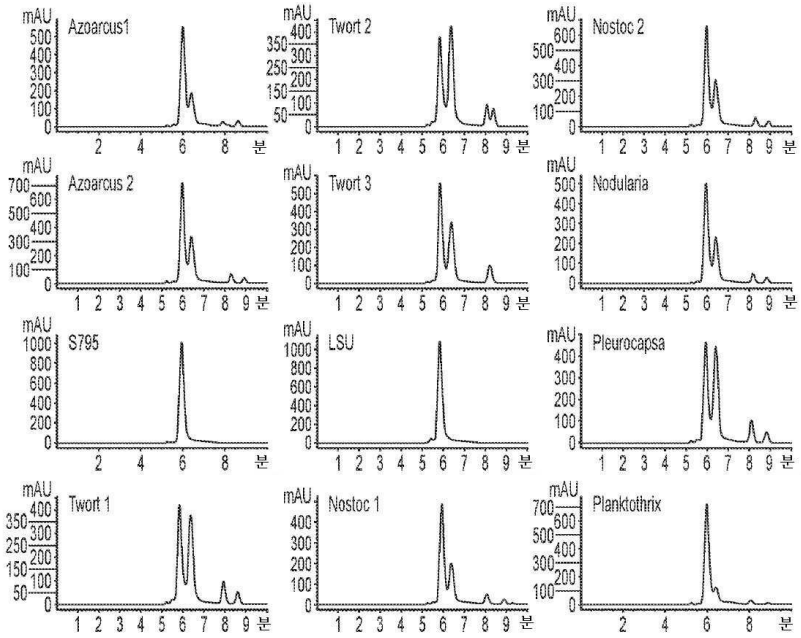
도면10c


PBMC
24 시간 발광 (RLU)
1x10^7
8x10^6
6x10^6
4x10^6
2x10^6
0
*
살리바이러스 FIIB (순수함)
살리바이러스 FIIB(순수하지 않음)
신형 5moU
모의
C

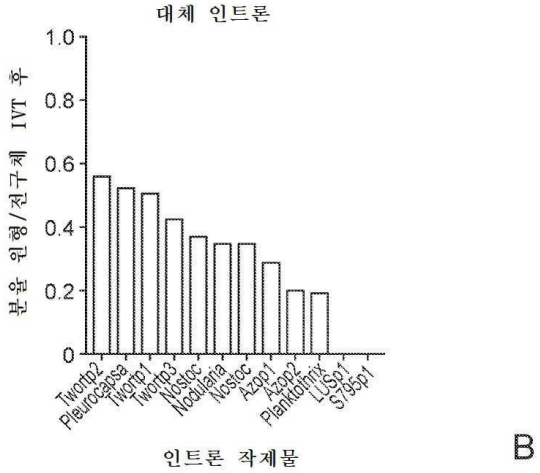
도면11


mAU
9a-1
전구체
CircRNA
인트론
1 2 3 4 5 6 7 8 9 분
9a-2
9a-4
9a-3
9a-5
A
인트론 순열
분율 원형/전구체
IVT 후
1.0
0.8
0.6
0.4
0.2
0
L9a-6 L9-8 L9-7 L9a-7 L6-4 L8-1 L5-1 L6-6 L5-2 L6-9 L6-4 L6-1 L5-4 L5-3 L2-1 L9-6 L9-3 L2-3
L9-9 L9a-5 L9a-8 L8-2 L8-5 L8-3 L6-2 L6-8 L6-7 L6-5 L9a-2 L6-3 L9a-1 L9a-3 L9a-4 L9-2 L2-2 L9-1
(대조군)
순열 작제물
B

도면12

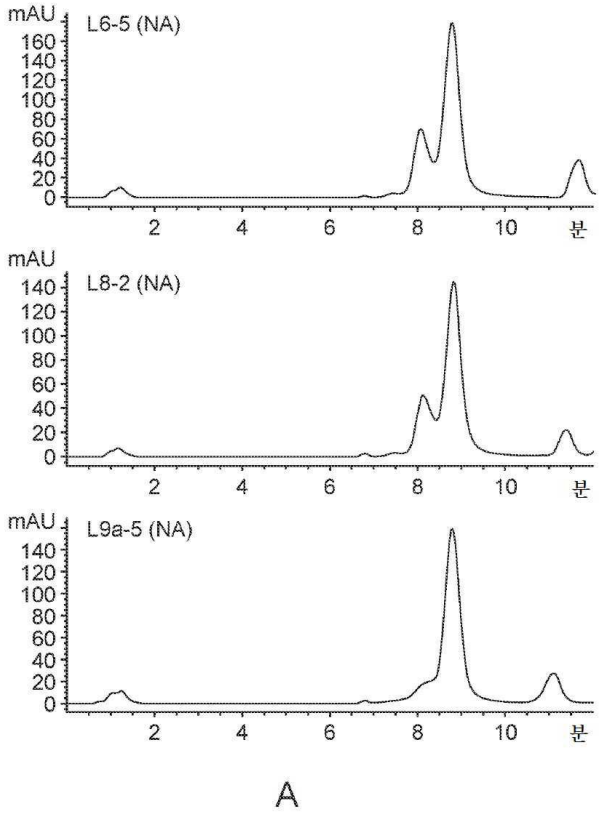

Azoarcus1
Twort 2
Nostoc 2
Azoarcus 2
Twort 3
Nodularia
S795
LSU
Pleurocapsa
Twort 1
Nostoc 1
Planktothrix
mAU
분

A


대체 인트론
분율 원형/천구체 IVT 후
1.0
0.8
0.6
0.4
0.2
0
Twortp2
Pleurocapsa
Twortp1
Twortp3
Nostoc
Nodularia
Nostoc
Azop1
Azop2
Planktothrix
LUSp1
S795p1
인트론 작제물

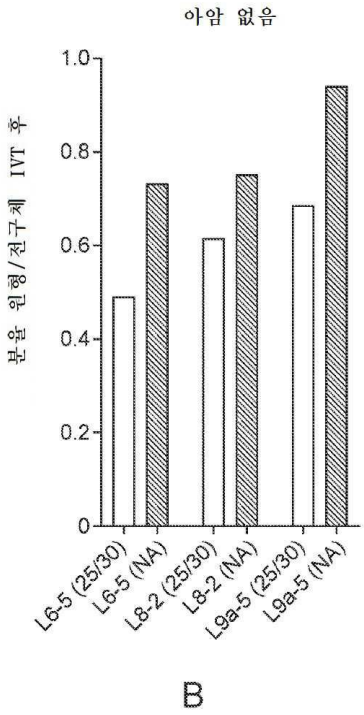
B

도면13a



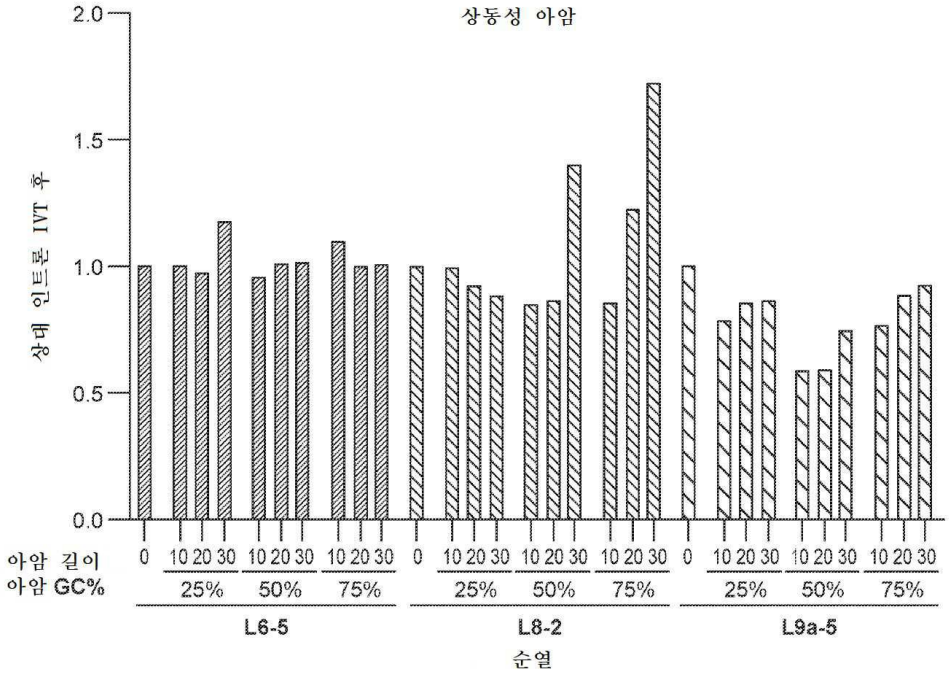
A

도면13b

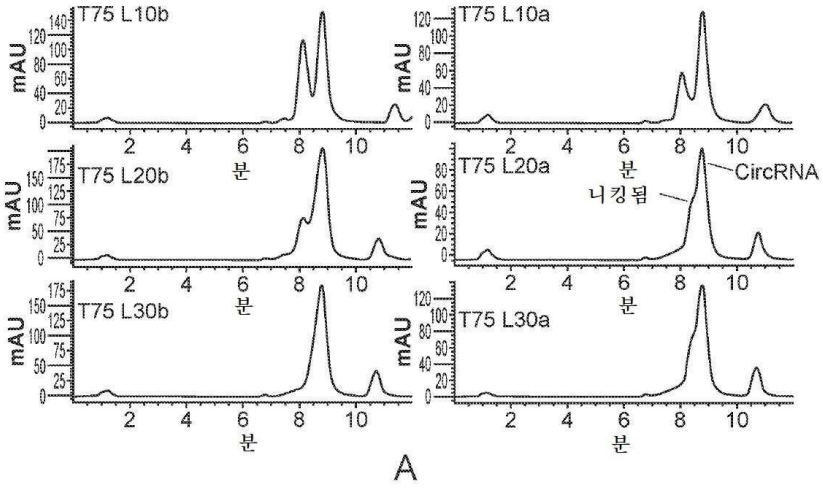


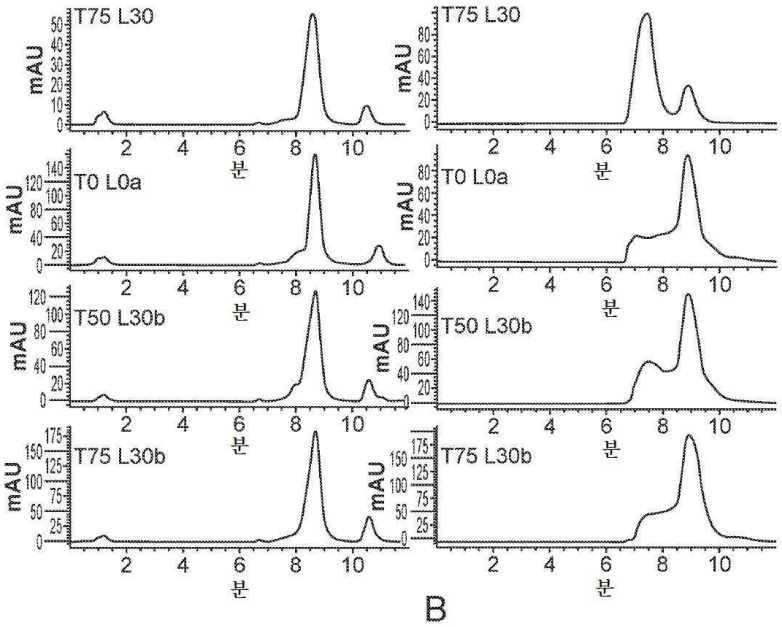
B

도면14

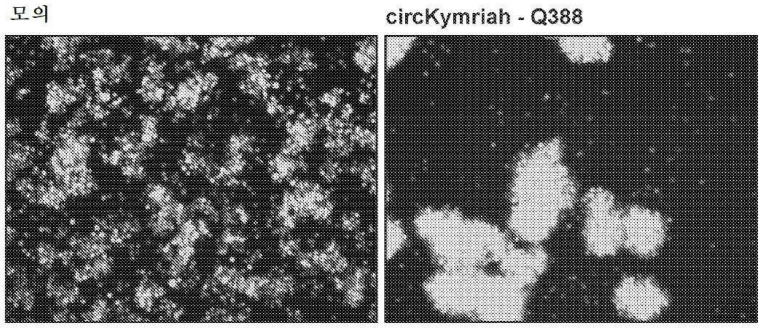

상동성 아암
2.0
1.5
1.0
0.5
0.0
상대 인트론 IVT 후
아암 길이
아암 GC%
0
10 20 30
25%
50%
75%
L6-5
L8-2
L9a-5
순열

도면15

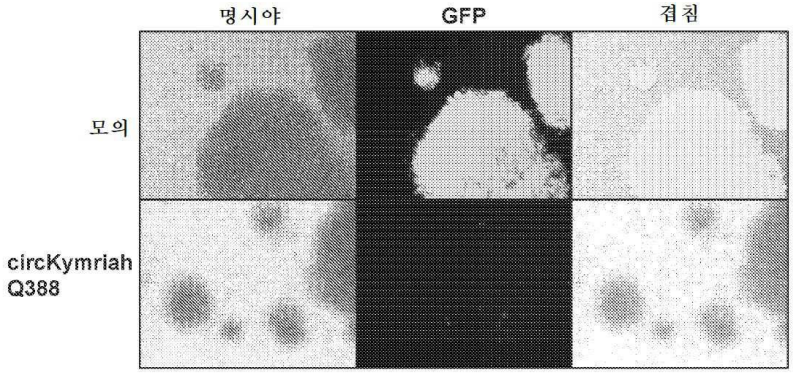

T75 L10b
T75 L10a
T75 L20b
T75 L20a
CircRNA
니킹됨
T75 L30b
T75 L30a
mAU
분
A


T75 L30
T0 L0a
T50 L30b
T75 L30b
mAU
분
B

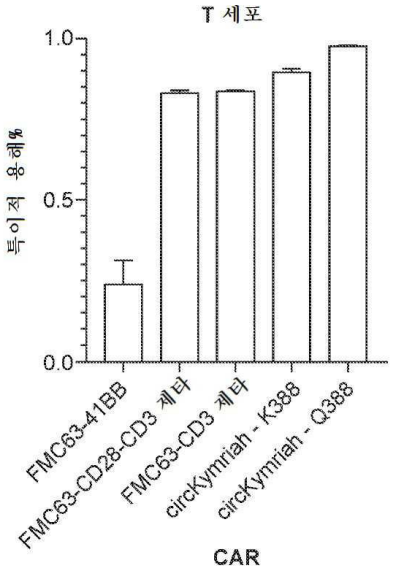
도면16


모의
circKymriah - Q388

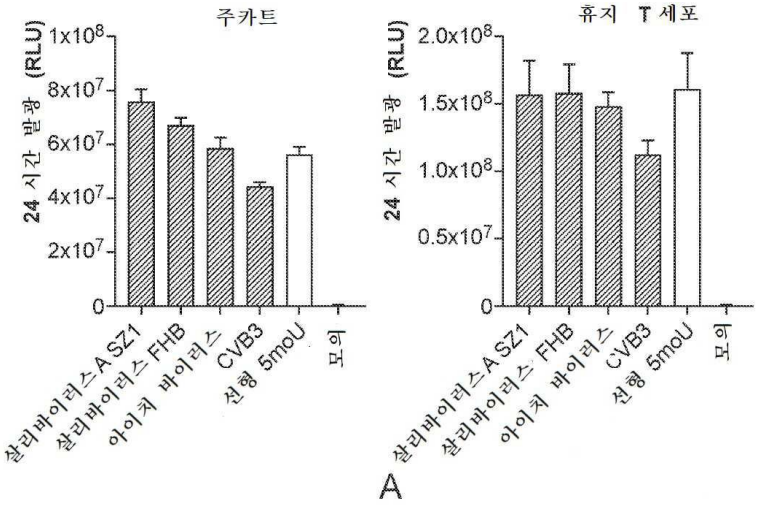
도면17

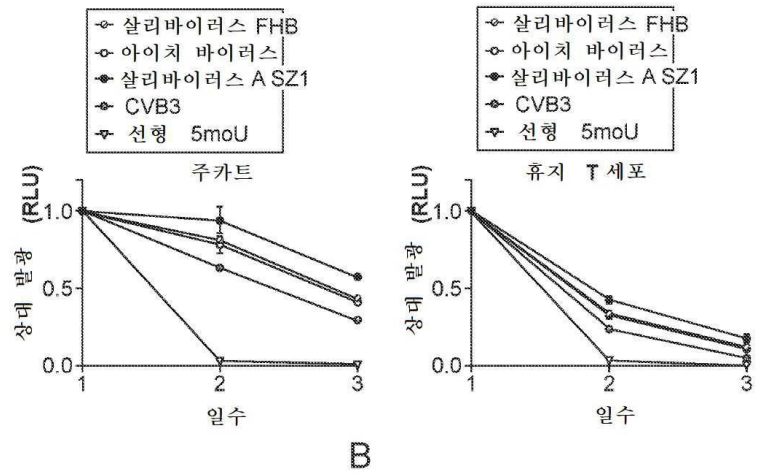

명시야
GFP
겹침
모의
circKymriah
Q388

도면18

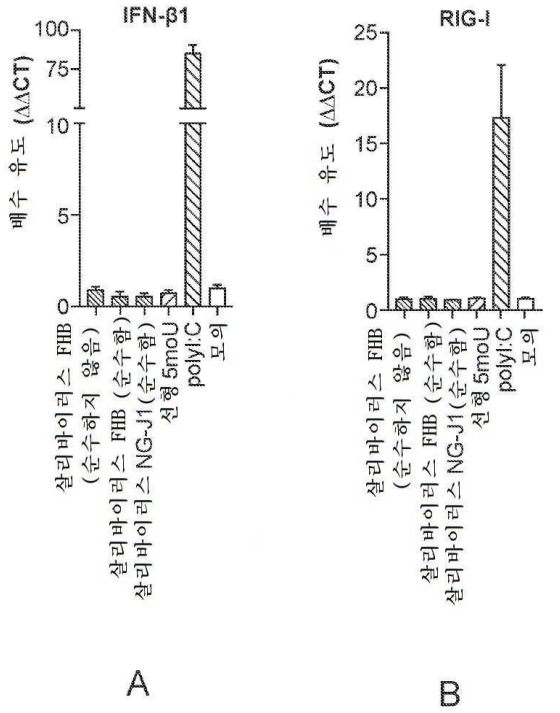

T 세포
1.0
0.5
0.0
특이적 용해%
FMC63-41BB
FMC63-CD28-CD3 제타
FMC63-CD3 제타
circKymriah - K388
circKymriah - Q388
CAR

도면19

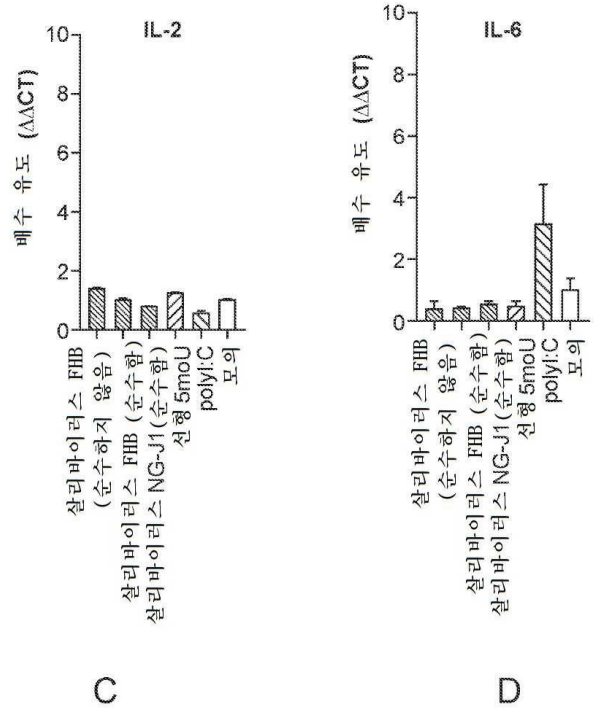

주카트
24 시간 발광 (RLU)
1x10^8
8x10^7
6x10^7
4x10^7
2x10^7
0
살리바이러스 A SZ1
살리바이러스 FHB
아이치 바이러스
CVB3
선형 5moU
모의
휴지 T 세포
24 시간 발광 (RLU)
2.0x10^8
1.5x10^8
1.0x10^8
0.5x10^7
0
살리바이러스 A SZ1
살리바이러스 FHB
아이치 바이러스
CVB3
선형 5moU
모의
A


살리바이러스 FHB
아이치 바이러스
살리바이러스 A SZ1
CVB3
선형 5moU
주카트
상대 발광 (RLU)
1.0
0.5
0.0
1
2
3
일수
살리바이러스 FHB
아이치 바이러스
살리바이러스 A SZ1
CVB3
선형 5moU
휴지 T 세포
상대 발광 (RLU)
1.0
0.5
0.0
1
2
3
일수
B

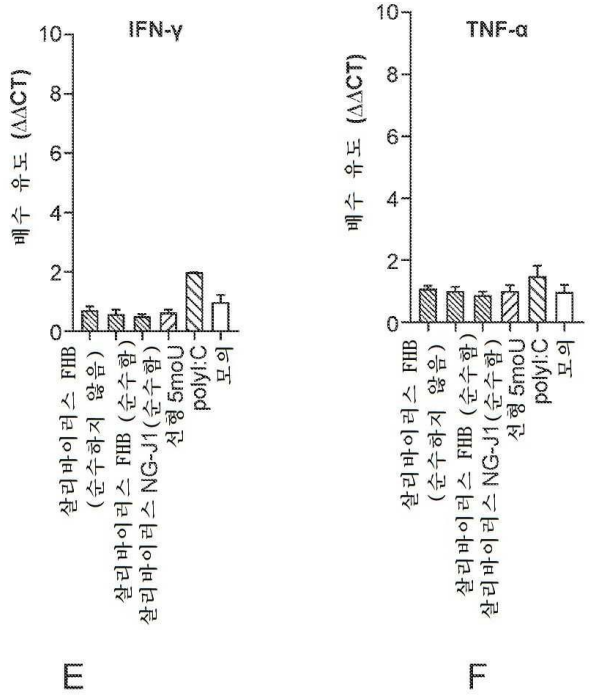
도면20ab



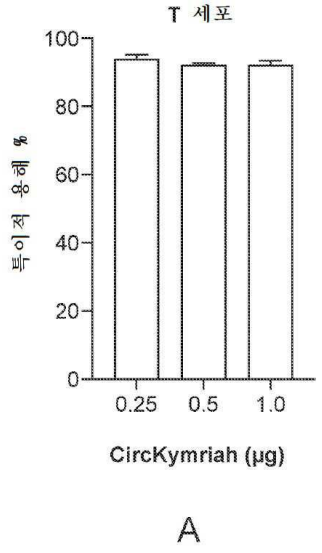
도면20cd



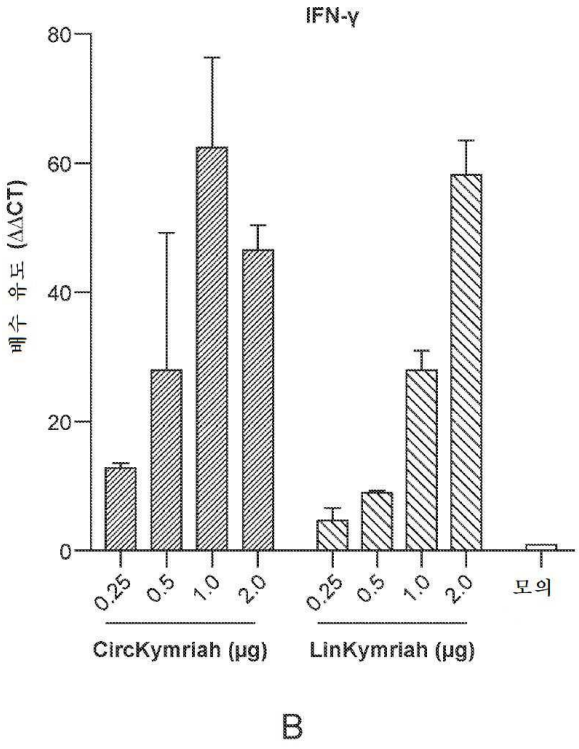
도면20ef



도면21a

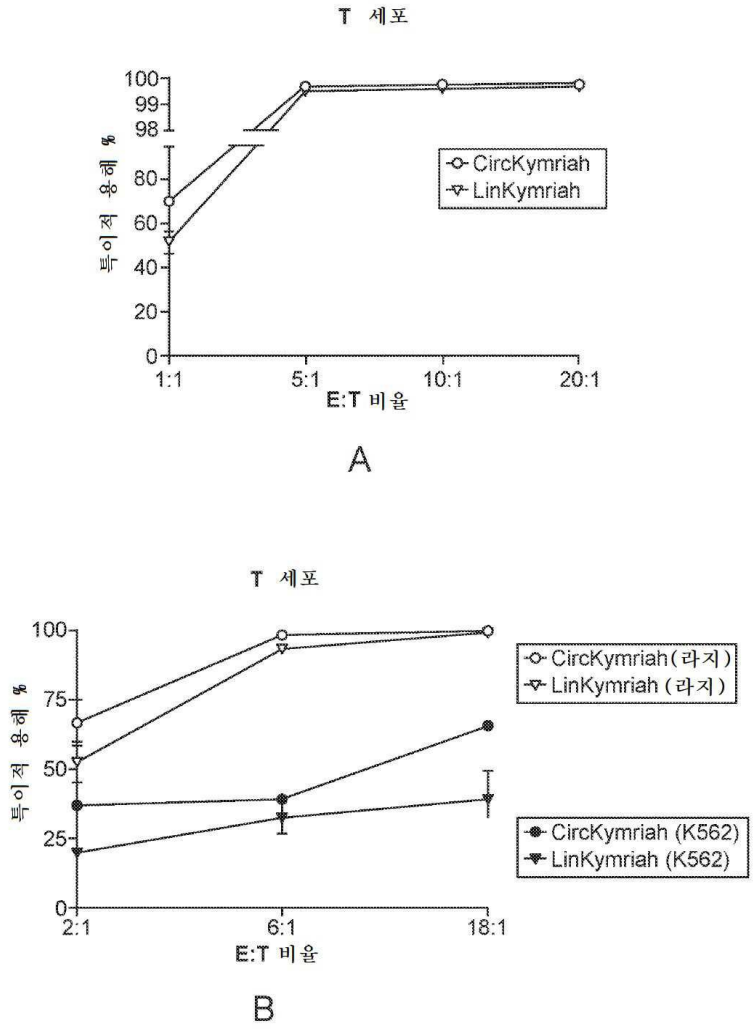


도면21b

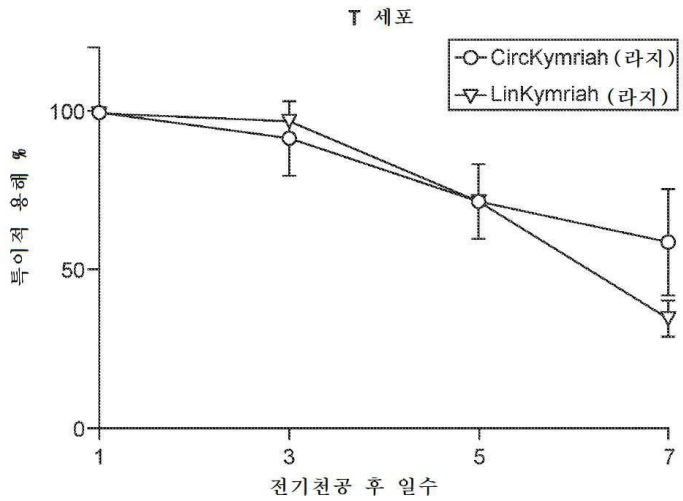

IFN-γ
배수 유도 (ΔΔCT)
80
60
40
20
0
0.25
0.5
1.0
2.0
0.25
0.5
1.0
2.0
모의
CircKymriah (μg)
LinKymriah (μg)

B

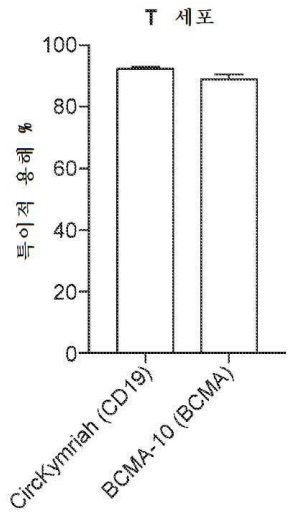
도면22


T 세포
100
99
98
80
60
40
20
0
특이적 용해 %
CircKymriah
LinKymriah
1:1
5:1
10:1
20:1
E:T 비율
A
T 세포
100
75
50
25
0
특이적 용해 %
CircKymriah(라지)
LinKymriah(라지)
CircKymriah (K562)
LinKymriah (K562)
2:1
6:1
18:1
E:T 비율
B

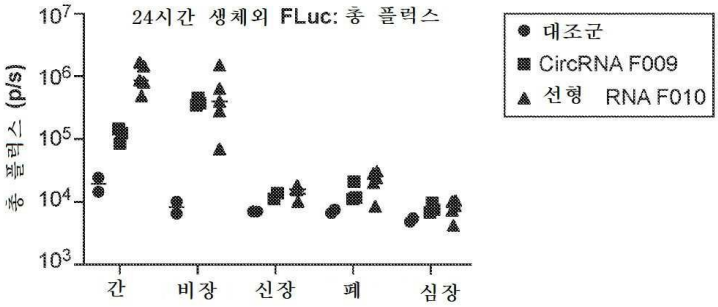
도면23


T 세포
CircKymriah(라지)
LinKymriah(라지)
100
50
0
특이적 용해 %
1
3
5
7
전기천공 후 일수

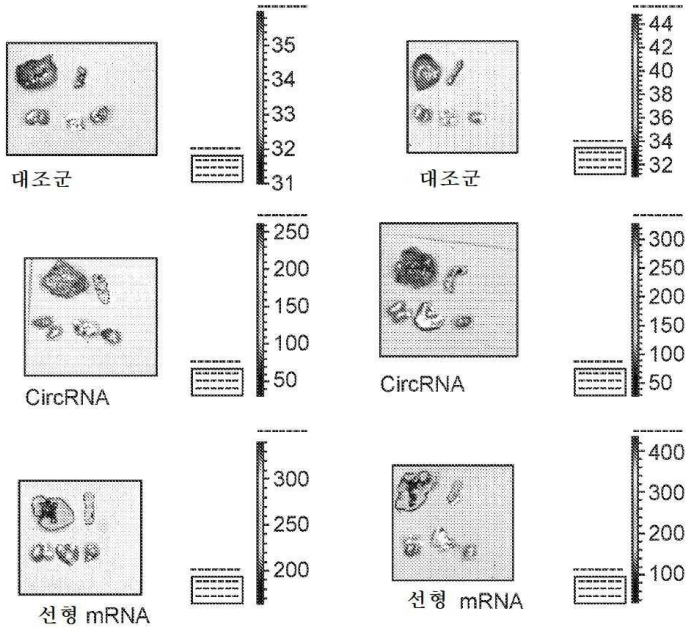
도면24



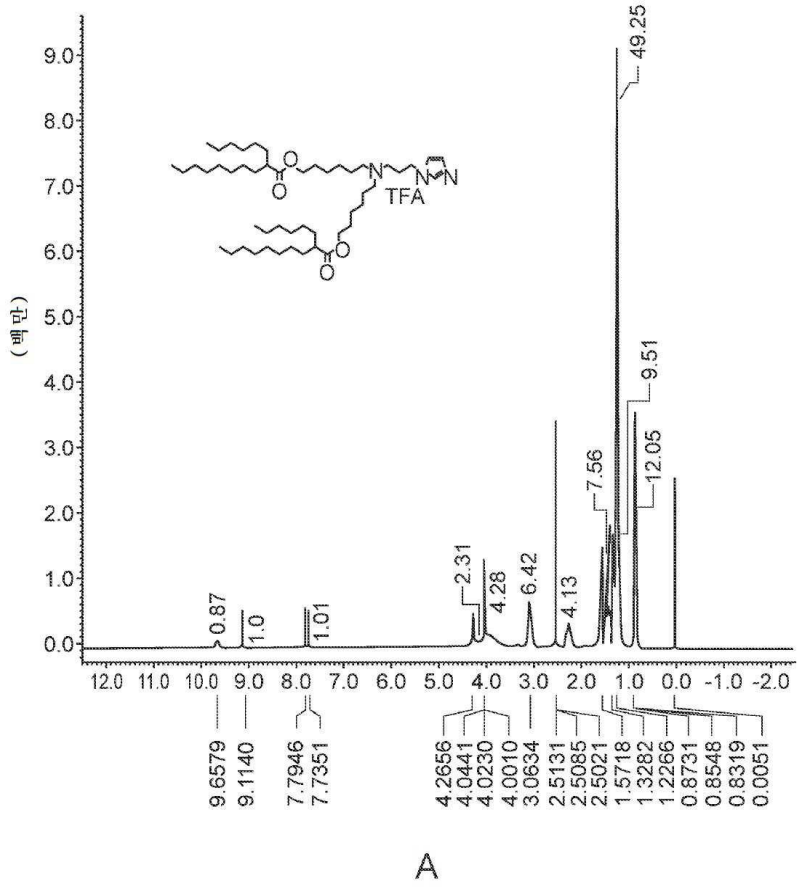
도면25



도면26

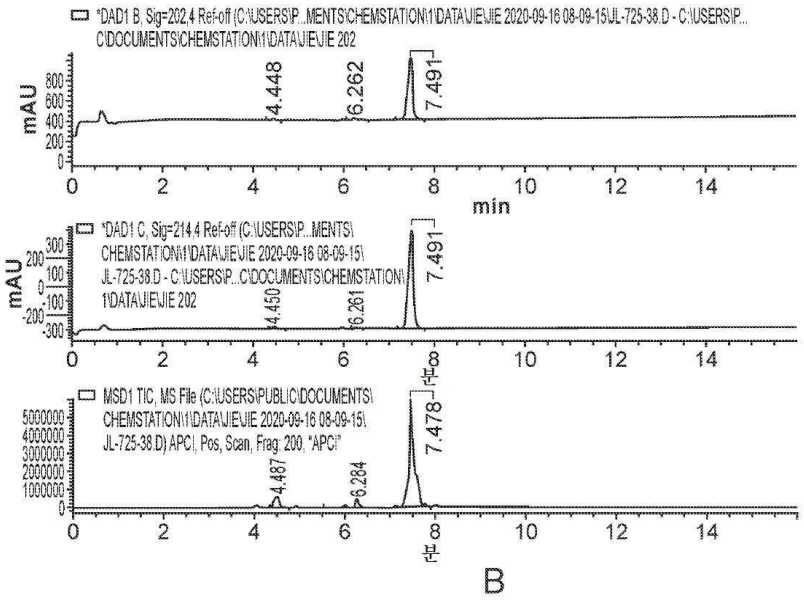


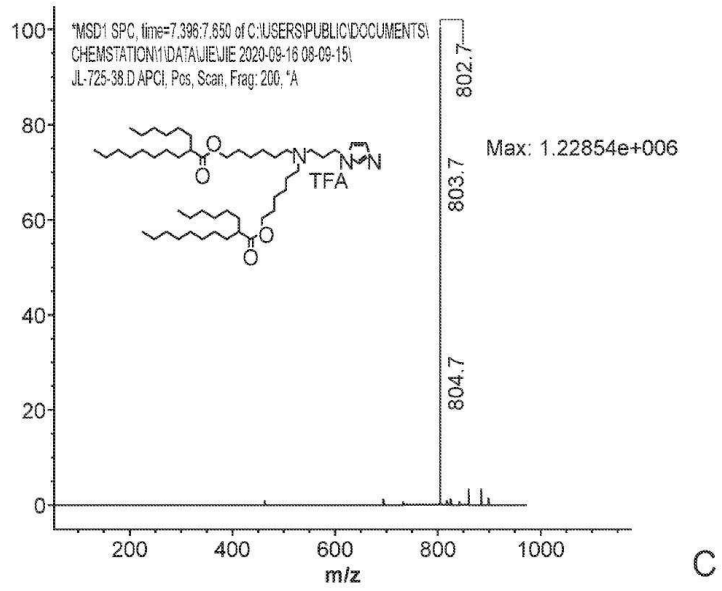
도면27a


(백만)
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
TFA
49.25
9.51
12.05
7.56
2.31
4.28
6.42
4.13
0.87
1.0
1.01
12.0 11.0 10.0 9.0 8.0 7.0 6.0 5.0 4.0 3.0 2.0 1.0 0.0 -1.0 -2.0
9.6579
9.1140
7.7946
7.7351
4.2656
4.0441
4.0230
4.0010
3.0634
2.5131
2.5085
2.5021
1.5718
1.3282
1.2266
0.8731
0.8548
0.8319
0.0051

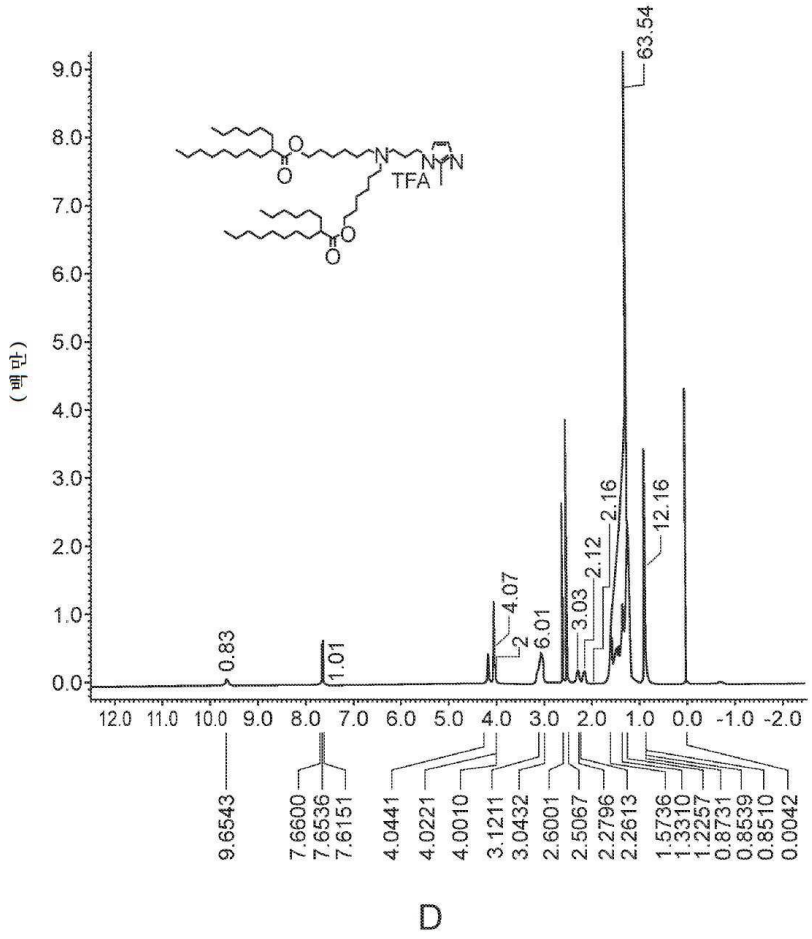
A

도면27bc


*DAD1 B, Sig=202,4 Ref=off (C:\USERS\P...MENTS\CHEMSTATION\1\DATA\JIEJIE 2020-09-16 08-09-15\JL-725-38.D - C:\USERS\P...
C\DOCUMENTS\CHEMSTATION\1\DATA\JIEJIE 202
4.448
6.262
7.491
mAU
min
*DAD1 C, Sig=214,4 Ref=off (C:\USERS\P...MENTS\
CHEMSTATION\1\DATA\JIEJIE 2020-09-16 08-09-15\
JL-725-38.D - C:\USERS\P...C\DOCUMENTS\CHEMSTATION\
1\DATA\JIEJIE 202
4.450
6.261
7.491
MSD1 TIC, MS File (C:\USERS\PUBLIC\DOCUMENTS\
CHEMSTATION\1\DATA\JIEJIE 2020-09-16 08-09-15\
JL-725-38.D) APCI, Pos, Scan, Frag: 200, "APCI"
4.487
6.284
7.478
분
B

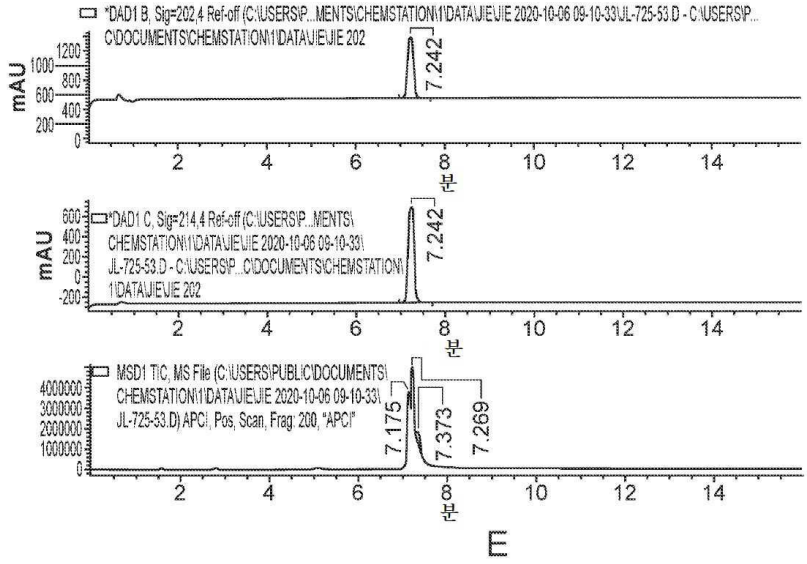

*MSD1 SPC, time=7.398:7.650 of C:\USERS\PUBLIC\DOCUMENTS\
CHEMSTATION\1\DATA\JIEJIE 2020-09-16 08-09-15\
JL-725-38.D APCI, Pos, Scan, Frag: 200, "A
TFA
802.7
803.7
804.7
Max: 1.22854e+006
m/z
C

도면27d

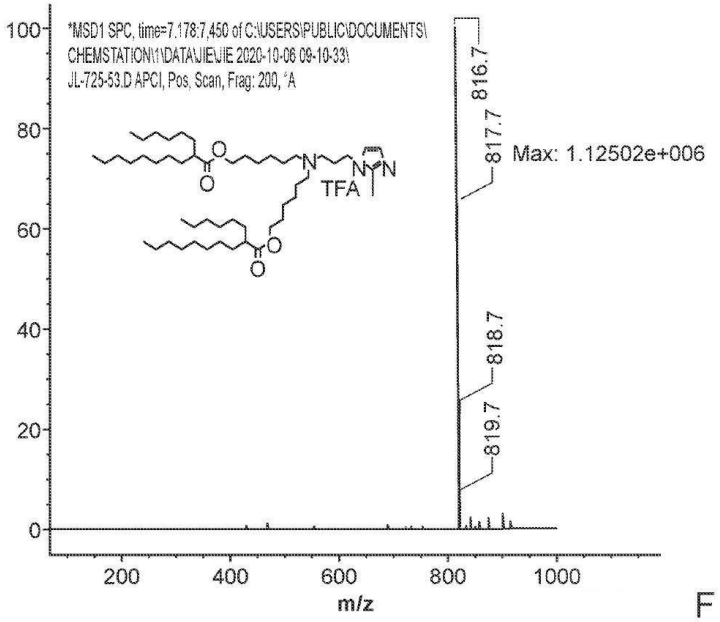

(백만)
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
TFA
63.54
12.16
2.16
2.12
3.03
6.01
4.07
2
1.01
0.83
12.0 11.0 10.0 9.0 8.0 7.0 6.0 5.0 4.0 3.0 2.0 1.0 0.0 -1.0 -2.0
9.6543
7.6600
7.6536
7.6151
4.0441
4.0221
4.0010
3.1211
3.0432
2.6001
2.5067
2.2796
2.2613
1.5736
1.3310
1.2257
0.8731
0.8539
0.8510
0.0042

D

도면27ef

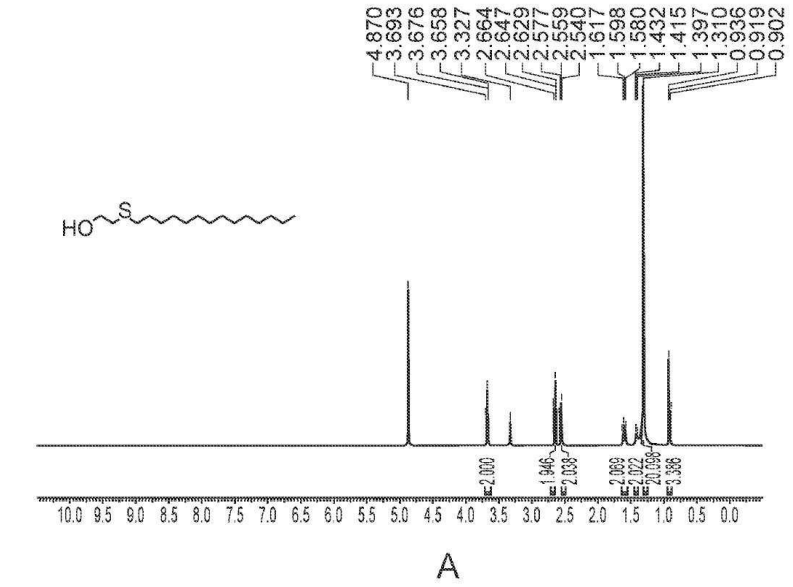

*DAD1 B, Sig=202,4 Ref-off (C:\USERS\P...MENTS\CHEMSTATION\1\DATA\JIEJIE 2020-10-06 09-10-33\JL-725-53.D - C:\USERS\P...
C\DOCUMENTS\CHEMSTATION\1\DATA\JIEJIE 202
mAU
분
7.242
*DAD1 C, Sig=214,4 Ref-off (C:\USERS\P...MENTS\
CHEMSTATION\1\DATA\JIEJIE 2020-10-06 09-10-33\
JL-725-53.D - C:\USERS\P...C\DOCUMENTS\CHEMSTATION\
1\DATA\JIEJIE 202
MSD1 TIC, MS File (C:\USERS\PUBLIC\DOCUMENTS\
CHEMSTATION\1\DATA\JIEJIE 2020-10-06 09-10-33\
JL-725-53.D) APCI, Pos, Scan, Frag: 200, "APCI"
7.175
7.373
7.269

E

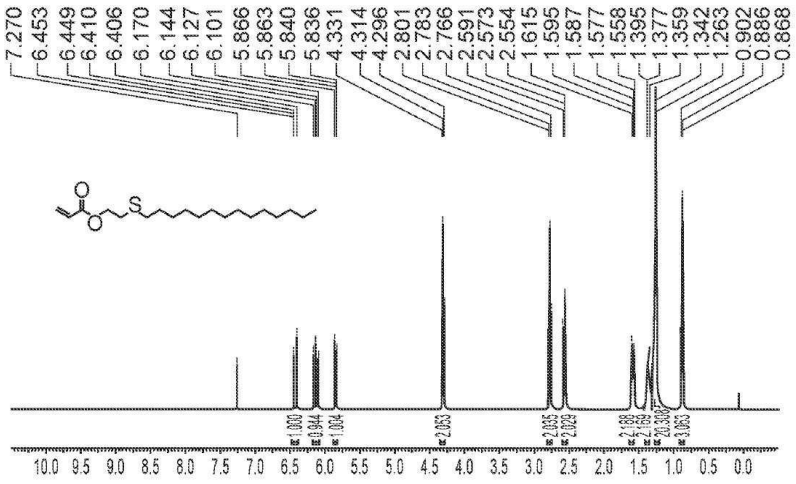

*MSD1 SPC, time=7.178:7.450 of C:\USERS\PUBLIC\DOCUMENTS\
CHEMSTATION\1\DATA\JIEJIE 2020-10-06 09-10-33\
JL-725-53.D APCI, Pos, Scan, Frag: 200, "A
TFA
816.7
817.7
818.7
819.7
Max: 1.12502e+006
m/z

F

도면28ab

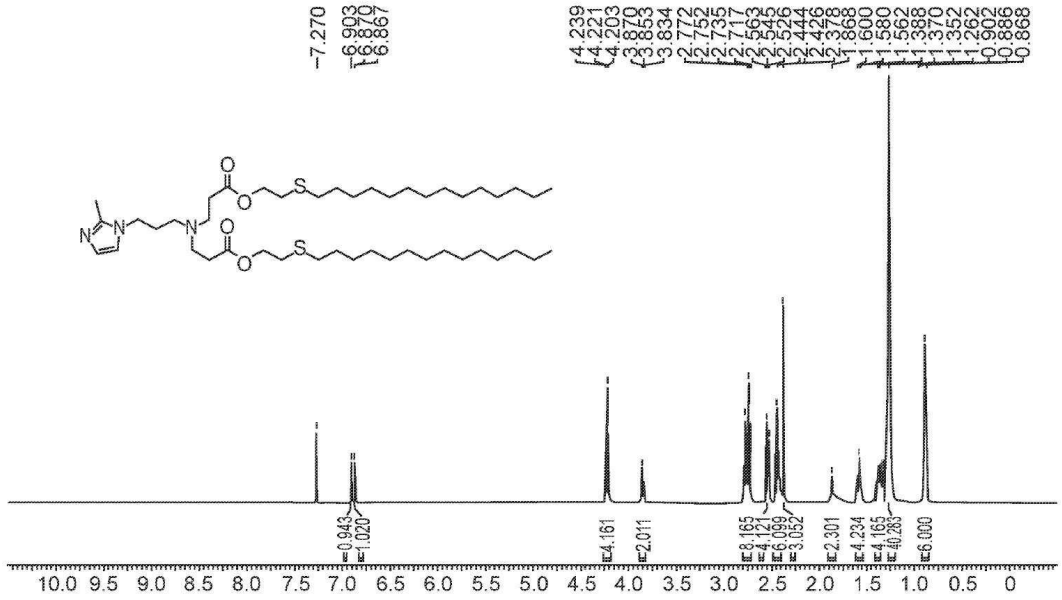

4.870
3.693
3.676
3.658
3.327
2.664
2.647
2.629
2.577
2.559
2.540
1.617
1.598
1.580
1.432
1.415
1.397
1.310
0.936
0.919
0.902
HO
S
2.000
1.946
2.038
2.069
2.022
20.059
3.386
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

A


7.270
6.453
6.449
6.410
6.406
6.170
6.144
6.127
6.101
5.866
5.863
5.840
5.836
4.331
4.314
4.296
2.801
2.783
2.766
2.591
2.573
2.554
1.615
1.595
1.587
1.577
1.558
1.395
1.377
1.359
1.342
1.263
0.902
0.886
0.868
O
O
S
1.000
0.944
1.004
2.053
2.035
2.020
2.188
2.169
20.336
3.053
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

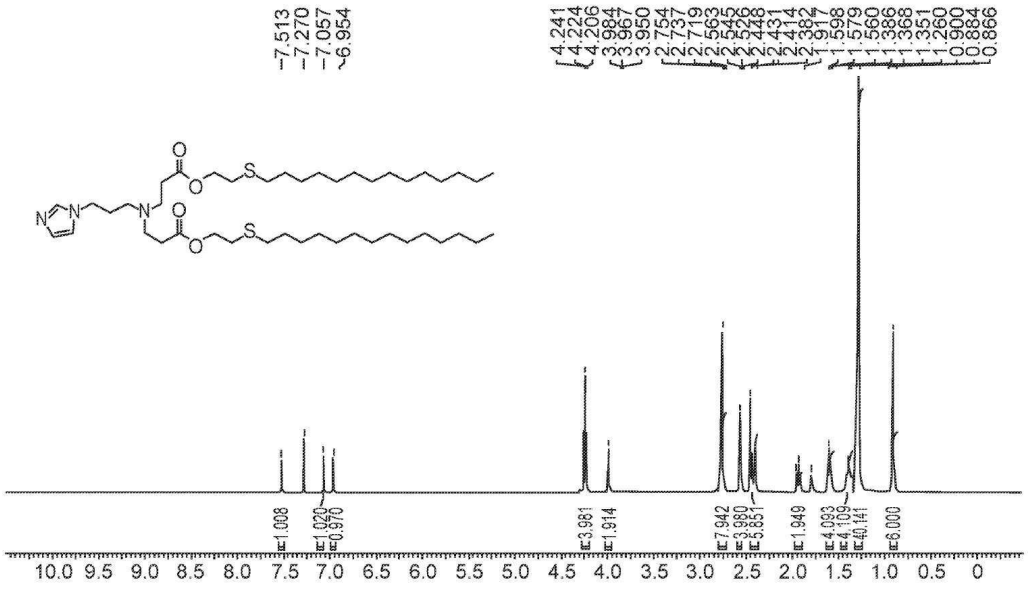
B

도면28c

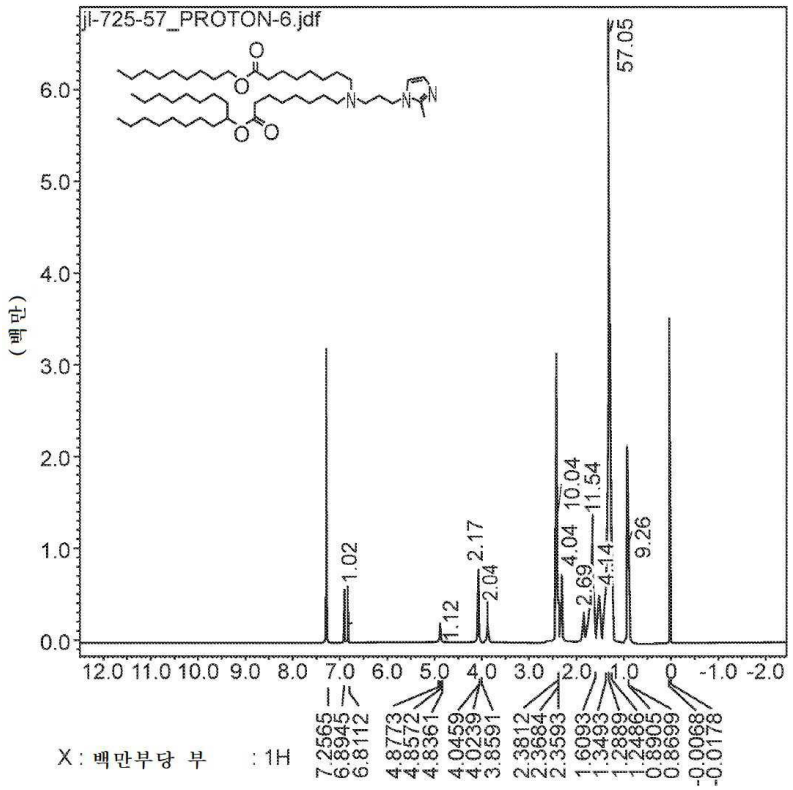

7.270
6.903
6.870
6.867
4.239
4.221
4.203
3.870
3.853
3.834
2.772
2.752
2.735
2.717
2.563
2.545
2.526
2.444
2.426
2.378
1.868
1.600
1.580
1.562
1.388
1.370
1.352
1.262
0.902
0.886
0.868
0.943
1.020
4.161
2.011
8.165
4.121
6.099
3.052
2.301
4.234
4.165
40.283
6.000
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0

C

도면29

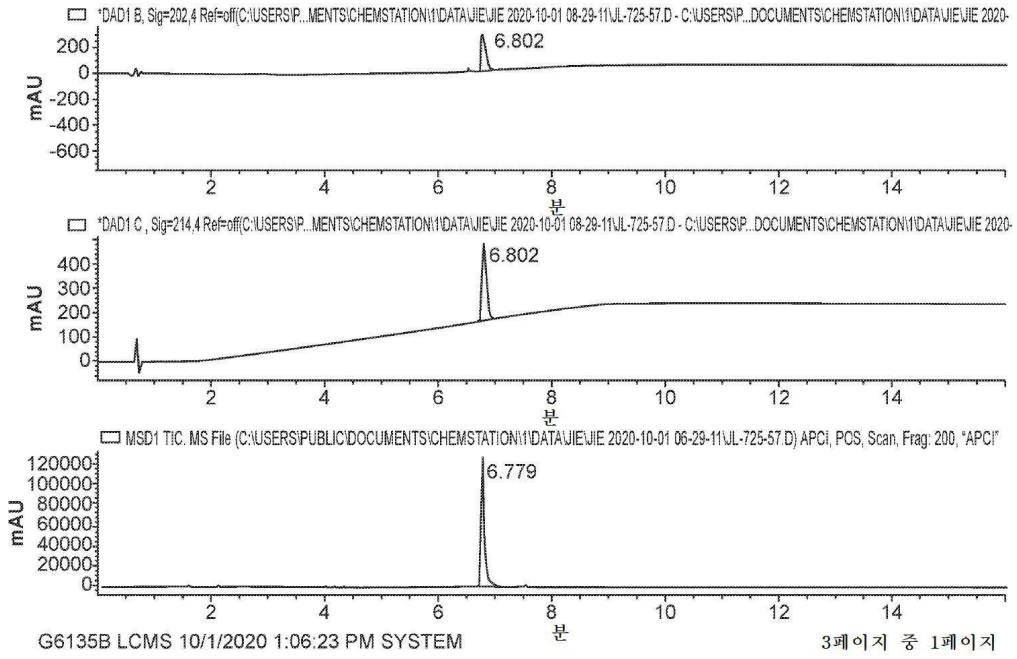

7.513
7.270
7.057
6.954
4.241
4.224
4.206
3.984
3.967
3.950
2.754
2.737
2.719
2.563
2.545
2.526
2.448
2.431
2.414
2.382
1.917
1.598
1.579
1.560
1.386
1.368
1.351
1.260
0.900
0.884
0.866
1.008
1.020
0.970
3.981
1.914
7.942
3.969
5.851
1.949
4.093
4.109
40.141
6.000
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0

도면30a



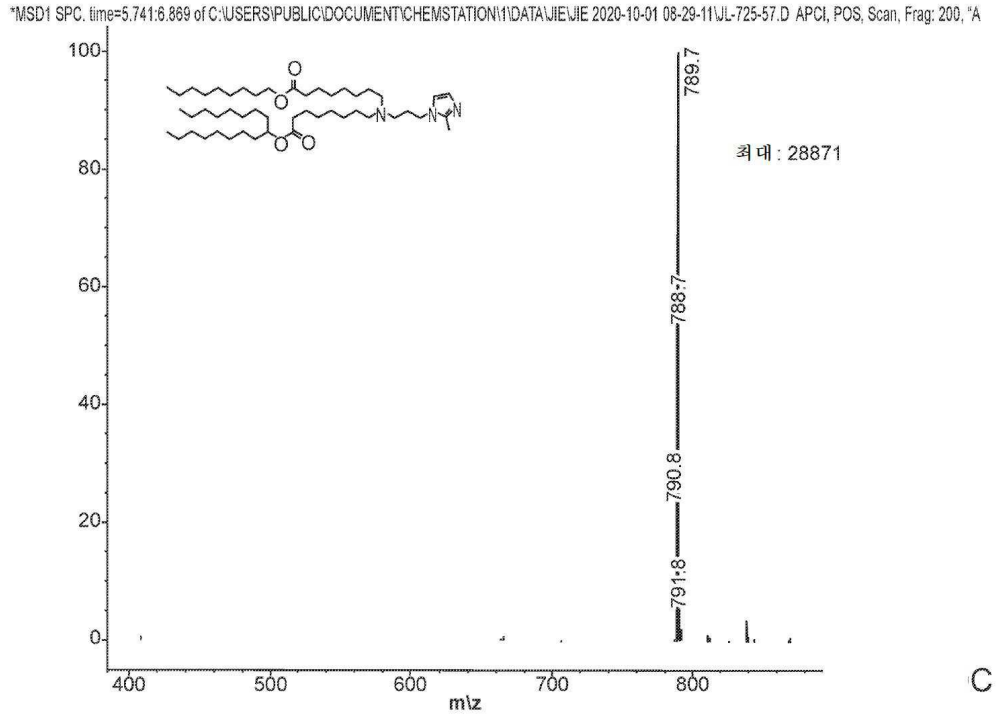
A

도면30b

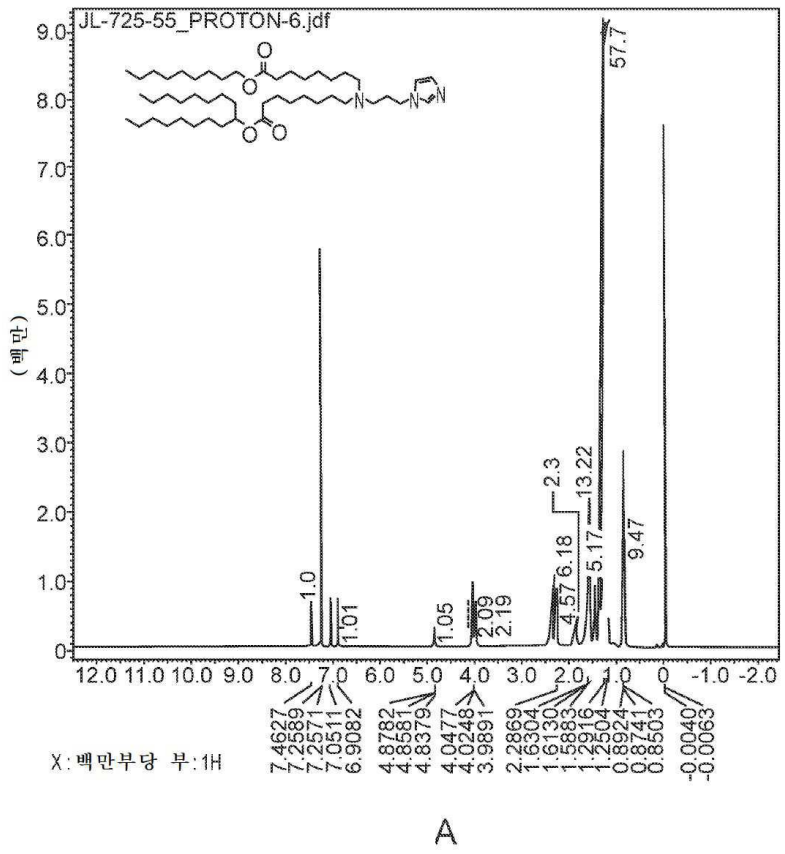


B

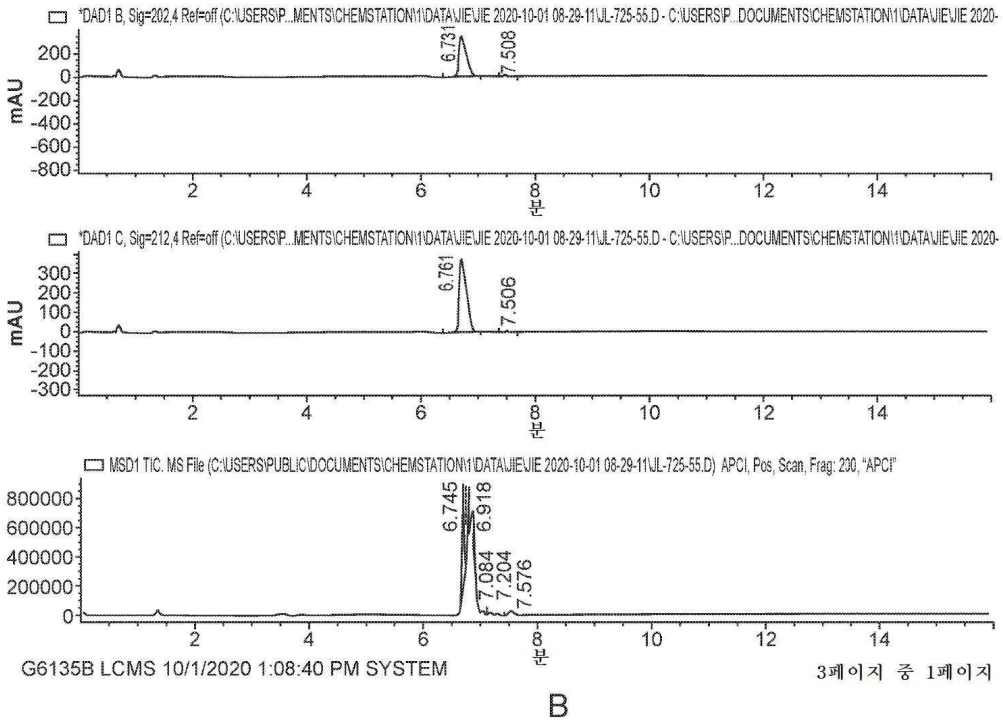
도면30c


*MSD1 SPC, time=5.741:6.869 of C:\USERS\PUBLIC\DOCUMENT\CHEMSTATION\1\DATA\JIE\JIE 2020-10-01 08-29-11\JL-725-57.D APCI, POS, Scan, Frag: 200, "A
최대 : 28871
789.7
788.7
790.8
791.8
m\z
C

도면31a

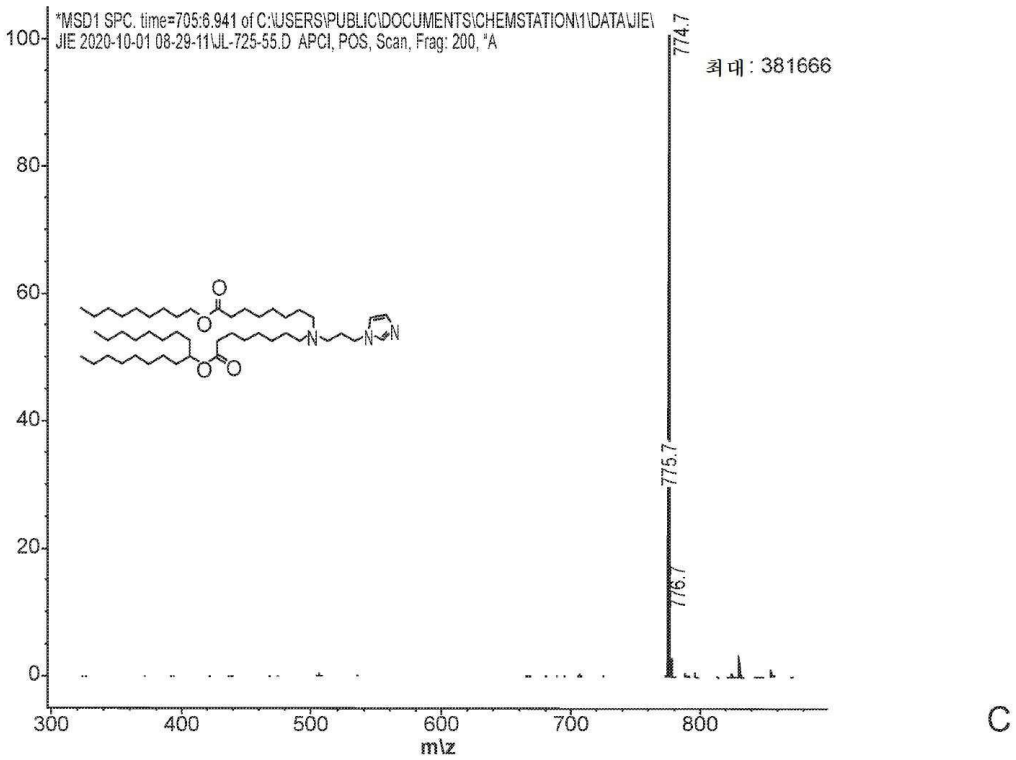

JL-725-55_PROTON-6.jdf
(백만)
X : 백만부당 부 : 1H
A

도면31b



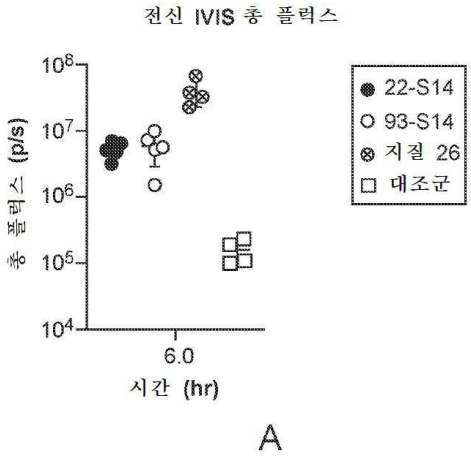
B

도면31c

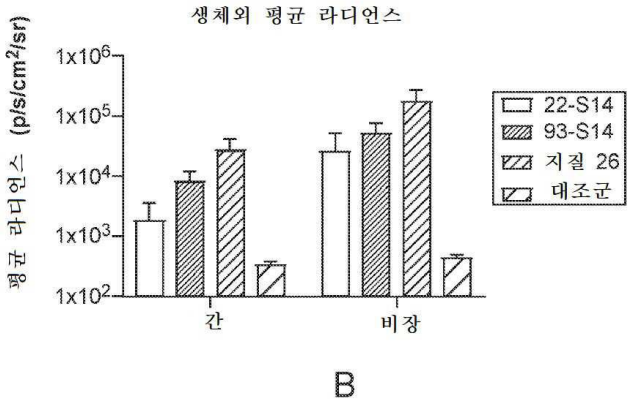


C

도면32

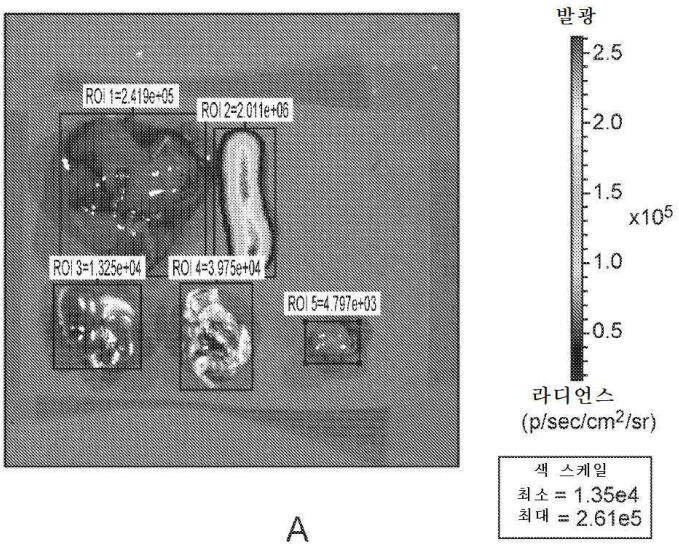

전신 IVIS 총 플럭스
총 플럭스 (p/s)
10^8
10^7
10^6
10^5
10^4
6.0
시간 (hr)
22-S14
93-S14
지질 26
대조군

A

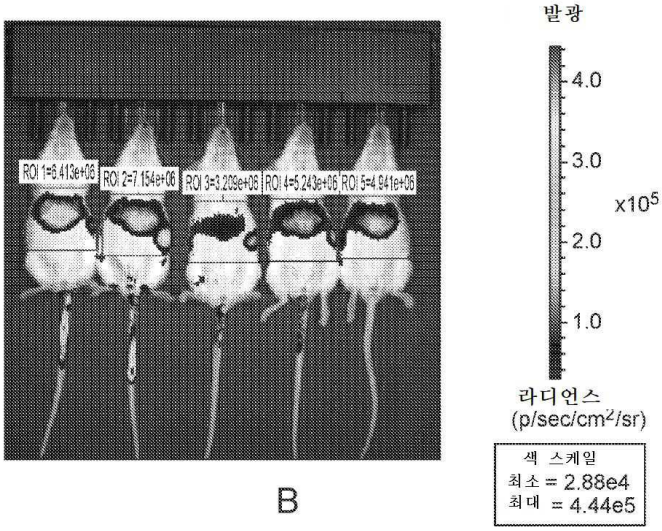

생체외 평균 라디언스
평균 라디언스 (p/s/cm²/sr)
1x10⁶
1x10⁵
1x10⁴
1x10³
1x10²
간
비장
22-S14
93-S14
지질 26
대조군

B

도면33

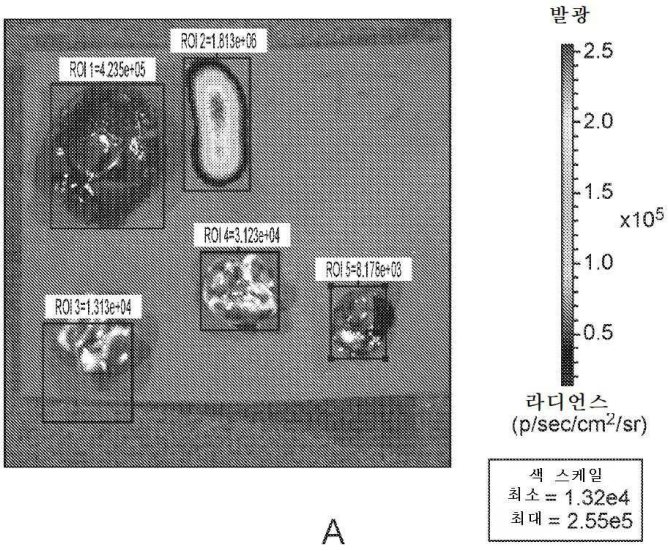

ROI 1=2.419e+05
ROI 2=2.011e+06
ROI 3=1.325e+04
ROI 4=3.975e+04
ROI 5=4.797e+03
발광
2.5
2.0
1.5
x10^5
1.0
0.5
라디언스
(p/sec/cm²/sr)
색 스케일
최소 = 1.35e4
최대 = 2.61e5

A

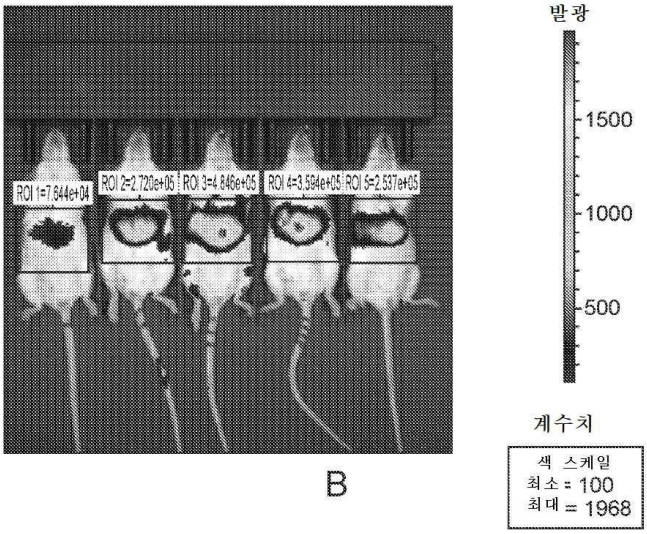

ROI 1=6.413e+03
ROI 2=7.154e+06
ROI 3=3.209e+06
ROI 4=5.243e+06
ROI 5=4.941e+06
발광
4.0
3.0
x10^5
2.0
1.0
라디언스
(p/sec/cm²/sr)
색 스케일
최소 = 2.88e4
최대 = 4.44e5

B

도면34

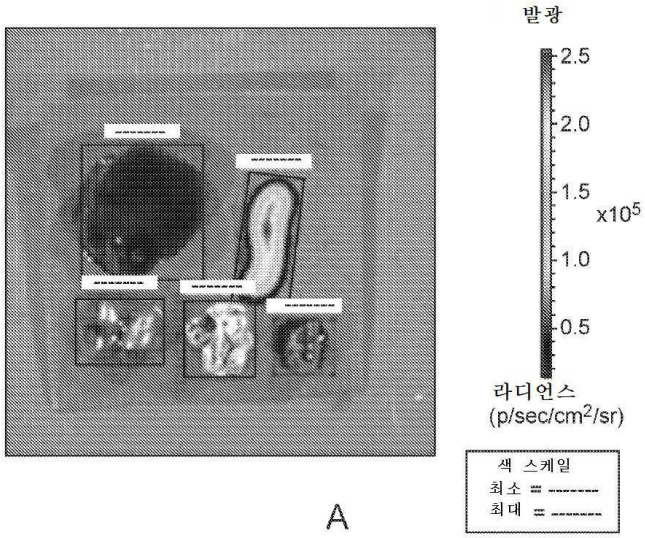

ROI 1=4.235e+05
ROI 2=1.813e+06
ROI 3=1.313e+04
ROI 4=3.123e+04
ROI 5=8.178e+03
발광
2.5
2.0
1.5
x10^5
1.0
0.5
라디언스
(p/sec/cm²/sr)
색 스케일
최소 = 1.32e4
최대 = 2.55e5

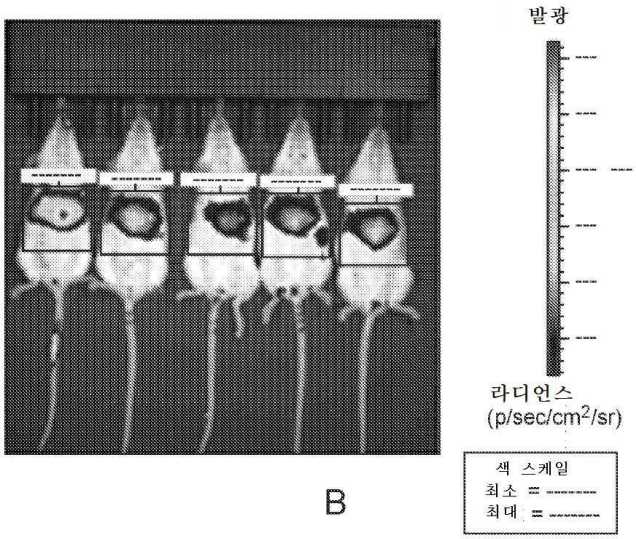
A


ROI 1=7.644e+04
ROI 2=2.720e+05
ROI 3=4.846e+05
ROI 4=3.594e+05
ROI 5=2.537e+05
발광
1500
1000
500
계수치
색 스케일
최소 = 100
최대 = 1968

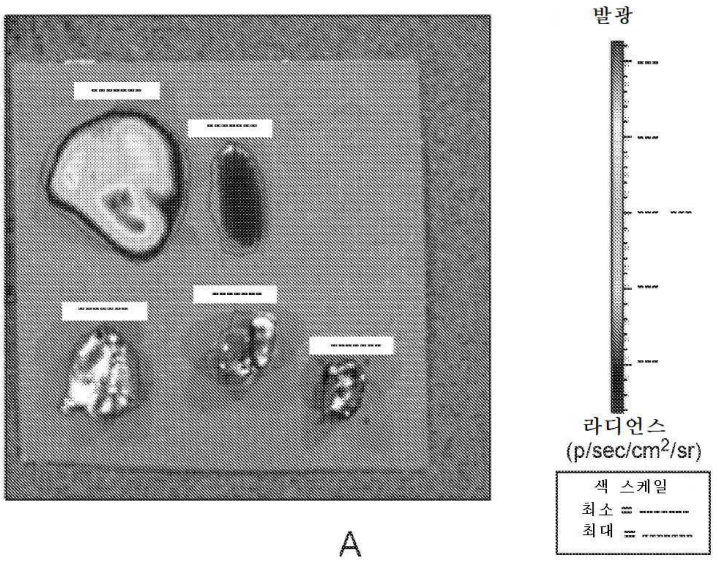
B

도면35

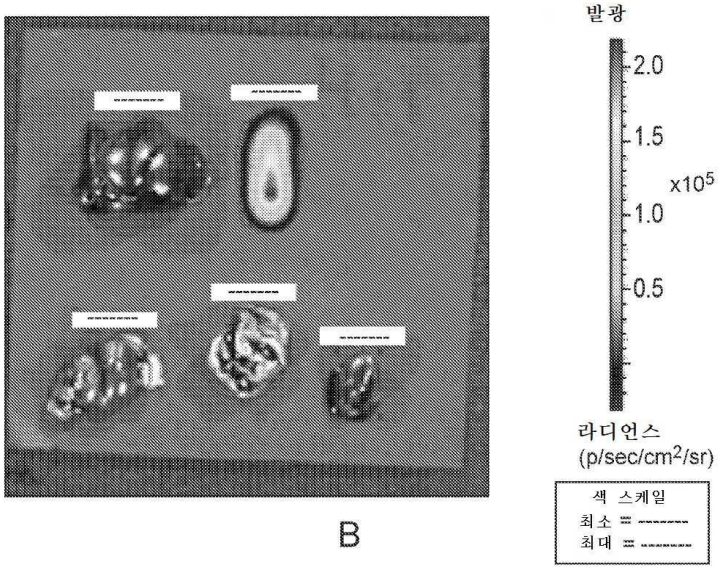

발광
2.5
2.0
1.5
x10^5
1.0
0.5
라디언스
(p/sec/cm²/sr)
색 스케일
최소 =
최대 =
A


발광
라디언스
(p/sec/cm²/sr)
색 스케일
최소 =
최대 =
B

도면36ab

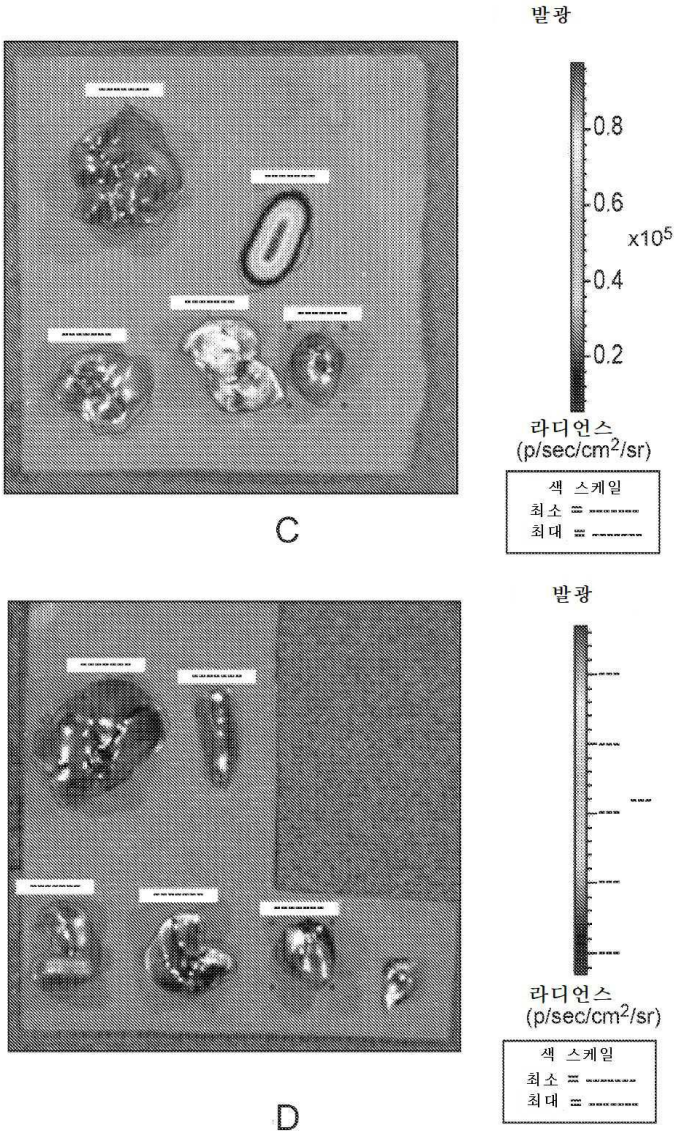

발광
라디언스
(p/sec/cm2/sr)
색 스케일
최소 =
최대 =

A

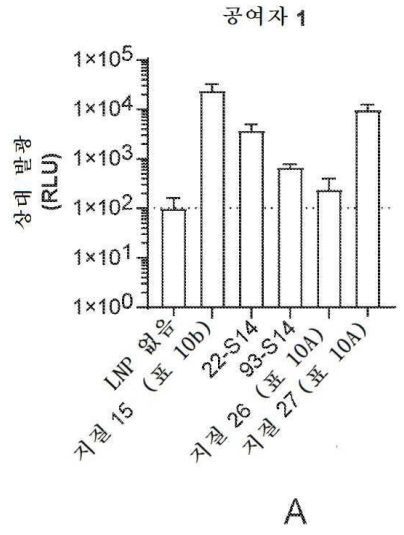

발광
2.0
1.5
x10^5
1.0
0.5
라디언스
(p/sec/cm2/sr)
색 스케일
최소 =
최대 =

B

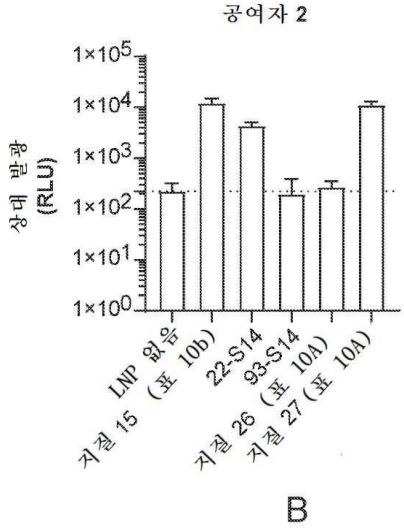
도면36cd


발광
0.8
0.6
x10^5
0.4
0.2
라디언스
(p/sec/cm²/sr)
색 스케일
최소 =
최대 =
C
발광
라디언스
(p/sec/cm²/sr)
색 스케일
최소 =
최대 =
D

도면37

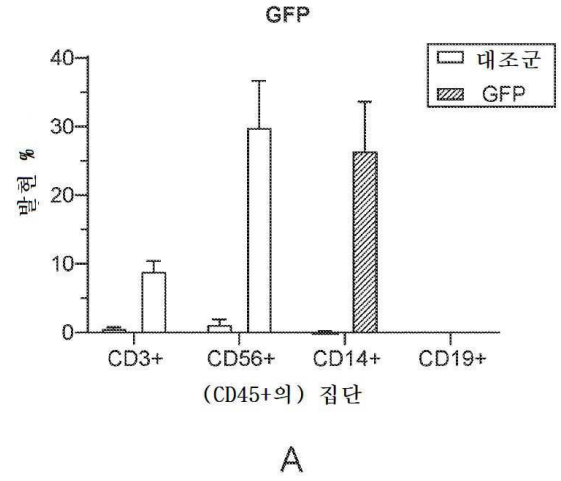

공여자 1
상대 발광 (RLU)
1×10^5
1×10^4
1×10^3
1×10^2
1×10^1
1×10^0
LNP 없음
지질 15
(포 10b)
22-S14
93-S14
지질 26
(포 10A)
지질 27
(포 10A)

A

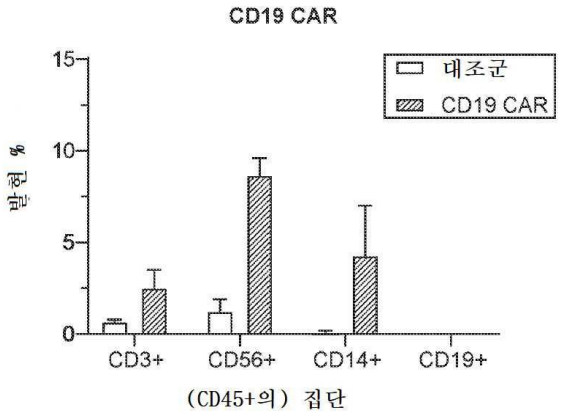

공여자 2
상대 발광 (RLU)
1×10^5
1×10^4
1×10^3
1×10^2
1×10^1
1×10^0
LNP 없음
지질 15
(포 10b)
22-S14
93-S14
지질 26
(포 10A)
지질 27
(포 10A)

B

도면38

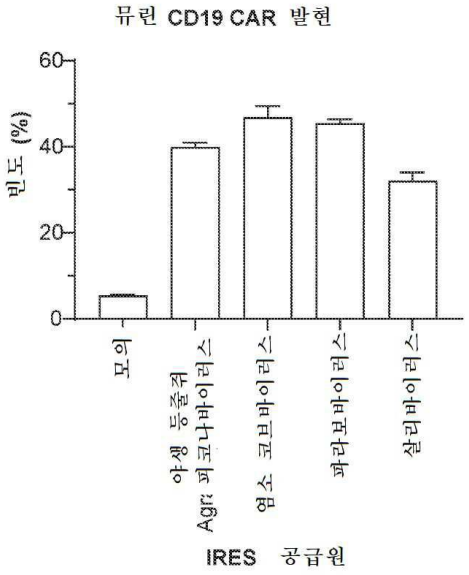

GFP
대조군
GFP
발현 %
40
30
20
10
0
CD3+
CD56+
CD14+
CD19+
(CD45+의) 집단

A

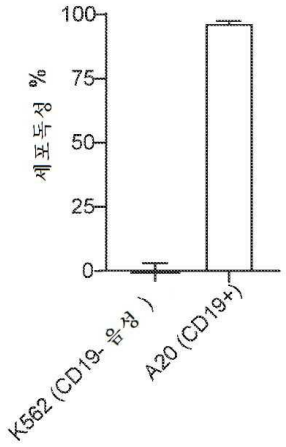

CD19 CAR
대조군
CD19 CAR
발현 %
15
10
5
0
CD3+
CD56+
CD14+
CD19+
(CD45+의) 집단

B

도면39

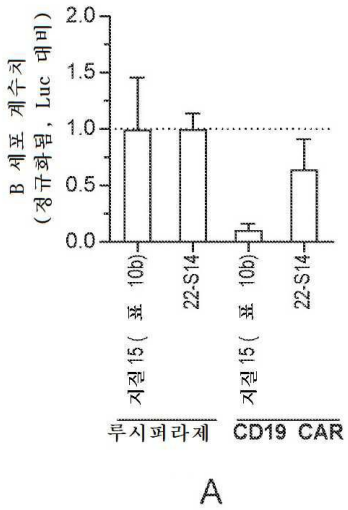

뮤린 CD19 CAR 발현
빈도 (%)
60
40
20
0
모의
아생 등줄쥐 피코나바이러스
염소 코브바이러스
파라보바이러스
살리바이러스
IRES 공급원

도면40



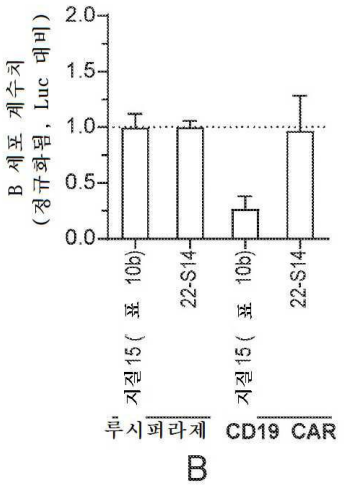
도면41ab

제8일 말초 혈액에서 B 세포 계수치



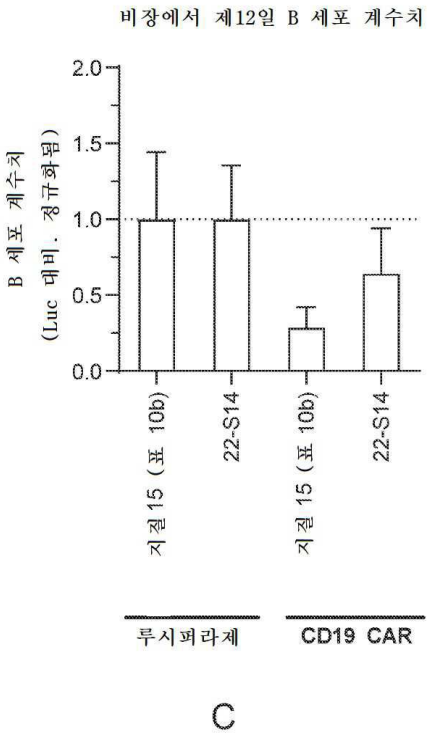
A

제12일 말초 혈액에서 B 세포 계수치



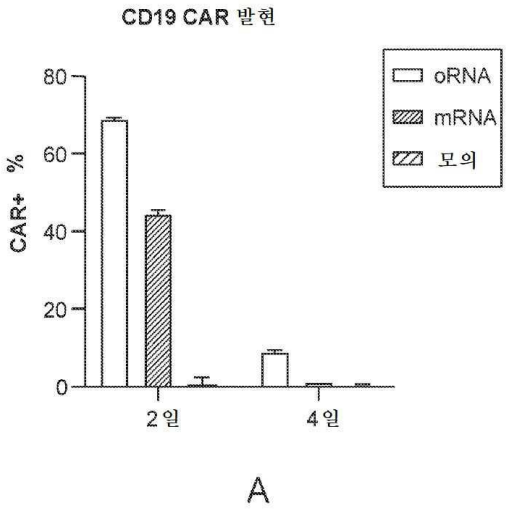
B

도면41c

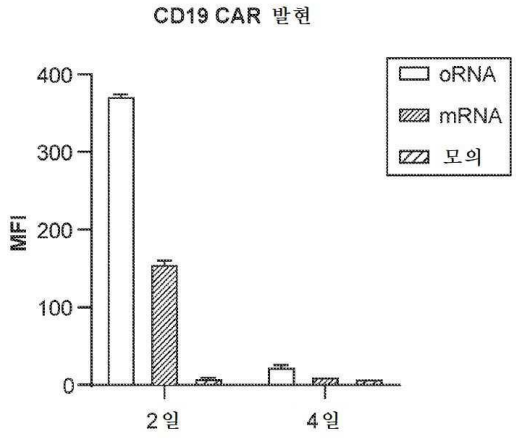

비장에서 제12일 B 세포 계수치
B 세포 계수치
(Luc 대비, 정규화됨)
2.0
1.5
1.0
0.5
0.0
지젤15 (표 10b)
22-S14
지젤15 (표 10b)
22-S14
루시퍼라제
CD19 CAR

C

도면42

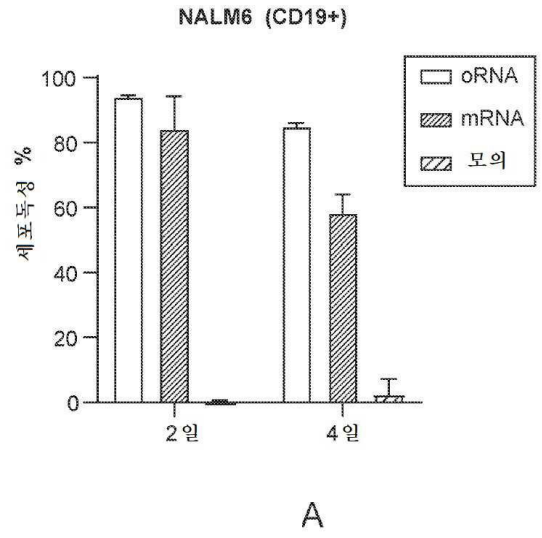

CD19 CAR 발현
oRNA
mRNA
모의
CAR+ %
80
60
40
20
0
2일
4일

A

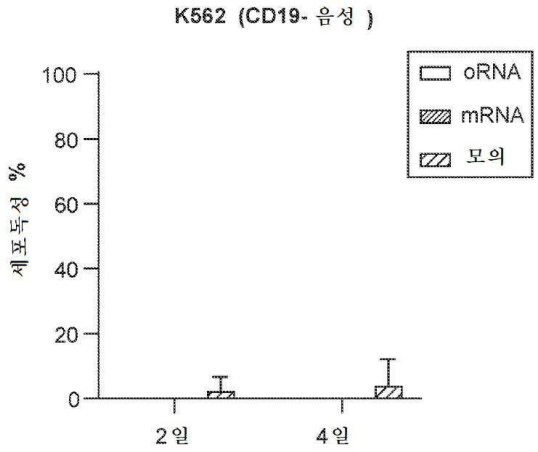

CD19 CAR 발현
oRNA
mRNA
모의
MFI
400
300
200
100
0
2일
4일

B

도면43

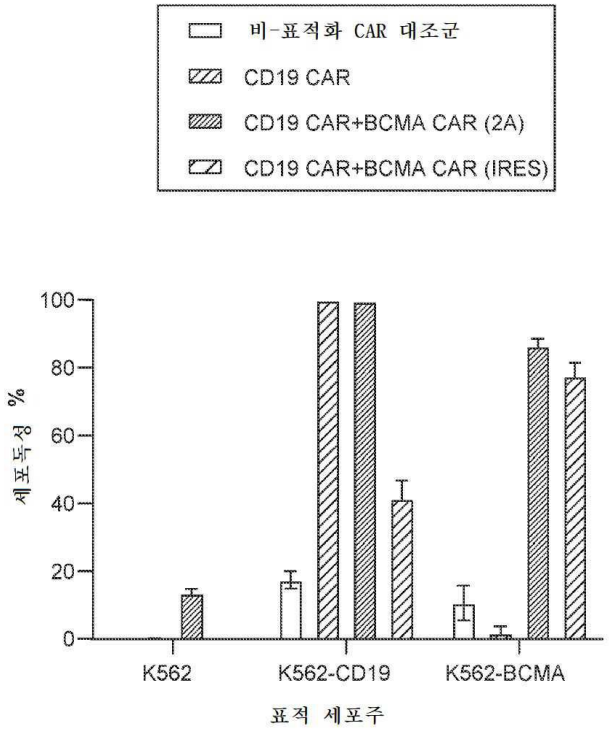

NALM6 (CD19+)
세포독성 %
100
80
60
40
20
0
2일
4일
oRNA
mRNA
모의

A

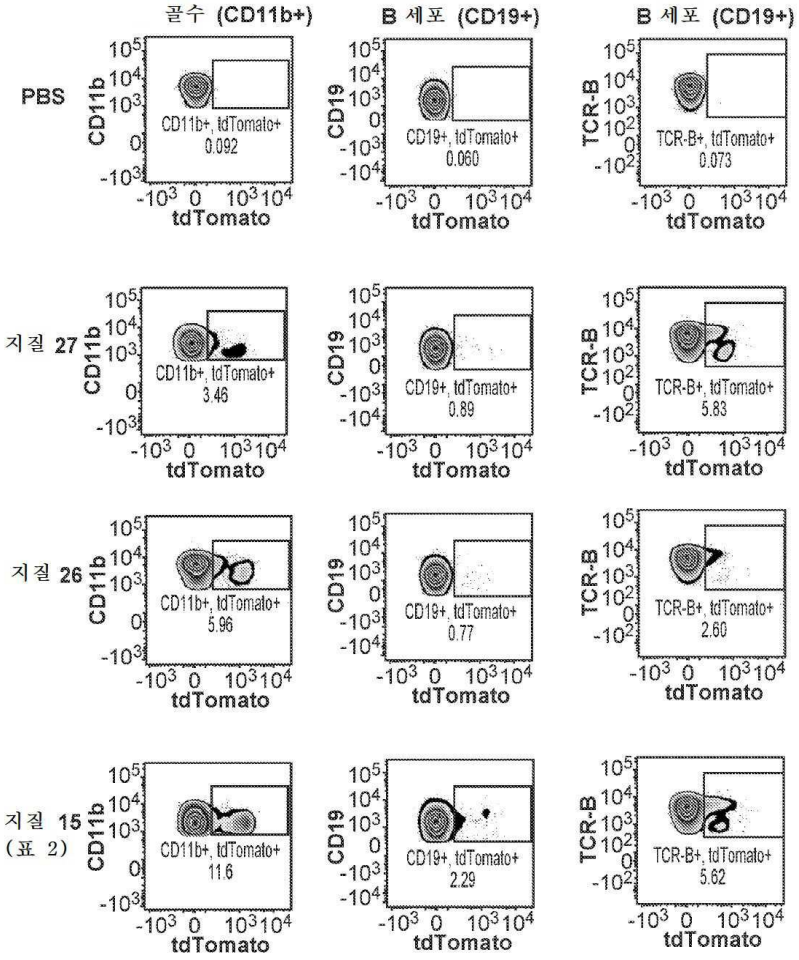

K562 (CD19- 음성)
세포독성 %
100
80
60
40
20
0
2일
4일
oRNA
mRNA
모의

B

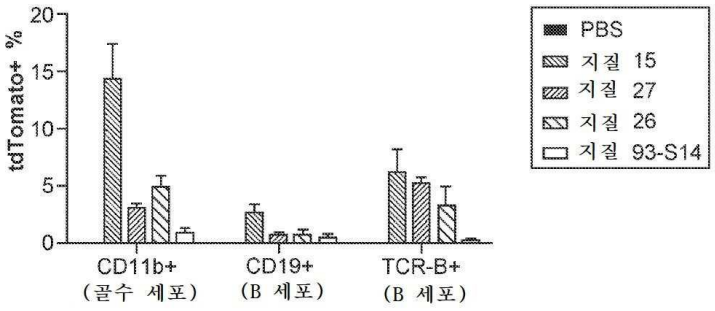
도면44


비-표적화 CAR 대조군
CD19 CAR
CD19 CAR+BCMA CAR (2A)
CD19 CAR+BCMA CAR (IRES)
100
80
60
40
20
0
세포독성 %
K562
K562-CD19
K562-BCMA
표적 세포주

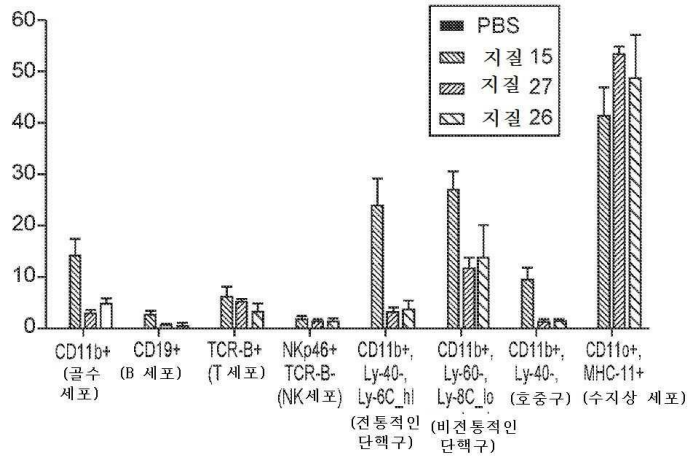
도면45a


골수 (CD11b+)
B 세포 (CD19+)
B 세포 (CD19+)
PBS
지질 27
지질 26
지질 15
(표 2)
CD11b
CD19
TCR-B
tdTomato
CD11b+, tdTomato+
0.092
CD19+, tdTomato+
0.060
TCR-B+, tdTomato+
0.073
CD11b+, tdTomato+
3.46
CD19+, tdTomato+
0.89
TCR-B+, tdTomato+
5.83
CD11b+, tdTomato+
5.96
CD19+, tdTomato+
0.77
TCR-B+, tdTomato+
2.60
CD11b+, tdTomato+
11.6
CD19+, tdTomato+
2.29
TCR-B+, tdTomato+
5.62

도면45bc

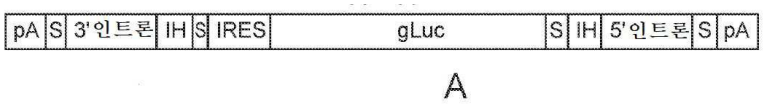

tdTomato+ %
20
15
10
5
0
PBS
지질 15
지질 27
지질 26
지질 93-S14
CD11b+
(골수 세포)
CD19+
(B 세포)
TCR-B+
(B 세포)

B

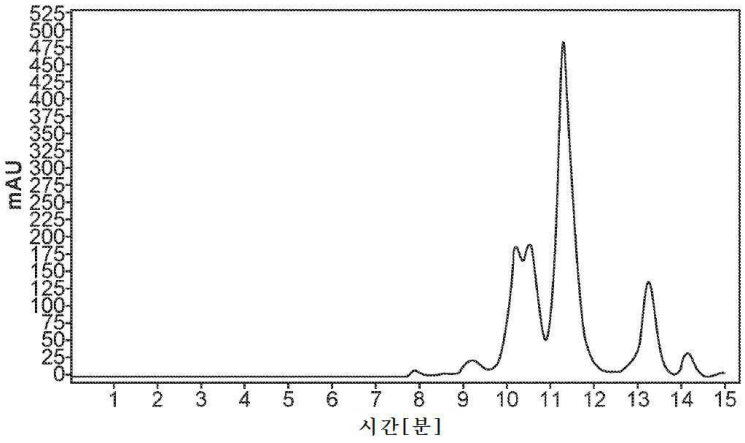

60
50
40
30
20
10
0
PBS
지질 15
지질 27
지질 26
CD11b+
(골수 세포)
CD19+
(B 세포)
TCR-B+
(T 세포)
NKp46+
TCR-B-
(NK세포)
CD11b+,
Ly-40-,
Ly-6C_hi
(전통적인 단핵구)
CD11b+,
Ly-60-,
Ly-8C_lo
(비전통적인 단핵구)
CD11b+,
Ly-40-,
(호중구)
CD11c+,
MHC-11+
(수지상 세포)

C

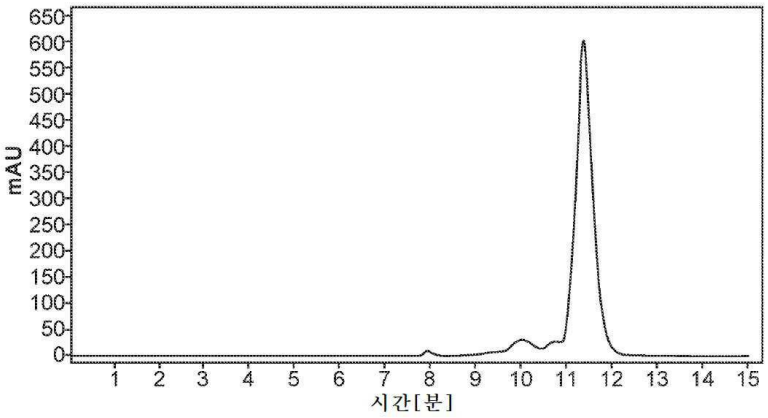
도면46ac


pA
S
3'인트론
IH
S
IRES
gLuc
S
IH
5'인트론
S
pA

A

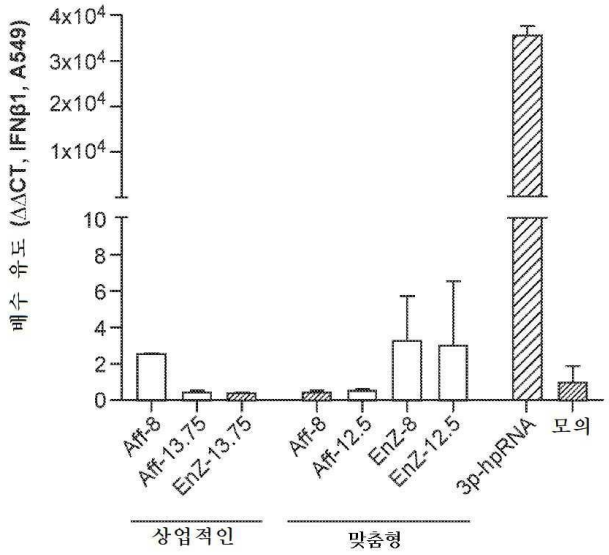

mAU
525
500
475
450
425
400
375
350
325
300
275
250
225
200
175
150
125
100
75
50
25
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
시간[분]

B


mAU
650
600
550
500
450
400
350
300
250
200
150
100
50
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
시간[분]

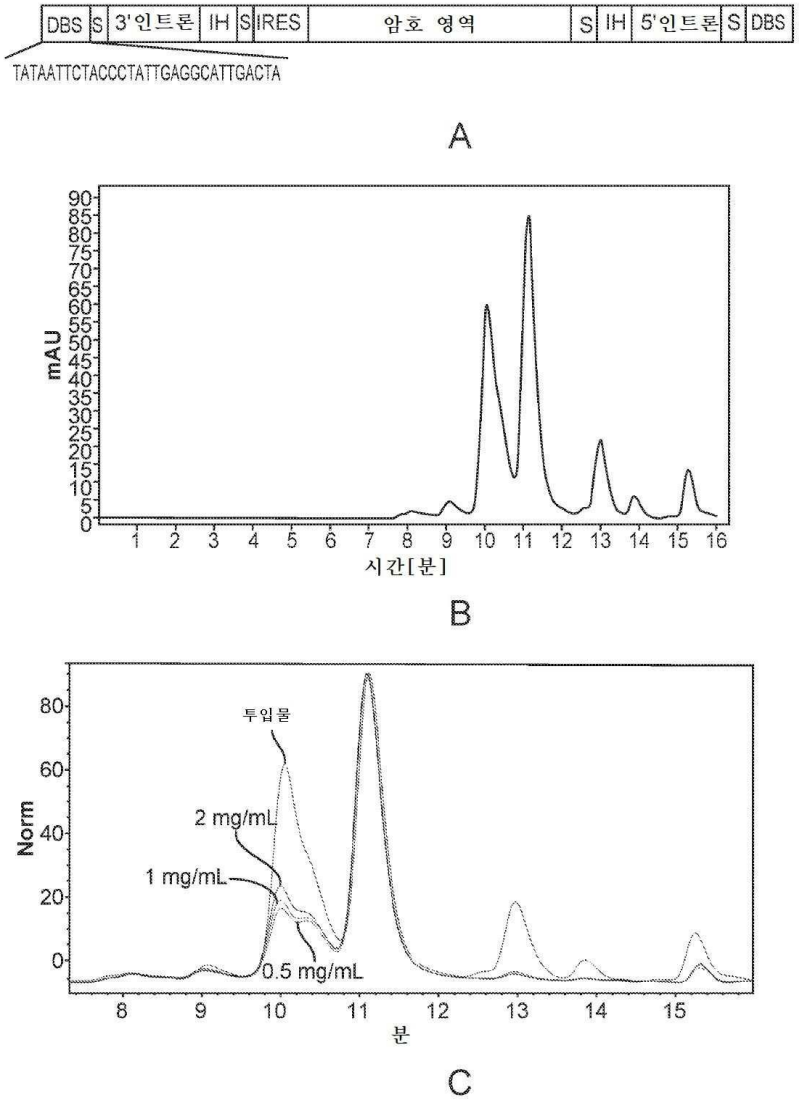
C

도면46d

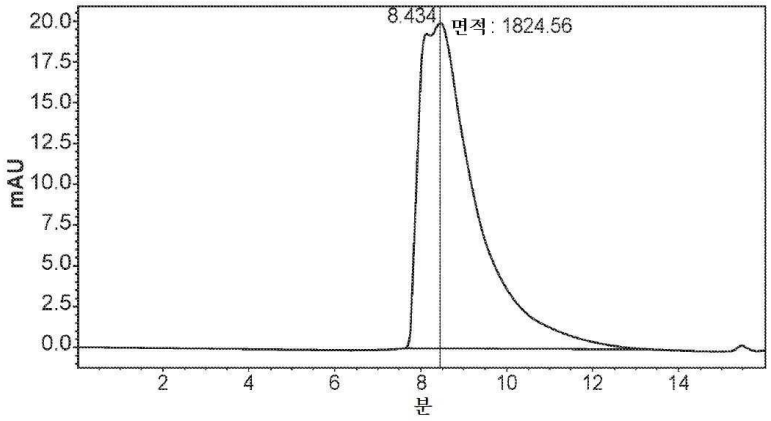

4x10^4
3x10^4
2x10^4
1x10^4
10
8
6
4
2
0
배수 유도 (ΔΔCT, IFNβ1, A549)
Aff-8
Aff-13.75
EnZ-13.75
Aff-8
Aff-12.5
EnZ-8
EnZ-12.5
3p-hpRNA
모의
상업적인
맞춤형

D

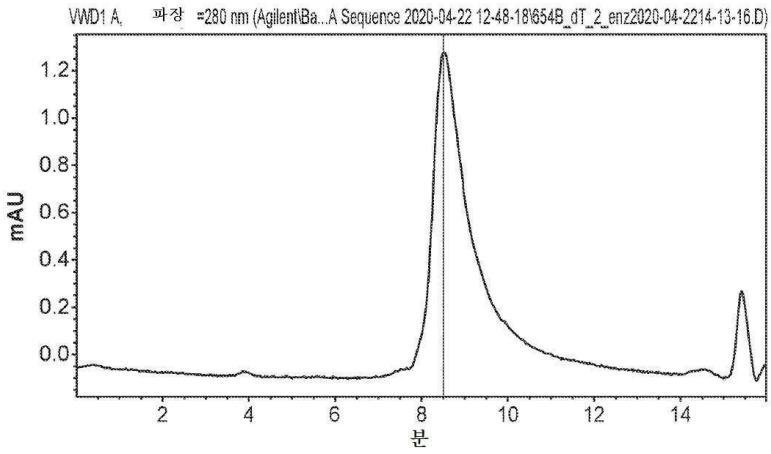
도면47


DBS
S
3'인트론
IH
S
IRES
암호 영역
S
IH
5'인트론
S
DBS
TATAATTCTACCCTATTGAGGCATTGACTA
A
mAU
시간[분]
B
Norm
투입물
2 mg/mL
1 mg/mL
0.5 mg/mL
분
C

도면48

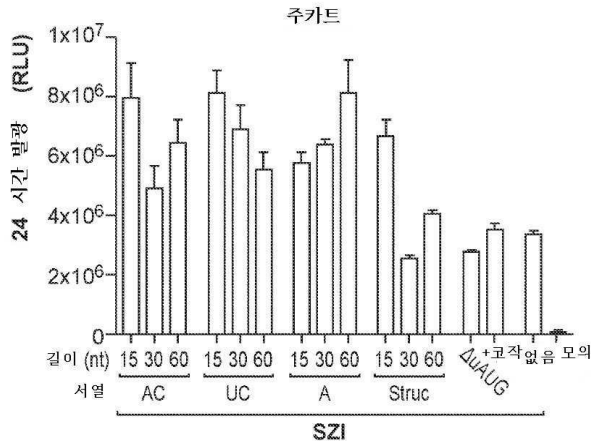

8.434
면적 : 1824.56
mAU
분

A

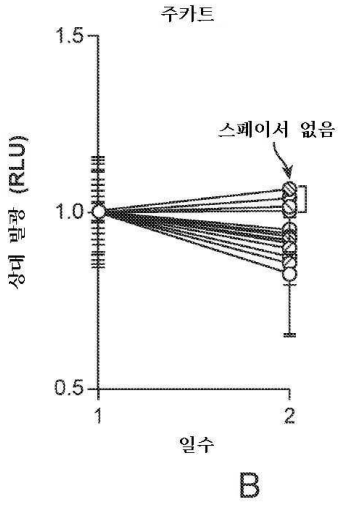

VWD1 A, 파장 =280 nm (Agilent\Ba...A Sequence 2020-04-22 12-48-18\654B_dT_2_enz2020-04-2214-13-16.D)
mAU
분

B

도면49

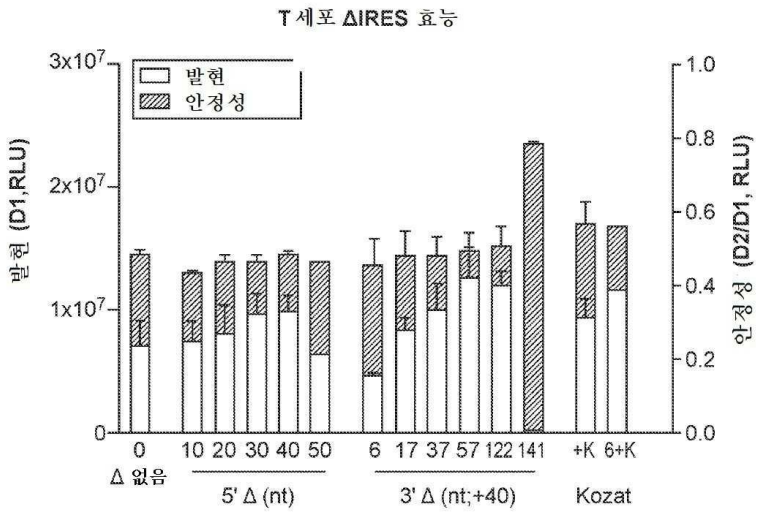

주카트
24 시간 발광 (RLU)
1x10^7
8x10^6
6x10^6
4x10^6
2x10^6
0
길이 (nt) 15 30 60 15 30 60 15 30 60 15 30 60 +코작 없음 모의
서열 AC UC A Struc ΔuAUG
SZI

A

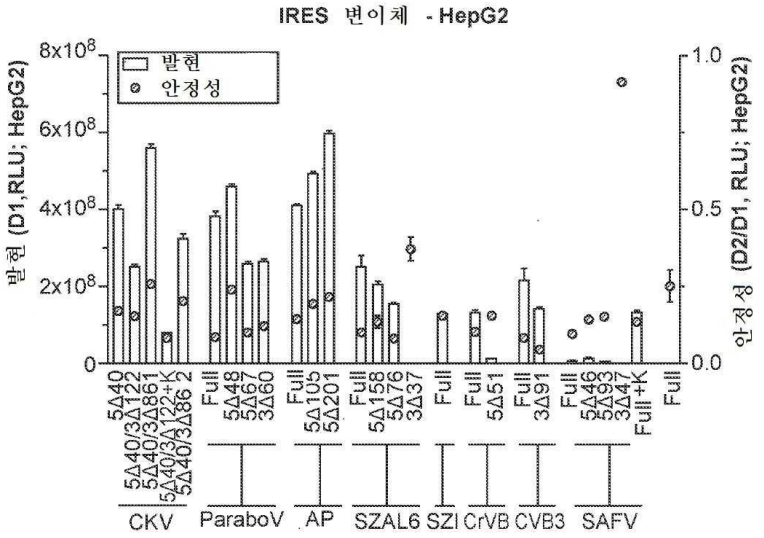

주카트
상대 발광 (RLU)
1.5
1.0
0.5
스페이서 없음
1
2
일수

B

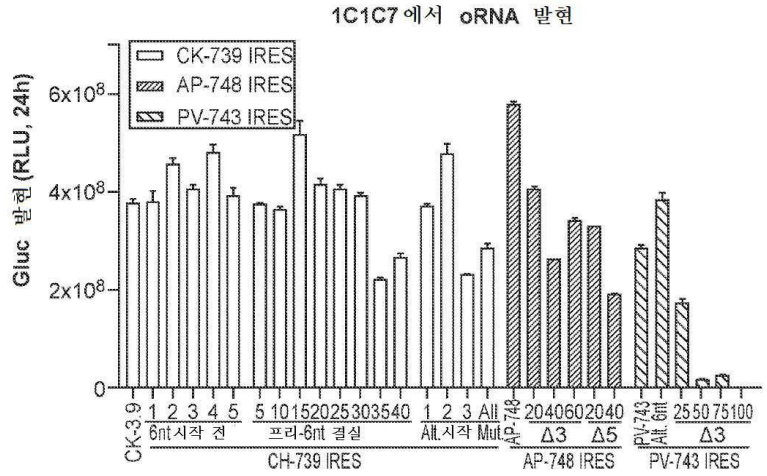
도면50


T 세포 ΔIRES 효능
발현 (D1,RLU)
안정성 (D2/D1, RLU)
발현
안정성
3x10^7
2x10^7
1x10^7
0
1.0
0.8
0.6
0.4
0.2
0.0
0 10 20 30 40 50 6 17 37 57 122 141 +K 6+K
Δ 없음
5' Δ (nt)
3' Δ (nt;+40)
Kozat

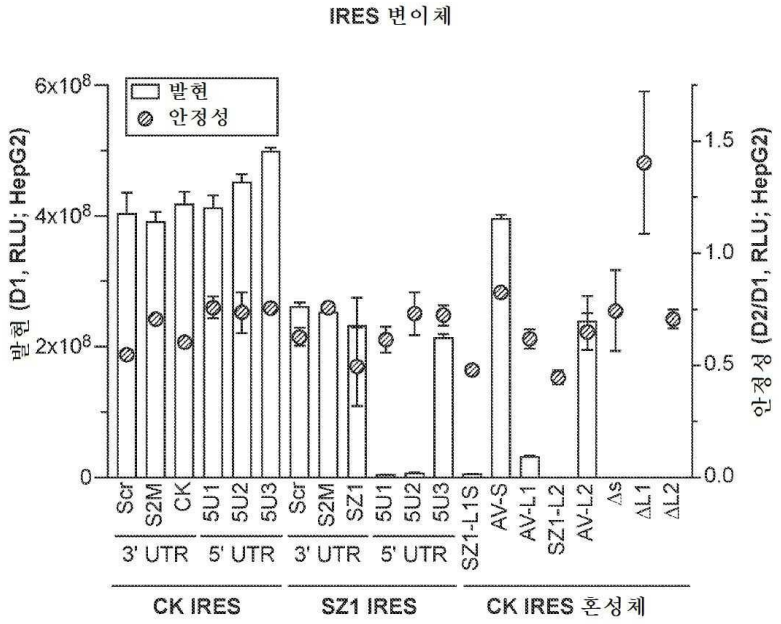
도면51


IRES 변이체 - HepG2
발현
안정성
발현 (D1,RLU; HepG2)
안정성 (D2/D1, RLU; HepG2)
8x10^8
6x10^8
4x10^8
2x10^8
0
1.0
0.5
0.0
5Δ40
5Δ40/3Δ122
5Δ40/3Δ861
5Δ40/3Δ122+K
5Δ40/3Δ86 2
Full
5Δ48
5Δ67
3Δ60
Full
5Δ105
5Δ201
Full
5Δ158
5Δ76
3Δ37
Full
Full
5Δ51
Full
3Δ91
Full
5Δ46
5Δ93
3Δ47
Full +K
Full
CKV
ParaboV
AP
SZAL6
SZI
CrVB
CVB3
SAFV

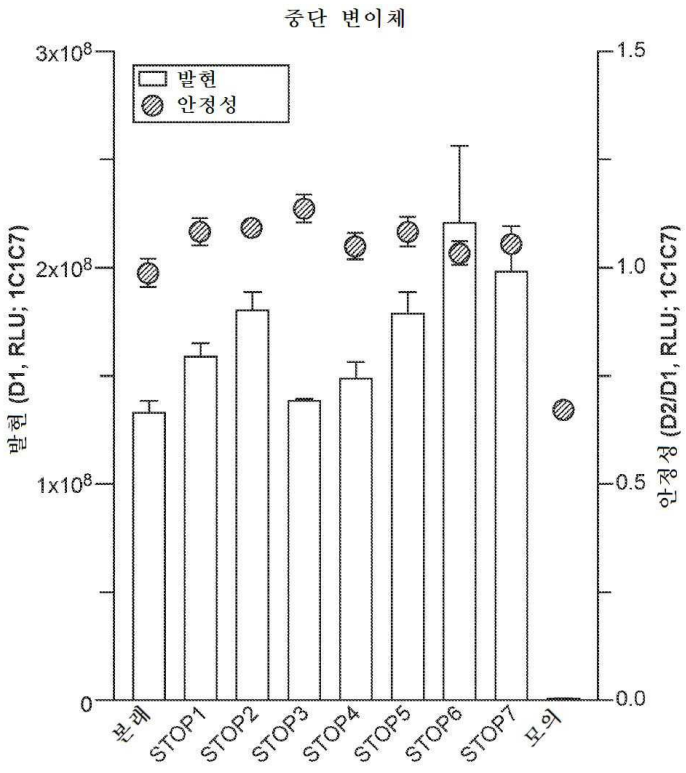
도면52


1C1C7 에서 oRNA 발현
CK-739 IRES
AP-748 IRES
PV-743 IRES
Gluc 발현 (RLU, 24h)
6x10^8
4x10^8
2x10^8
0
CK-3 9
1 2 3 4 5
5 10 15 20 25 30 35 40
1 2 3 All
Ap 748
20 40 60 20 40
PV-743
25 50 75 100
6nt시작 전
프리-6nt 결실
AUL시작
Mut
Δ3
Δ5
Δ3
CH-739 IRES
AP-748 IRES
PV-743 IRES

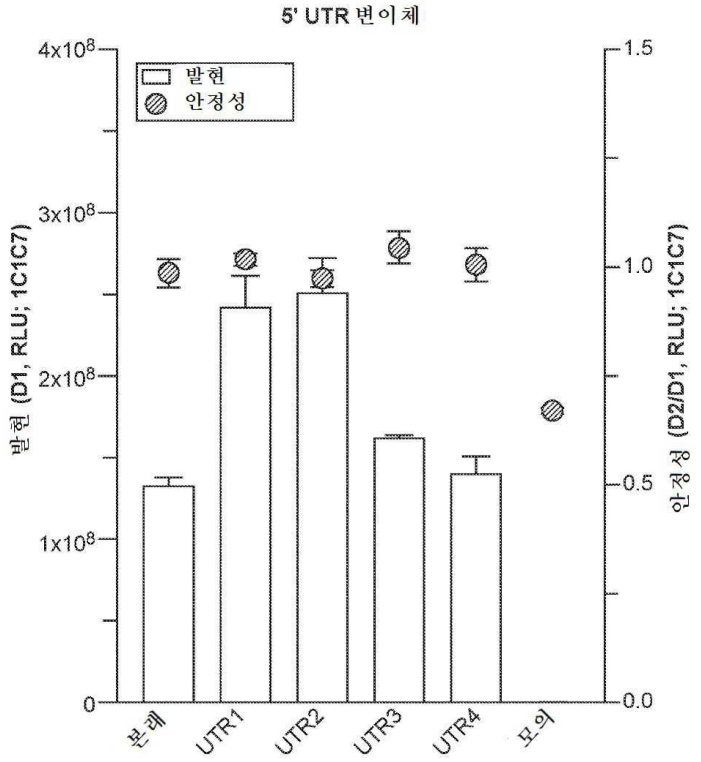
도면53


IRES 변이체
발현
안정성
발현 (D1, RLU; HepG2)
안정성 (D2/D1, RLU; HepG2)
6x10^8
4x10^8
2x10^8
0
1.5
1.0
0.5
0.0
Scr
S2M
CK
5U1
5U2
5U3
Scr
S2M
SZ1
5U1
5U2
5U3
SZ1-L1S
AV-S
AV-L1
SZ1-L2
AV-L2
ΔS
ΔL1
ΔL2
3' UTR
5' UTR
3' UTR
5' UTR
CK IRES
SZ1 IRES
CK IRES 혼성체

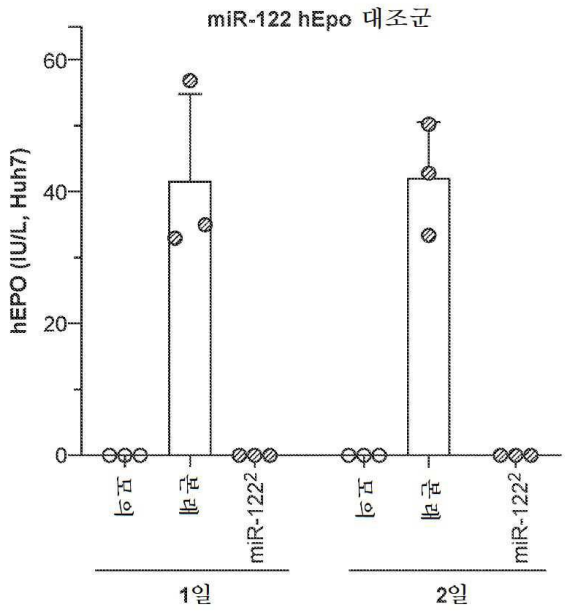
도면54


중단 변이체
발현
안정성
발현 (D1, RLU; 1C1C7)
안정성 (D2/D1, RLU; 1C1C7)
3x10^8
2x10^8
1x10^8
0
1.5
1.0
0.5
0.0
본래
STOP1
STOP2
STOP3
STOP4
STOP5
STOP6
STOP7
프리

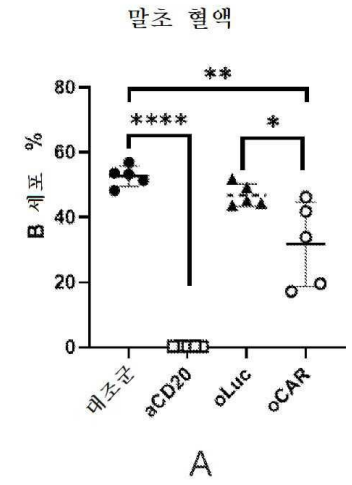
도면55


5' UTR 변이체
발현
안정성
4x10^8
3x10^8
2x10^8
1x10^8
0
1.5
1.0
0.5
0.0
발현 (D1, RLU; 1C1C7)
안정성 (D2/D1, RLU; 1C1C7)
본래
UTR1
UTR2
UTR3
UTR4
모의

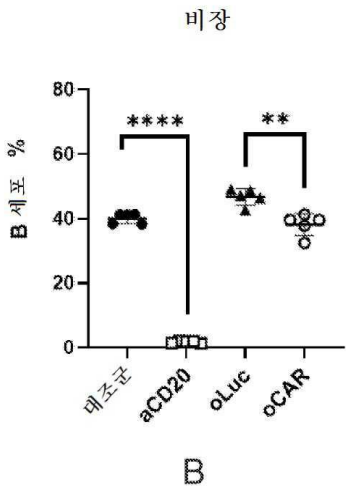
도면56


miR-122 hEpo 대조군
60
40
20
0
hEPO (IU/L, Huh7)
모의
본래
miR-122^2
모의
본래
miR-122^2
1일
2일

도면57

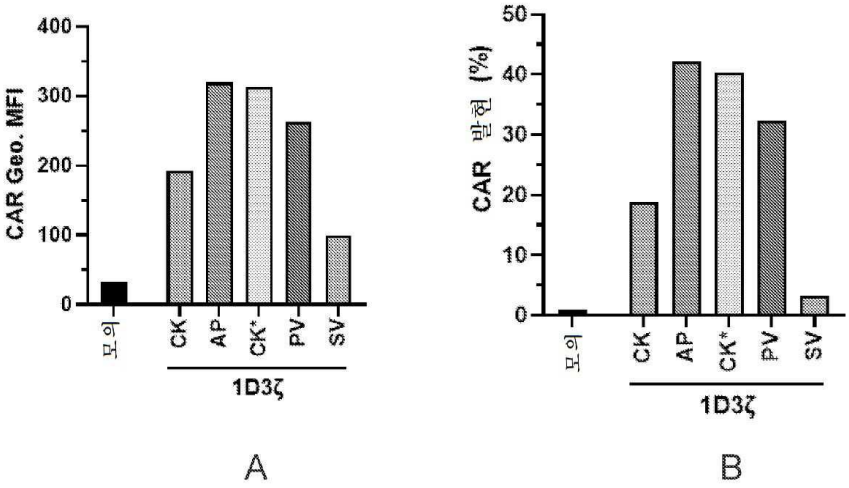

말초 혈액
**
****
*
B 세포 %
80
60
40
20
0
대조군
aCD20
oLuc
oCAR

A

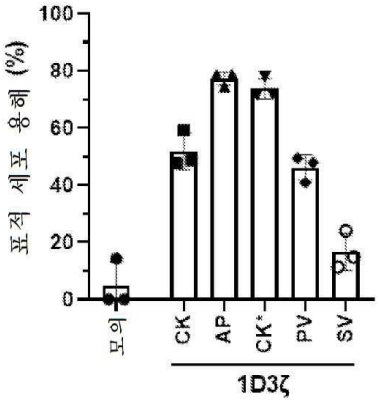

비장
****
**
B 세포 %
80
60
40
20
0
대조군
aCD20
oLuc
oCAR

B

도면58

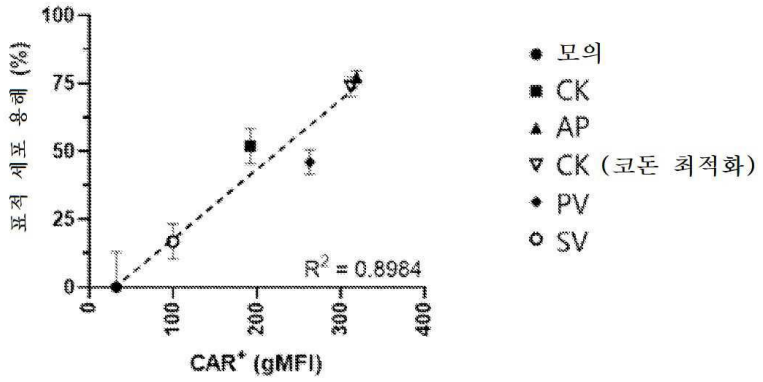


A B

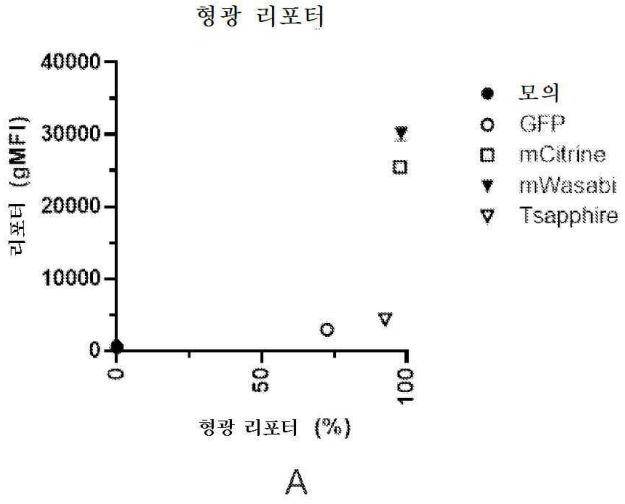


C

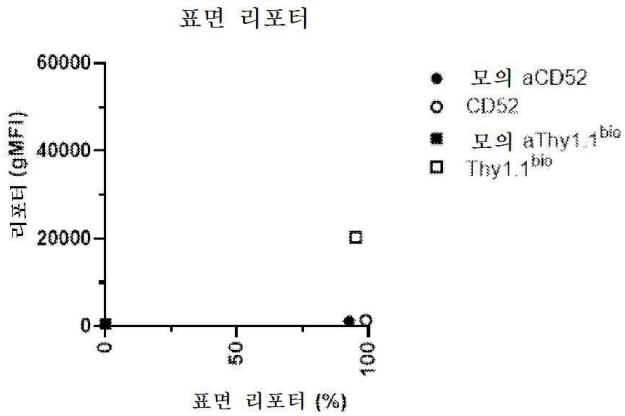
도면59



도면60

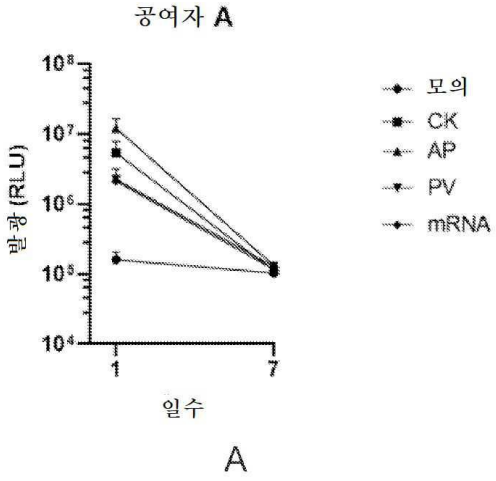

형광 리포터
40000
30000
20000
10000
0
리포터 (gMFI)
0
50
100
형광 리포터 (%)
모의
GFP
mCitrine
mWasabi
Tsapphire

A

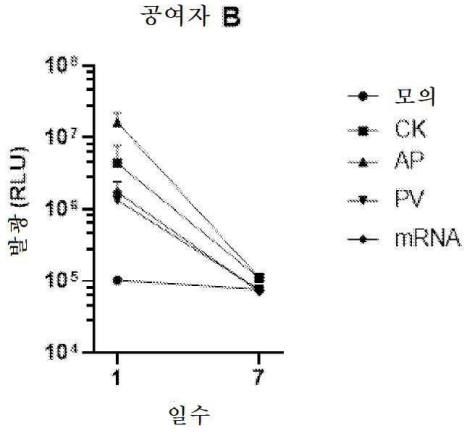

표면 리포터
60000
40000
20000
0
리포터 (gMFI)
0
50
100
표면 리포터 (%)
모의 aCD52
CD52
모의 aThy1.1bio
Thy1.1bio

B

도면61ab

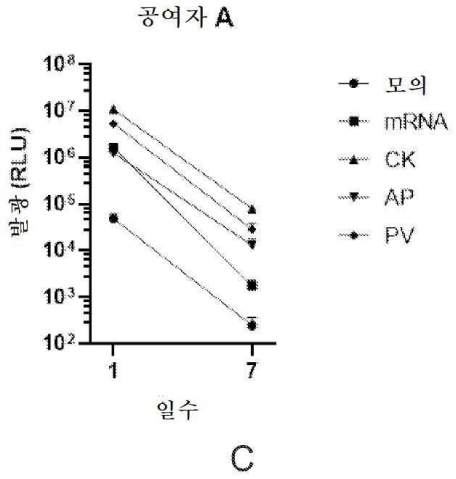

공여자 A
발광 (RLU)
10^8
10^7
10^6
10^5
10^4
1
7
일수
모의
CK
AP
PV
mRNA

A

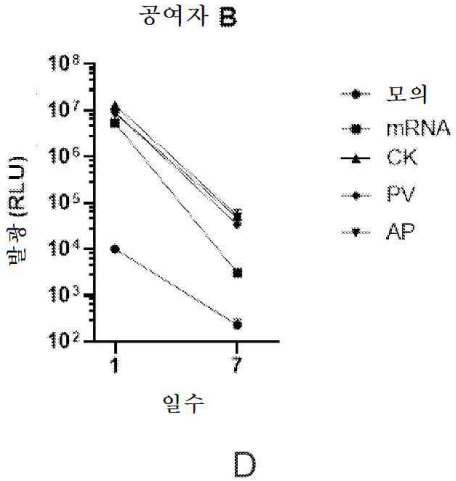

공여자 B
발광 (RLU)
10^8
10^7
10^6
10^5
10^4
1
7
일수
모의
CK
AP
PV
mRNA

B

도면61cd

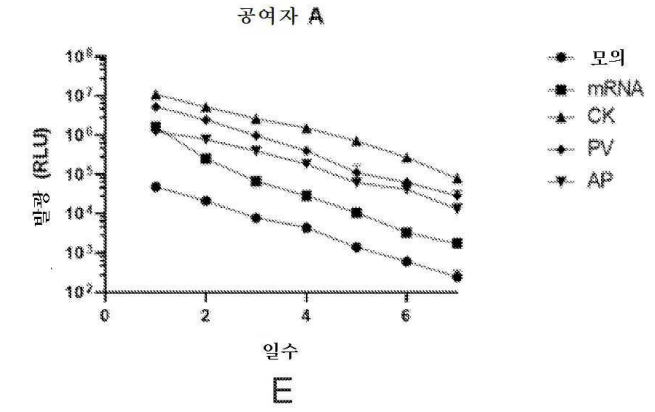

공여자 A
발광 (RLU)
10^8
10^7
10^6
10^5
10^4
10^3
10^2
1
7
일수
모의
mRNA
CK
AP
PV

C

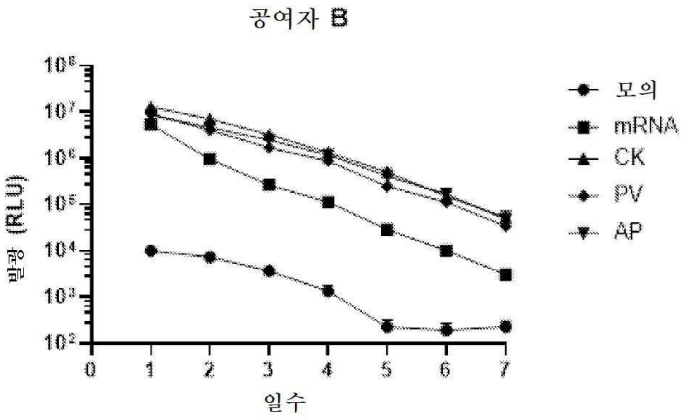

공여자 B
발광 (RLU)
10^8
10^7
10^6
10^5
10^4
10^3
10^2
1
7
일수
모의
mRNA
CK
PV
AP

D

도면61ef

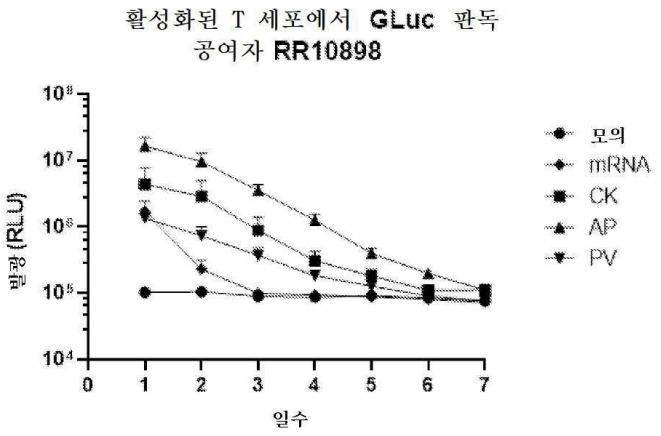


E

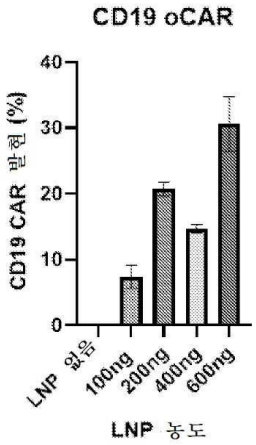


F

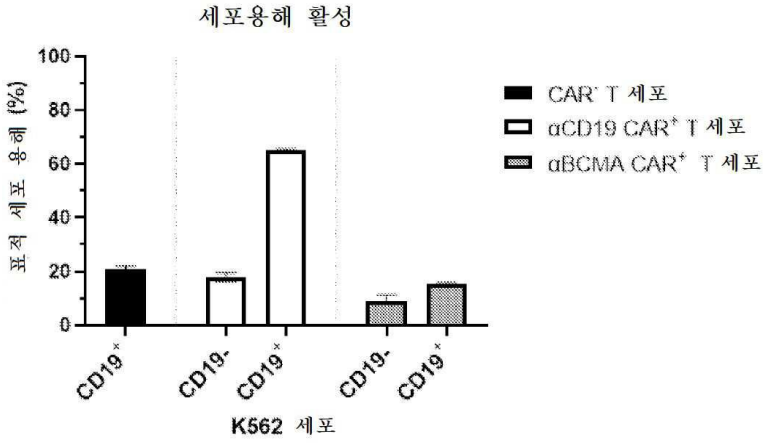
도면61g



도면62

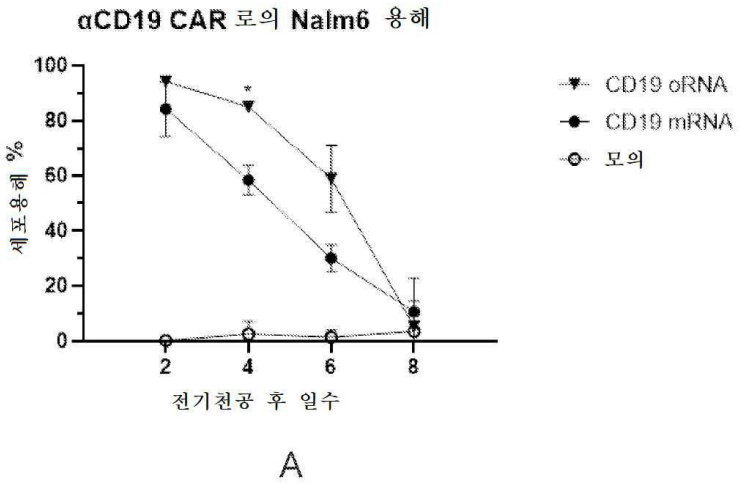

CD19 oCAR
CD19 CAR 발현 (%)
40
30
20
10
0
LNP 없음
100ng
200ng
400ng
600ng
LNP 농도

A

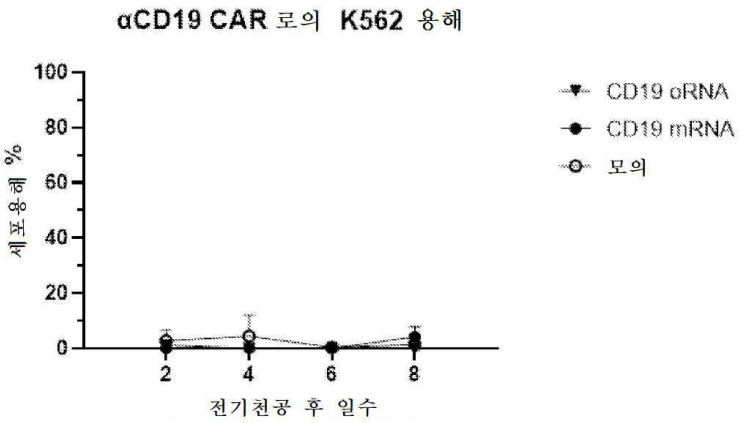

세포용해 활성
표적 세포 용해 (%)
100
80
60
40
20
0
CD19+
CD19-
CD19+
CD19-
CD19+
K562 세포
CAR- T 세포
αCD19 CAR+ T 세포
αBCMA CAR+ T 세포

B

도면63

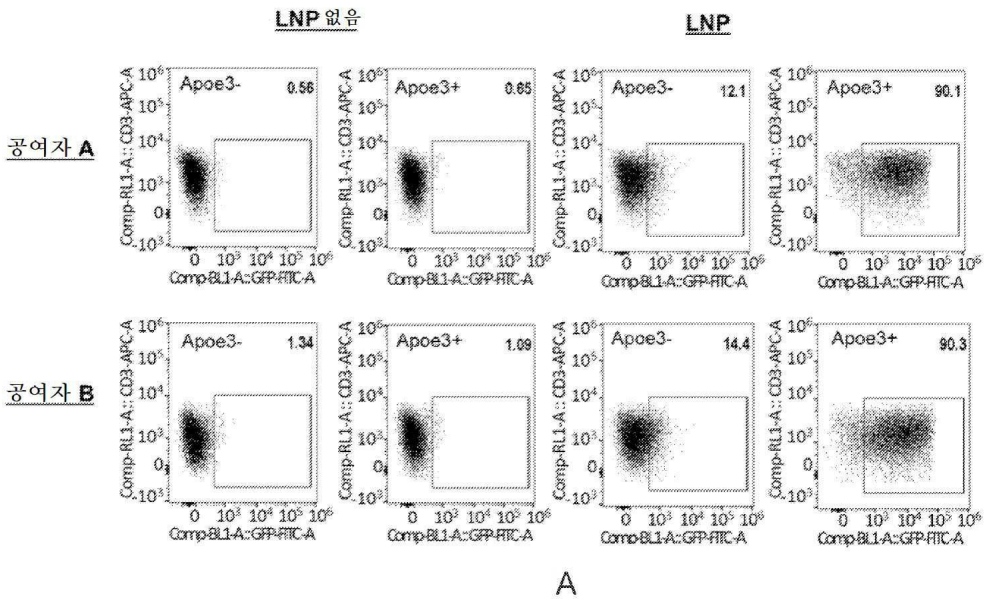


A



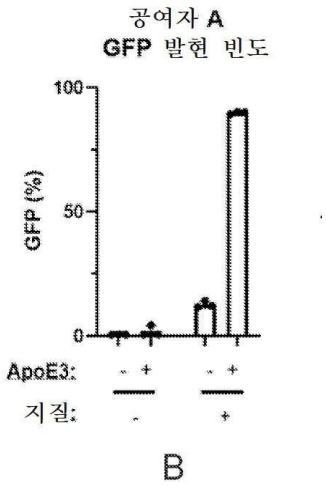
B

도면64a

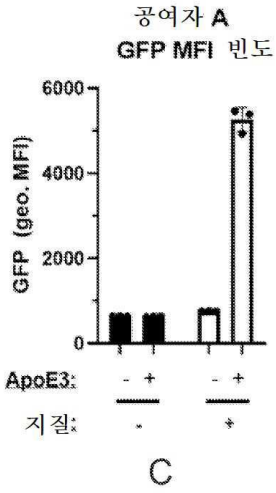


A

도면64bc

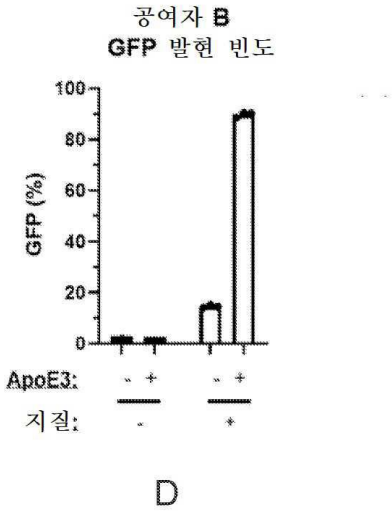

공여자 A
GFP 발현 빈도
100
50
0
GFP (%)
ApoE3: - + - +
지질: - +

B

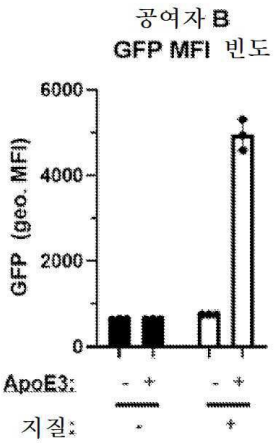

공여자 A
GFP MFI 빈도
6000
4000
2000
0
GFP (geo. MFI)
ApoE3: - + - +
지질: - +

C

도면64de



D



E

서 열 목 록

SEQUENCE LISTING

<110> ORNA THERAPEUTICS, INC.

<120> CIRCULAR RNA COMPOSITIONS AND METHODS

<130> WO/2021/189059

<140> PCT/US2021/023540

<141> 2021-03-22

<150> US 62/992,518

<151> 2020-03-20

<160> 793

<170> PatentIn version 3.5

<210> 1

<211> 585

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Encephalomyocarditis virus sequence"

<400> 1

ccccctctc cctccccccc taacgttact ggccgaagcc gcttggaata aggccggtgt 60

gcgtttgtct atatgttatt ttccaccata ttgccgtctt ttggcaatgt gagggcccgg 120

aaacctggcc ctgtcttctt gacgagcatt cctaggggtc tttcccctct cgccaaagga 180

atgcaaggtc tgttgaatgt cgtgaaggaa gcagttcctc tggaagcttc ttgaagacaa 240

acaacgtctg tagcgaccct ttgcaggcag cggaaccccc cacctggcga caggtgcctc 300

tgcggccaaa agccacgtgt ataagataca cctgcaaagg cggcacaacc ccagtgccac 360

gttgtgagtt ggatagttgt ggaaagagtc aaatggctct cctcaagcgt attcaacaag 420

gggctgaagg atgcccagaa ggtaccccat tgtatgggat ctgatctggg gcctcggtgc 480

acatgcttta catgtgttta gtcgaggtta aaaaacgtct aggccccccg aaccacgggg 540

acgtggtttt cctttgaaaa acacgatgat aatatggcca caacc 585

<210> 2

<211> 568

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Encephalomyocarditis virus sequence"

<400> 2

ctccccctcc ccccccttac tatactggcc gaagccactt ggaataaggc cggtgtgcgt 60

ttgtctacat gctattttct accgcattac cgtcttatgg taatgtgagg gtccagaacc 120

tgaccctgtc ttcttgacga acactcctag gggtctttcc cctctcgaca aaggagtgta 180

aggtctgttg aatgtcgtga aggaagcagt tcctctggaa gcttcttaaa gacaaacaac 240

gtctgtagcg acccttttgca ggcagcggaa cccccacct ggtgacaggt gcctctgcgg 300

ccaaaagcca cgtgtataag atacacctgc aaaggcggca caaccccagt gccacgttgt 360

gagttggata gttgtggaaa gagtcaaatg gctctcctca agcgtattca acaaggggct 420

gaaggatgcc cagaaggtac cccattgtat gggatctgat ctggggcctc ggtgcacgtg 480

ctttacacgt gttgagtcga ggtgaaaaaa cgtctaggcc ccccgaacca cggggacgtg 540

gttttccttt gaaaaccacg attacaat 568

<210> 3

<211> 696

<212> DNA

<213> Unknown

<

220><221> source

<223> /note="Description of Unknown:

Encephalomyocarditis virus sequence"

<400> 3

ttgccagtct gctcgatatc gcaggctggg tccgtgacta cccactcccc ctttcaacgt 60

gaaggctacg atagtgccag ggcgggtact gccgtaagtg ccaccccaaa caacaacaac 120

aaaacaaact ccccctcccc ccccttacta tactggccga agccacttgg aataaggccg 180

gtgtgcgttt gtctacatgc tattttctac cgcattaccg tcttatggta atgtgagggt 240

ccagaacctg accctgtctt cttgacgaac actcctaggg gtctttcccc tctcgacaaa 300

ggagtgtaag gtctgttgaa tgtcgtgaag gaagcagttc ctctggaagc ttcttaaaga 360

caaacaacgt ctgtagcgac cctttgcagg cagcggaacc ccccacctgg tgacaggtgc 420

ctctgcggcc aaaagccacg tgtataagat acacctgcaa aggcggcaca accccagtgc 480

cacgttgtga gttggatagt tgtggaaaga gtcaaatggc tctcctcaag cgtattcaac 540

aaggggctga aggatgccca gaaggtaccc cattgtatgg gatctgatct ggggcctcgg 600

tgcacgtgct ttacacgtgt tgagtcgagg tgaaaaaacg tctaggcccc ccgaaccacg 660

gggacgtggt tttcctttga aaaccacgat tacaat 696

<210> 4

<211> 697

<212> DNA

<213> Unknown

<

220><221> source

<223> /note="Description of Unknown:

Encephalomyocarditis virus sequence"

<400> 4

ttgccagtct gctcgatatc gcaggctggg tccgtgacta cccactcccc ctttcaacgt 60

gaaggctacg atagtgccag ggcgggtact gccgtaagtg ccaccccaaa acaacaacaa 120

ccccccctct ccctcctccc cccctaacgt tactggccga agccgcttgg aataaggccg 180

gtgtgcgttt gtctatatgt tattttccac catattgccg tcttttggca atgtgagggc 240

ccggaaacct ggccctgtct tcttgacgag cattcctagg ggtctttccc ctctcgccaa 300

aggaatgcaa ggtctgttga atgtcgtgaa ggaagcagtt cctctggaag cttcttgaag 360

acaaacaacg tctgtagcga ccctttgcag gcagcggaac cccccacctg gcgacaggtg 420

cctctgcggc caaaagccac gtgtataaga tacacctgca aaggcggcac aaccccagtg 480

ccacgttgtg agttggatag ttgtggaaag agtcaaatgg ctctcctcaa gcgtattcaa 540

caaggggctg aaggatgccc agaaggtacc ccattgtatg ggatctgatc tggggcctcg 600

gtgcacatgc tttacatgtg tttagtcgag gttaaaaaac gtctaggccc cccgaaccac 660

ggggacgtgg ttttcctttg aaaaacacga tgataat 697

<210> 5

<211> 562

<212> DNA

<213> Unknown

<

220><221> source

<223> /note="Description of Unknown:

Encephalomyocarditis virus sequence"

<400> 5

cccccccccc aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta 60

tatgttattt tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc 120

tgtcttcttg acgagcattc ctaggggtct ttcccctctc gccaaaggaa tgcaaggtct 180

gttgaatgtc gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt 240

agcgaccctt tgcaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa 300

gccacgtgta taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg 360

gatagttgtg gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga 420

tgcccagaag gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgctttac 480

atgtgtttag tcgaggttaa aaaacgtcta ggccccccga accacgggga cgtggttttc 540

ctttgaaaaa cacgatgata at 562

<210> 6

<211> 156

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Picobirnavirus sp. sequence"

<400> 6

gtaaattaaa tgctatttac aaaatttaaa cagaaaggag agatgttatg aaccggtttt 60

acaaggtttc atacatcgaa aatagcacta cctggggcag ccgacacact aacatcgtct 120

gtttaaccag aagtgttact gaaaggaggt tattta 156

<210> 7

<211> 340

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hepatitis C virus sequence"

<400> 7

acctgcccct aataggggcg acactccgcc atgaatcact cccctgtgag gaactactgt 60

cttcacgcag aaagcgtcta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120

ccccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg 180

aagactgggt cctttcttgg ataaacccac tctatgcccg gacatttggg cgtgcccccg 240

caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300

tgcttgcgag tgccccggga ggtctcgtag accgtgcatc 340

<210> 8

<211> 315

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human Cosavirus sequence"

<400> 8

ctacaagctt tgtgtaaaca aacttttgtt tggcttttct caagcttctc tcacatcagg 60

ccccaaagat gtcctgaagg taccccgtgt atctgaggat gagcaccatc gactacccgg 120

acctgcaaaa ttttgcaaac gcatgtggta tcccagcccc ctcctctcgg ggaggggggct 180

ttgctcactc agcacaggat ctgatcagga gatccacctc cggtgcttta caccggggcg 240

tggatttaaa aattgcccaa ggcctggcgc acaacctagg ggactaggtt ttccttatat 300

tttaaagctg tcaat 315

<210> 9

<211> 182

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human Cosavirus sequence"

<400> 9

gtcttaggac gacgcatgtg gtatcccagc ccccgcctac attggcgggg gcttttgaag 60

caccagacac tggatctgat caggaggagg gtagctgctt tacagcccct cttaaaaatt 120

gcccaaggtc cggccaccca acctagggga ctaggttttc cttttatttt taaattgtca 180

tt 182

<210> 10

<211> 759

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human Cosavirus sequence"

<400> 10

acatggggga gactgcatgt ggcagtcttg aaacgtgtgg tttgacgtct accttatatg 60

gcagtgggtg gagtactgca aagatgtcac cgtgctttac acggtttttg aaccccacac 120

cggctgtttg acgctcgtag ggcagcaggt ttattttcat taaaattctt actttctagc 180

tgcatgagtt ctattcatgc agacggagtg atactcccgt tccttcttgg acaggttgcc 240

tccacgccct ttgtggatct taaggtgacc aagtcactgg tgttggaggt gaagatagag 300

agtcctcttg ggaatgtcat gtggctgtgc caggggttgt agcgatgcca ttcgtgtgtg 360

cggatttcct ctcgtggtga cacgagcctc acaggccaaa agccccgtcc gaaaggaccc 420

gaatggtgga gtgaccctga ctccccccctg catagttttg tgattaggaa cttgaggaat 480

ttctgtcata aatctctatc acatcaggcc ccaaagatgt cctgaaggta ccctgtgtat 540

ctgaggatga gcaccaccga ctacccggac ttgcattagc agacacatgt ggttgcccag 600

ccccacctct tcagaggtgg ggctttgctc actcagcaca ggatctgatc aggagccccg 660

ctcgtgtgct ttacactcga cgcggggtta aaaattgccc aaggcctggc acaacaacct 720

aggggactag gttttcctat ttttgtaaat tatgtcaat 759

<210> 11

<211> 478

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Human rhinovirus sequence"

<400> 11

gtgacaatca gccagattgt taacggtcaa gcacttctgt ttccccggta cccttgtata 60

cgcttcaccc gaggcgaaaa gtgaggttat cgttatccgc aaagtgccta cgagaagcct 120

agtagcactt ttgaagccta tggctggtcg ctcaactgtt tacccagcag tagacctggc 180

agatgaggct agatgttccc caccagcgat ggtgatctag cctgcgtggc tgcctgcaca 240

ctctattgag tgtgaagcca gaaagtggac aaggtgtgaa gagcctattg tgctcacttt 300

gagtcctccg gcccctgaat gtggctaatc ctaaccccgt agctgttgca tgtaatccaa 360

catgtctgca gtcgtaatgg gcaactatgg gatggaacca actactttgg gtgtccgtgt 420

ttcttgtttt tctttatgct tgcttatggt gacaactgta gttattacat ttgttacc 478

<210> 12

<211> 628

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Human rhinovirus sequence"

<400> 12

ttaaaacagc ggatgggtat cccaccattc gacccattgg gtgtagtact ctggtactat 60

gtacctttgt acgcctgttt ctccccaacc acccttcctt aaaattccca cccatgaaac 120

gttagaagct tgacattaaa gtacaatagg tggcgccata tccaatggtg tctatgtaca 180

agcacttctg tttcccagga gcgaggtata ggctgtaccc actgccaaaa gcctttaacc 240

gttatccgcc aaccaactac gtaacagtta gtaccatctt gttcttgact ggacgttcga 300

tcaggtggat tttccctcca ctagtttggt cgatgaggct aggaattccc cacgggtgac 360

cgtgtcctag cctgcgtggc ggccaaccca gcttatgctg ggacgccctt ttaaggacat 420

ggtgtgaaga ctcgcatgtg cttggttgtg agtcctccgg cccctgaatg cggctaacct 480

taaccctaga gccttatgcc acgatccagt ggttgtaagg tcgtaatgag caattccggg 540

acgggaccga ctactttggg tgtccgtgtt tctcattttt cttcatattg tcttatggtc 600

acagcatata tatacatata ctgtgatc 628

<210> 13

<211> 618

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human rhinovirus sequence"

<400> 13

ttaaaactgg gagtgggttg ttcccactca ctccacccat gcggtgttgt actctgttat 60

tacggtaact ttgtacgcca gtttttccca cccttcccca taatgtaact tagaagtttg 120

tacaatatga ccaataggtg acaatcatcc agactgtcaa aggtcaagca cttctgtttc 180

cccggtcaat gaggatatgc tttacccaag gcaaaaacct tagagatcgt tatccccaca 240

ctgcctacac agagcccagt accatttttg atataattgg gttggtcgct ccctgcaaac 300

ccagcagtag acctggcaga tgaggctgga cattccccac tggcgacagt ggtccagcct 360

gcgtggctgc ctgctcaccc ttcttgggtg agaagcctaa ttattgacaa ggtgtgaaga 420

gccgcgtgtg ctcagtgtgc ttcctccggc ccctgaatgt ggctaacctt aaccctgcag 480

ccgttgccca taatccaatg ggtttgcggt cgtaatgcgt aagtgcggga tgggaccaac 540

tactttgggt gtccgtgttt cctgtttttc ttttgattgc attttatggt gacaatttat 600

agtgtataga ttgtcatc 618

<210> 14

<211> 610

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Human rhinovirus sequence"

<400> 14

ttaaaactgg gtacaggttg ttcccacctg tatcacccac gtggtgtggt gctcttgtat 60

tccggtacac ttgcacgcca gtttgccacc cctcacccgt cgtaacttag aagctaacaa 120

ctcgaccaac aggcggtggt aaaccatacc acttacggtc aagcactcct gtttccccgg 180

tatgcgagga atagactcct acagggttga agcctcaagt atcgttatcc gcattggtac 240

tacgcaaagc ttagtagtgc cttgaaagtc ccttggttgg tcgctccgct agtttccct 300

agtagacctg gcagatgagg caggacactc cccactggcg acagtggtcc tgcctgcgtg 360

gctgcctgcg cacccttagg ggtgcgaagc caagtgacag acaaggtgtg aagagccccg 420

tgtgctacca atgagtcctc cggcccctga atgcggctaa tccaacccca cagctattgc 480

acacaagcca gtgtgtatgt agtcgtaatg agcaattgtg ggacggaacc gactactttg 540

ggtgtccgtg tttcctttta ttcttatcat tctgcttatg gtgacaatac tgtgaaatag 600

tgttgttacc 610

<210> 15

<211> 614

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Human rhinovirus sequence"

<400> 15

taaaactgga tccaggttgt tcccacctgg atctcctatt gggagttgta ctctattatt 60

ccggtaattt tgtacgccag ttttatcttc cccctcccca attgtaactt agaaggttat 120

caatacgacc aataggtggt agttagccaa actaccaaag gtcaagcact tctgtttccc 180

cggtcaaagt tgatatgctc caacagggca aaaacaactg agatcgttat ccgcaaagtg 240

cctacgcaaa gcctagtaac acctttgaag atttatggtt ggtcgttccg ctatttccca 300

tagtagacct ggcagatgag gctagaaatc ccccactggc gacagtgctc tagcctgcgt 360

ggctgcctgc gcacccccttg ggtgcgaagc catacattgg acaaggtgtg aagagccccg 420

tgtgctcact ttgagtcctc cggcccctga atgtggctaa ccttaaccct gcagctagtg 480

catgtaatcc aacatgttgc tagtcgtaat gagtaattgc gggacgggac caactacttt 540

gggtgtccgt gtttcacttt ttccttttaa tattgcttat ggtgacaata tatatagcta 600

tatatattga cacc 614

<210> 16

<211> 624

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus A sequence"

<400> 16

ttcccctgca accattacgc ttactcgcat gtgcattgag tggtgcatgt gttgaacaaa 60

cagctacact cacatggggg cgggttttcc cgccctacgg cttctcgcga ggcccacccc 120

tccccttttct cccataacta cagtgctttg gtaggtaagc atcctgatcc cccgcggaag 180

ctgctcacgt ggcaactgtg gggacccaga caggttatca aaggcacccg gtctttccgc 240

cttcaggagt atccctgcta gcgaattcta gtagggctct gcttggtgcc aacctccccc 300

aaatgcgcgc tgcgggagtg ctcttcccca actcacccta gtatcctctc atgtgtgtgc 360

ttggtcagca tatctgagac gatgttccgc tgtcccagac cagtccagta atggacgggc 420

cagtgtgcgt agtcgtcttc cggcttgtcc ggcgcatgtt tggtgaaccg gtggggtaag 480

gttggtgtgc ccaacgcccg tactcagggg atacctcaag gcacccagga atgccaggga 540

ggtaccccgc ttcacagcgg gatctgaccc tggggtaaat gtctgcgggg ggtcttcttg 600

gcccacttct cagtactttt cagg 624

<210> 17

<211> 730

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus FHB sequence"

<400> 17

acatgggggg tctgcggacg gcttcggccc acccgcgaca agaatgccgt catctgtcct 60

cattacccgt attccttccc ttccccccgca accaccacgc ttactcgcgc acgtgttgag 120

tggcacgtgc gttgtccaaa cagctacacc cacacccttc ggggcgggtt tgtcccgccc 180

tcgggttcct cgcggaaccc ccccctccct ctctctcttt ctatccgccc tcacttccca 240

taactacagt gctttggtag gtgagcaccc tgacccccg cggaagctgc taacgtggca 300

actgtgggga tccaggcagg ttatcaaagg cacccggtct ttccgccttc aggagtatct 360

ctgccggtga attccggtag ggctctgctt ggtgccaacc tcccccaaat gcgcgctgcg 420

ggagtgctct tccccaactc atcttagtaa cctctcatgt gtgtgcttgg tcagcatatc 480

tgaggcgacg ttccgctgtc ccagaccagt ccagcaatgg acgggccagt gtgcgtagtc 540

gctttccggt tttccggcgc atgtttggcg aaacgctgag gtaaggttgg tgtgcccaac 600

gcccgtaatt tggtgatacc tcaagaccac ccaggaatgc cagggaggta ccccacttcg 660

gtgggatctg accctgggct aattgtctac ggtggttctt cttgcttcca cttctctttt 720

ttctggcatg 730

<210> 18

<211> 709

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus NG-J1 sequence"

<400> 18

tatggcaggc gggcttgtgg acggcttcgg cccacccaca gcaagaatgc catcatctgt 60

cctcaccccc aattttccct tttcttcccc tgcaaccatt acgcttactc gcatgtgcat 120

tgagtggtgc atgtgttgaa caaacagcta cactcacatg ggggcgggtt ttcccgccct 180

acggcctctc gcgaggccca ccccttccct ccccttataa ctacagtgct ttggtaggta 240

agcatcctga tcccccgcgg aagctgctca cgtggcaact gtggggaccc agacaggtta 300

tcaaaggcac ccggtctttc cgccttcagg agtatcccta ctagtgaatt ctagcggggc 360

tctgcttggt gccaacctcc cccaaatgcg cgctgcggga gtgctcttcc ccaactcacc 420

ctagtatcct ctcatgtgtg tgcttggtca gcatatctga gacgatgttc cgctgtccca 480

gaccagtcca gtaatggacg ggccagtgcg tgtagtcgtc ttccggcttg tccggggcat 540

gtttggtgaa ccggtggggt aaggttggtg tgcccaacgc ccgtactttg gtgacacctc 600

aagaccaccc aggaatgcca gggaggtacc ccacctcacg gtgggatctg accctgggct 660

aattgtctac ggtggttctt cttgcttcca cttctttctt ctgttcacg 709

<210> 19

<211> 708

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human Parechovirus sequence"

<400> 19

tttgaaaggg gtctcctaga gagcttggcc gtcgggcctt ataccccgac ttgctgagtt 60

tctctaggag agccctttc ccagccctga ggcggctggt caataaaagc ctcaaacgta 120

actaacacct aagaagatca tgtaaaccct atgcctggtc tccactattc gaaggcaact 180

tgcaataaga agagtgggat caagacgctt aaagcataga gacagttttc ttttctaacc 240

cacatttgtg tggggtggca gatggcgtgc cataactcta atagtgagat accacgcttg 300

tggaccttat gctcacacag ccatcctcta gtaagtttgt gagacgtctg gtgacgtgtg 360

ggaacttatt ggaaacaaca ttttgctgca aagcatccta ctgccagcgg aaaaacacct 420

ggtaacaggt gcctctgggg ccaaaagcca aggtttaaca gacccttag gattggttct 480

aaacctgaga tgttgtggaa gatatttagt acctgctgat ctggtagtta tgcaaacact 540

agttgtaagg cccatgaagg atgcccagaa ggtacccgta ggtaacaagt gacactatgg 600

atctgatttg gggccagata cctctatctt ggtgatctgg ttaaaaaaca tctaatgggc 660

caaacccggg ggggatcccc ggtttcctct tattctatca atgccact 708

<210> 20

<211> 508

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Crohivirus B sequence"

<400> 20

gtataagaga caggtgtttg ccttgtcttc ggactggcat cttgggacca acccccctttt 60

tccccagcca tgggttaaat ggcaataaag gacgtaacaa ctttgtaacc attaagcttt 120

gtaattttgt aaccactaag ctttgtgcac ataatgtaac catcaagctt gttagtccca 180

gcaggaggtt tgcatgcttg tagccgaaat ggggctcgac cccccatagt aggatacttg 240

attttgcatt ccattgtgga cctgcaaact ctacacatag aggctttgtc ttgcatctaa 300

acacctgagt acagtgtgta cctagaccct atagtacggg aggaccgttt gtttcctcaa 360

taaccctaca taataggcta ggtgggcatg cccaatttgc aagatcccag actgggggtc 420

ggtctgggca gggttagatc cctgttagct actgcctgat agggtggtgc tcaaccatgt 480

gtagtttaaa ttgagctgtt catatacc 508

<210> 21

<211> 248

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Picornaviridae sp. sequence"

<400> 21

actgaagatc ctacagtaac tactgcccca atgaacgcca cagatgggtc tgctgatgac 60

tacctatctt agtgctagtt gaggtttgaa gtgagccggt ttttagaaga accagtttct 120

gaacattatc atccccagca tctattctat acgcacaaga tagatagtca tcagcagaca 180

catctgtgct actgcttgat agagttgcgg ctggtcaact tagattggta taaccagttg 240

agtggcaa 248

<210> 22

<211> 516

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Rosavirus M-7 sequence"

<400> 22

tatgcatcac tggacggcct aacctcggtc gtggcttctt gccgatttca gcgctaccag 60

gctttctggt ctcgccaggc gttgattagt aggtgcactg tctaagtgaa gacagcagtg 120

ctctctgtga aaagttgatg acactcttca ggtttgtagc gatcactcaa ggctagcgga 180

tttccccgtg tggtaacaca cgcctctagg cccagaaggc acggtgttga cagcacccct 240

tgagtggctg gtcttcccca ccagcacctg atttgtggat tcttcctagt aacggacaag 300

catggctgct cttaagcatt cagtgcgtcc ggggctgaag gatgcccaga aggtacccgc 360

aggtaacgat aagctcactg tggatctgat ctggggctgc gggctgggtg tctttccacc 420

cagccaaaac ccgtaaaacg gtagtcgcag ttaaaaaacg tctaggcccc acccccccag 480

ggatgggggg ttcccttaaa ccctcacaag ttcaac 516

<210> 23

<211> 478

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Shanbavirus A sequence"

<400> 23

tgaaaagggg gcgcagggtg gtggtggtta ctaaataccc accatcgccc tgcacttccc 60

ttttcccctg tggctcaggg tcacttagcc ccctctttgg gttaccagta gttttctacc 120

cctgggcaca gggttaacta tgcaagacgg aacaacaatc tcttagtccc cctcgccgat 180

agtgggctcg acccccatgt gtaggagtgg ataagggacg gagtgagccg atacggggaa 240

gagtgtgcgg tcacacctta attccatgag cgctgcgaag aaggaagctg tgaacaatgg 300

cgacctgaac cgtacacatg gagctccaca ggcatggtac tcgttagact acgcagcctg 360

gttgggagtg ggtataccct gggtgagccg ccagtgaatg ggagttcact ggttaacaca 420

cactgcctga tagggtcagg gcctcctgtc cccgccgtaa tgaggtagac catatgcc 478

<210> 24

<211> 403

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Pasivirus A sequence"

<400> 24

gcggctggat attctggccg tgcaactgct tttgaccagt ggctctgggt aacttagcca 60

aagtgtcctt ctccctttcc ctattatatg ttttatggct ttgtctggtc ttgtttagtt 120

tatatataag atcctttccg ccgatataga cctcgacagt ctagtgtagg aggattggtg 180

atattaattt gccccagaag agtgaccgtg acacatagaa accatgagta catgtgtatc 240

cgtggaggat cgcccgggac tggattccat atcccattgc catcccaaca agcggagggt 300

atacccacta tgtgcacgtc tgcagtggga gtctgcagat ttagtcatac tgcctgatag 360

ggtgtgggcc tgcactctgg ggtactcagg ctgtttatat aat 403

<210> 25

<211> 404

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Pasivirus A sequence"

<400> 25

gctggacttt ctggctgcgc aactgctttt aaccagtggc tctgggttac ttagccaaaa 60

ccccctttcc ccgtacccta gtttgtgtgt gtattattat tttgttgttg ttttgtaaat 120

ttttatataa gatcctttcc gccgatatag acctcgacag tctagtgtag gaggattggt 180

gatattaata tgccccagaa gagtgaccgt gacacataga aaccatgagt acatgtgtat 240

ccgtggagga tcgcccggga ctggattcca tatcccattg ccatcccaac aaacggaggg 300

tatacccgct atgtgcgcgt ctacagtggg aatctgtaga tttagtcata ctgcctgata 360

gggtgtgggc ctgcactctg gggtactcag gctgtttata taat 404

<210> 26

<211> 744

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Echovirus E14 sequence"

<400> 26

ttaaaacagc ctgtgggttg ttcccatcca cagggcccac tgggcgccag cactctggta 60

ttgcggtacc ttagtgcgcc tgttttatat acccgtcccc caaacgtaac ttagacgcat 120

gtcaacgaag accaatagta agcgcagcac accagctgtg ttccggtcaa gcacttctgt 180

taccccggac cgagtatcaa taagctactc acgtggctga aggagaaaac gttcgttacc 240

cgaccaatta cttcaagaaa cctagtaaca ccatgaaggt tgcgcagtgt ttcgctccgc 300

acaaccccag tgtagatcag gtcgatgagt caccgcattc cccacgggtg accgtggcgg 360

tggctgcgct ggcggcctgc ccatggggaa acccatggga cgcttcaata ctgacatggt 420

gcgaagagtc tattgagcta attggtagtc ctccggcccc tgaatgcggc taatcctaac 480

tgcggagcag atacccacac accagtgggc agtctgtcgt aacgggcaac tctgcagcgg 540

aaccgactac tttgggtgtc cgtgtttctc tttatcctta tactggctgc ttatggtgac 600

aattgagaga ttgttaccat atagctattg gattggccat ccggtgacaa atagagcaat 660

tgtgtatttg tttgttggtt tcgtgccatt aaattacaag gttctaaaca cccttaatct 720

tattatagca ttcaacacaa caaa 744

<210> 27

<211> 663

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human Parechovirus sequence"

<400> 27

gtacattaga tgcgtcatct gcaactttag tcaataaatt acctccaatg tcattaccaa 60

cattccctac cttttcacta acacctaaga caacaagtac ctatgcctgg tctccactat 120

tcgaaggcaa cttgcaataa gaagagtgga attaagacgc ttaaagcata gagctagtta 180

tcttttctaa cccacaaagt tttgtggggt ggcagatggc gtgccataac tctattagtg 240

agataccatg cttgtggatc ttatgctcac acagccatcc tctagtaagt tgataaggtg 300

tctggtgata tgtgggaact cacatgaacc attaatttac cgtaaggtat cctatagcca 360

gcggaatcac atctggtgac agatgcctct ggggccgaaa gccaaggttt aacagaccct 420

ataggattgg tttcaaaacc tgaattgatg tggattgtgt atagtacctg ttgatctggt 480

aacagtgtca acactagttg taaggcccac gaaggatgcc cagaaggtac ccgtaggtaa 540

caagtgacac tatggatctg atctggggcc agctacctct atcatggtga gttggttaaa 600

aaacgtctag tgggccaaac ccaggggggga tccctggttt ccttttacct aatcaaagcc 660

act 663

<210> 28

<211> 744

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Aichi virus sequence"

<400> 28

tttgaaaagg gggtgggggg gcctcggccc cctcaccctc ttttccggtg gtctggtccc 60

ggaccaccgt tactccattc agcttcttcg gaacctgttc ggaggaatta aacgggcacc 120

catactcccc ccacccccct tttgtaacta agtatgtgtg ctcgtgatct tgactcccac 180

ggaacggacc gatccgttgg tgaacaaaca gctaggtcca catcctccct tcccctggga 240

gggcccccgc cctcccacat cctccccca gcctgacgta tcacaggctg tgtgaagccc 300

ccgcgaaagc tgctcacgtg gcaattgtgg gtcccccctt catcaagaca ccaggtcttt 360

cctccttaag gctagccccg gcgtgtgaat tcacgttggg caactagtgg tgtcactgtg 420

cgctcccaat ctcggccgcg gagtgctgtt ccccaagcca aacccctggc ccttcactat 480

gtgcctggca agcatatctg agaaggtgtt ccgctgtggc tgccaacctg gtgacaggtg 540

ccccagtgtg cgtaaccttc ttccgtctcc ggacggtagt gattggttaa gatttggtgt 600

aaggttcatg tgccaacgcc ctgtgcggga tgaaacctct actgccctag gaatgccagg 660

caggtacccc acctccgggt gggatctgag cctgggctaa ttgtctacgg gtagtttcat 720

ttccaatcct tttatgtcgg agtc 744

<210> 29

<211> 734

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hepatitis A virus sequence"

<400> 29

ttcaagaggg gtctccggag ttttccggaa cccctcttgg aagtccatgg tgaggggact 60

tgatacctca ccgccgtttg cctaggctat aggctaaatt tccctttccc tgtccttccc 120

ctatttcctt ttgttttgtt tgtaaatatt aattcctgca ggttcagggt tctttaatct 180

gtttctctat aagaacactc aatttttcac gctttctgtc tcctttcttc cagggctctc 240

cccttgccct aggctctggc cgttgcgccc ggcggggtca actccatgat tagcatggag 300

ctgtaggagt ctaaattggg gacgcagatg tttgggacgt cgccttgcag tgttaacttg 360

gctttcatga acctctttga tcttccacaa ggggtaggct acgggtgaaa cctcttaggc 420

taatacttca atgaagagat gccttggata gggtaacagc ggcggatatt ggtgagttgt 480

taagacaaaa accattcaac gccggaggac tggctctcat ccagtggatg cattgaggga 540

attgattgtc agggctgtct ctaggtttaa tctcagacct ctctgtgctt agggcaaaca 600

ctatttggcc ttaaatggga tcctgtgaga gggggtccct ccattgacag ctggactgtt 660

ctttggggcc ttatgtggtg tttgcctctg aggtactcag gggcatttag gtttttcctc 720

attcttaaat aata 734

<210> 30

<211> 638

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Phopivirus sequence"

<400> 30

gggagtaaac ctcaccaccg tttgccgtgg tttacggcta cctatttttg gatgtaaata 60

ttaattcctg caggttcagg tctcttgaat tatgtccacg ctagtggcac tctcttaccc 120

ataagtgacg ccttagcgga acctttctac acttgatgtg gttaggggtt acattatttc 180

cctgggcctt ctttggccct ttttcccctg cactatcatt ctttcttccg ggctctcagc 240

atgccaatgt tccgaccggt gcgcccgccg gggttaactc catggttagc atggagctgt 300

aggccctaaa agtgctgaca ctggaactgg actattgaag catacactgt taactgaaac 360

atgtaactcc aatcgatctt ctacaagggg taggctacgg gtgaaacccc ttaggttaat 420

actcatattg agagatactt ctgataggtt aaggttgctg gataatggtg agtttaacga 480

caaaaaccat tcaacagctg tgggccaacc tcatcaggta gatgcttttg gagccaagtg 540

cgtaggggtg tgtgtggaaa tgcttcagtg gaaggtgccc tcccgaaagg tcgtaggggt 600

aatcaggggc agttaggttt ccacaattac aatttgaa 638

<210> 31

<211> 743

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CVA10 sequence"

<400> 31

gctcttccga tctgggttgt tcccacccac agggcccact gggcgccagc actctgattc 60

cacggaatct ttgtgcgcct gttttacaac ccttcccaat ttgtaacgta gaagcaatac 120

acactactga tcaatagtag gcatggcgcg ccagtcatgt catgatcaag cacttctgtt 180

ccccggact gagtatcaat agactgctca cgcggttgaa ggagaaaacg ttcgttaccc 240

ggctaactac ttcgagaaac ctagtagcac catggaagct gcggagtgtt tcgctcagca 300

ctttccccgt gtagatcagg tcgatgagtc actgcaatcc ccacgggcga ccgtggcagt 360

ggctgcgttg gcggcctgcc tatggggcaa cccataggac gctctaatgt ggacatggtg 420

cgaagagtct attgagctag ttagtagtcc tccggcccct gaatgcggct aatcctaact 480

gcggagcaca tgccttcaac ccaggaggtg gtgtgtcgta acgggtaact ctgcagcgga 540

accgactact ttgggtgtcc gtgtttcctt ttatccttat attggctgct tatggtgaca 600

atcacggaat tgttgccata tagctattgg attggccatc cggtgtctaa cagagctatt 660

gtatacctat ttgttggatt tactccccta tcatacaaat ctctgaacac tttgtgcttt 720

atactgaact taaacacacg aaa 743

<210> 32

<211> 746

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Enterovirus C sequence"

<400> 32

ttaaaacagc tctggggttg ttcccacccc agaggcccac gtggcggcca gtacaccggt 60

accacggtac ccttgtacgc ctgttttata ctcccctccc cgtaaactag aagcacgaaa 120

cacaagttca atagaagggg gtacagacca gtaccaccac gaacaagcac ttctgttccc 180

ccggtgaggt cacatagact gtccccacgg tcaaaagtga ctgatccgtt atccgctcac 240

gtacttcgga aagcctagta ccaccttgga atctacgatg cgttgcgctc agcactcgac 300

cccggagtgt agcttaggct gatgagtctg gacgttcccc actggtgaca gtggtccagg 360

ctgcgttggc ggcctacctg tggtccaaaa ccacaggacg ctagtagtga acaaggtgtg 420

aagagcccac tgagctacct gagaatcctc cggcccctga atgcggctaa tcccaaccac 480

ggagcaggta atcgcaaacc agcggtcagc ctgtcgtaac gcgtaagtct gtggcggaac 540

cgactacttt gggtgtccgt gtttcctttt atttttatgg tggctgctta tggtgacaat 600

catagattgt tatcataaag caaattggat tggccatccg gagtgagcta aactatctat 660

ttctctgagt gttggattcg tttcacccac attctgaaca atcagcctca ttagtgttac 720

cctgttaata agacgatatc atcacg 746

<210> 33

<211> 726

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Enterovirus D sequence"

<400> 33

ttaaaacagc tctggggttg ttcccacccc agaggcccac gtggcggcta gtactccggt 60

accccggtac ccttgtacgc ctgttttata ctccctttcc caagtaactt tagaagaaat 120

aaactaatgt tcaacaggag ggggtacaaa ccagtaccac cacgaacaca cacttctgtt 180

tccccggtga agttgcatag actgtaccca cggttgaaag cgatgaatcc gttacccgct 240

taggtacttc gagaagccta gtatcatctt ggaatcttcg atgcgttgcg atcagcactc 300

taccccgagt gtagcttggg tcgatgagtc tggacacccc acaccggcga cgtggtccag 360

gctgcgttgg cggcctaccc atggctagca ccatgggacg ctagttgtga acaaggtgcg 420

aagagcctat tgagctacct gagagtcctc cggcccctga atgcggctaa tcccaaccac 480

ggagcaaatg ctcacaatcc agtgagtggt ttgtcgtaat gcgcaagtct gtggcggaac 540

cgactacttt gggtgtccgt gtttcctttt atttttatta tggctgctta tggtgacaat 600

ctgagattgt tatcatatag ctattggatt agccatccgg tgatatcttg aaattttgcc 660

ataacttttt cacaaatcct acaacattac actacacttt ctcttgaata attgagacaa 720

ctcata 726

<210> 34

<211> 737

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Enterovirus J sequence"

<400> 34

ttaaaatagc ctcagggttg ttcccaccct gagggcccac gtggtgtagt actctggtat 60

tacggtacct ttgtacgcct attttatacc cccttcccca agtaatttag aagcaagcac 120

aaaccagttc agtagtaagc agtacaatcc agtactgtaa tgaacaagta cttctgttac 180

cccggaaggg tctatcggta agctgtaccc acggctgaag aatgacctac cgttaaccgg 240

ctacctactt cgagaagcct agtaatgccg ttgaagtttt attgacgtta cgctcagcac 300

actaccccgt gtgtagtttt ggctgatgag tcacggcact ccccacgggc gaccgtggcc 360

gtggctgcgt tggcggccaa ccaaggagtg caagctcctt ggacgtcata ttacagacat 420

ggtgtgaaga gcctattgag ctaggtggta gtcctccggc ccctgaatgc ggctaatcct 480

aactccggag catatcggtg cgaaccagca cttggtgtgt tgtaatacgt aagtctggag 540

cggaaccgac tactttgggt gtccgtgttt cctgttttaa cttttatggc tgcttatggt 600

gacaatttaa cattgttacc atatagctgt tgggttggcc atccggattt tgttataaaa 660

ccatttcctc gtgccttgac ctttaacaca tttgtgaact tctttaaatc ccttttatta 720

gtccttaaat actaaga 737

<210> 35

<211> 327

<212> DNA

<213> Pegivirus sp.

<400> 35

aactgttgtt gtagcaatgc gcatattgct acttcggtac gcctaattgg taggcgcccg 60

gccgaccggc cccgcaaggg cctagtagga cgtgtgacaa tgccatgagg gatcatgaca 120

ctggggtgag cggaggcagc accgaagtcg ggtgaactcg actcccagtg cgaccacctg 180

gcttggtcgt tcatggaggg catgcccacg ggaacgctga tcgtgcaaag ggatgggtcc 240

ctgcactggt gccatgcgcg gcaccactcc gtacagcctg atagggtggc ggcgggcccc 300

cccagtgtga cgtccgtgga gcgcaac 327

<210> 36

<211> 281

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GB virus C sequence"

<400> 36

tgacgtgggg gggttgattt tccccccccg gcactgggtg caagccccag aaaccgacgc 60

ctatctaagt agacgcaatg actcggcgcc gactcggcga ccggccaaaa ggtggtggat 120

gggtgatgac agggttggta ggtcgtaaat cccggtcatc ctggtagcca ctataggtgg 180

gtcttaagag aaggtcaaga ttcctcttac gcctgcggcg agaccgcgca cggtccacag 240

gtgttggccc taccggtgtg aataagggcc cgacatcagg c 281

<210> 37

<211> 459

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GB virus C sequence"

<400> 37

gacgtggggg ggttgatccc ccccctttgg cactgggtgc aagccccaga aaccgacgcc 60

tatttaaaca gacgttaaga accggcgccg acccggcgac cggccaaaag gtggtggatg 120

ggtgatgcca gggttggtag gtcgtaaatc ccggtcatct tggtagccac tataggtggg 180

tcttaagggt tggttaaggt ccctctggcg cttgtggcga gaaagcgcac ggtccacagg 240

tgttggccct accggtgtga ataagggccc gacgtcaggc tcgtcgttaa accgagccca 300

ctacccacct gggcaaacaa cgcccacgta cggtccacgt cgcccttcaa tgtctctctt 360

gaccaatagg cttagccggc gagttgacaa ggaccagtgg gggctgggcg gtaggggaag 420

gacccctgcc gctgcccttc ccggtggagt gggaaatgc 459

<210> 38

<211> 350

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GB virus C sequence"

<400> 38

tgacgtgggg gggttgatcc gccccccccg gcactcggtg caagccccat aaaccgacgc 60

ctatctaagt agacgcaatg actcggcgcc gactcggcga ccggccaaaa ggtggtggat 120

gggtggtgac agggttggta ggtcgtaaat cccggtcatc ctggtagcca ctataggtgg 180

gtcttaagag aaggtcaaga ctcctcttgt gcctgcggcg agaccgcgca cggtccacag 240

gtgctggccc taccggtgtg aataagggcc cgacgtcagg ctcgtcgtta aaccgagccc 300

gtcacccacc tgggcaaacg acgcccacgt acggtccacg tcgcccttca 350

<210> 39

<211> 318

<212> DNA

<213> Pegivirus A

<400> 39

tgtagcaatg cgcatattgc tacttcggta cgcctaattg gtaggcgccc ggccgaccgg 60

ccccgcaagg gcctagtagg acgtgtgaca atgccatgcg ggatcatgac actggggtga 120

gcggaggcag caccgaagtc gggtgaactc gactcccagt gcgaccacct ggcttggtcg 180

ttcatggagg gcatgcccac gggaacgctg atcgtgcaaa gggatgggtc cctgcactgg 240

tgccatgcgc ggcaccactc cgtacagcct gatagggtgg cggcgggccc ccccagtgtg 300

acgtccgtgg agcgcaac 318

<210> 40

<211> 399

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Pasivirus A sequence"

<400> 40

attttctggc cgtgtagctg cttttgacca gtggctctgg gttacttagc caaatccccc 60

ttccttcacc cttttaaatt tgatggtctg tgttgtttgt tttgtcttgt ctaaataata 120

tataagatcc ttcccgccga tacagacctc gacagtctgg tgtaggaggg ttggtgttat 180

taatttgccc cagaagagtg accgtgacac atagaaacca tgagtacatg tgtatccgtg 240

gaggatcgcc cgggactgga ttccatatcc cattgccatc ccaacaagcg gagggtatac 300

ccactatgtg cgcgtttgca gtgggaatct gcaaatttag tcatactgcc tgatagggtg 360

tgggcctgca ctctggggta ctcaggctgt tcatataat 399

<210> 41

<211> 430

<212> DNA

<213> Sapelovirus sp.

<400> 41

cccctccacc cttaaggtgg ttgtatccca catccccac cctcccttcc aaagtggacg 60

gacaactgga ttttgactaa cggcaagtct gaatggtatg atttggatac gtttaaacgg 120

cagtagcgtg gcgagctatg gaaaaatcgc aattgtcgat agccatgtta gtgacgcgct 180

tcggcgtgct cctttggtga ttcggcgact ggttacagga gagtaggcag tgagctatgg 240

gcaaacctct acagtattac ttagagggaa tgtgcaattg agacttgacg agcgtctctt 300

tgagatgtgg cgcatgctct tggcattacc atagtgagct tccaggttgg gaaacctgga 360

ctgggcctat actacctgat agggtcgcgg ctggccgcct gtaactagta tagtcagttg 420

aaaccccccc 430

<210> 42

<211> 504

<212> DNA

<213> Rosavirus B

<400> 42

gtctctttag tgtctatgct tcagagagcg gtgaactgac accgttgctt cttgcacagc 60

ccttcgtgcc ggtctttccg gttctcgaca gcgttgggca tcatggctag ttaggctaag 120

atagtggatg atctagtgaa cagttttgga ttgtttggag ttttgtagcg atgctagtag 180

tgtgtgtgga cctccccacg tggtaacacg tgccccacag gccaaaagcc aaggtgttga 240

aagcacccct actagtccca gactcaccca tctgggaact cctctcatga aaaatcttag 300

taacttttga ttcggctatt catcaacctc tctagtcaag ggctgaagga tgcccggaag 360

gtacccgcag gtaacgataa gctcactgtg gatctgatcc ggggctttgg tgcgaccgtc 420

tgtccggcgt agccagagtt aaaaaacgtc taggcccttc caccccaagg gattggggtt 480

tccccaatca tttgaaagtt cact 504

<210> 43

<211> 532

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bakunsa virus sequence"

<400> 43

ttttgaacgc cacctcggag cgatatccgg ggacccccctc ccctttttcc ttcctacctt 60

cttcccaaat ttccctcttc ccttgttatt ttggtttgga tttcctggac atgactcgga 120

cggatctatc tcatttgctt tgtgtctgct ccaccagtgg catggtcgaa agatcatcaa 180

cactggacgt gtactgtaat ggccaaacgt gcccacaggg gaaaccatgc cggtcgctgt 240

agcggcgggt ggacgtggtg gacccctctc cctgctcata aactttgggt aggtgaaggg 300

ttcaagcgac gcttgccgtg agggcgcatc cggatggtgg gaaccaacaa actaggctgt 360

aatggccgac ctcaggtgga tgagctaggg ctgctgcacc aaaagggact cgattcgata 420

tcccggcctg gtagcctagt gcagtggact cgtagttggg aatctacgac tggcctagta 480

cagggtgata gccccgtttc ccacgcccac ctgttgtagg gacacccccc cc 532

<210> 44

<211> 494

<212> DNA

<213> Tremovirus A

<400> 44

tttgaaagag gcctccggag tgtccggagg ctctctttcg acccaaccca tactgggggg 60

tgtgtgggac cgtacctgga gtgcacggta tatatgcatt cccgcatggc aagggcgtgc 120

taccttgccc cttgacgcat ggtatgcgtc atcatttgcc ttggttaagc cccatagaaa 180

cgaggcgtca cgtgccgaaa atccctttgc gtttcacaga accatcctaa ccatgggtgt 240

agtatgggaa tcgtgtatgg ggatgattag gatctctcgt agagggatag gtgtgccatt 300

caaatccagg gagtactctg gctctgacat tgggacattt gatgtaaccg gacctggttc 360

agtatccggg ttgtcctgta ttgttacggt gtatccgtct tggcacactg aaagggtatt 420

tttgggtaat cctttcctac tgcctgatag ggtggcgtgc ccggccacga gagattaagg 480

gtagcaattt aaac 494

<210> 45

<211> 378

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Swine Pasivirus 1 sequence"

<400> 45

gcttttgacc agtggctctg ggttacttag ccaagtccct ttctcttatt ttcactagtt 60

tatgttgtgt gttgtctgtt ttgttttgtt taaattgtat acaagatcct tcccgccgac 120

acagacctcg acagtctggt gtaggagggt tggtgatatt aatttgcccc aaaagagtga 180

ccgtgatacg tggaaaccat gagtacatgt gtatccgtgg aggatcgccc gggactggat 240

tccatatccc attgccatcc caacaaacgg agggtatacc caccacgtgc gcgtttgcag 300

tgggaatctg caaatttagt catactgcct gatagggtgt gggcctgcac tttggggtac 360

tcaggctgtt catataat 378

<210> 46

<211> 319

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Parechovirus-like virus PLV-CHN sequence"

<400> 46

acatggggta tgttgtctgt cctgttttgt tgaaacaata tataagatcc tttccgccga 60

tatagacctc gacagtctag tgtaggagga ttggtgatag taacttgccc cagaagagtg 120

accgtgacac atagaaacca tgagtacatg tgtatccgtg gaggatcgcc cgggactgga 180

ttccatatcc cattgccatc ccaacaaacg gagggtatac ccactatgtg cgcgtttgca 240

gtgggagcct gcaaatttag tcatactgcc tgatagggtg tgggcctgca ctctggggta 300

ctcaggctgt ttatataat 319

<210> 47

<211> 420

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Pasivirus A sequence"

<400> 47

tgaaaaagtg gttgtgcagc tggattttcc ggctgtgcaa ctgcttttga ccagtggctc 60

tgggttactt agccaaattc ctttccctta tccctattgg tttgtgttgt gtgttgtttg 120

ttttgttttg tcttaactat atacaagatc cttcccgccg atacagacct cgacagtctg 180

gtgtaggagg gttggtgtta ttaatttgcc ccaaaagagt gaccgtgaca cgtggaaacc 240

atgagtacat gtgtatccgt ggaggatcgc ccgggactgg attccatatc ccattgccat 300

cccaacaaac ggagggtata cccaccacgt gcgcgtttgc agtgggaatc tgcaaattta 360

gtcatactgc ctgatagggt gtgggcctgc actttggggt actcaggctg tttatataat 420

<210> 48

<211> 373

<212> DNA

<213> Sicinivirus sp.

<400> 48

gtgtcattaa ggtgtgtttg gaagttcgaa ttagctggtt tgtggtgatt agtagacccc 60

ctggaggtac ccaattcgga tctgaccagg gacccgtgac tataccgctc cggtaattcg 120

ggtttaaaac aatgaacgtc accacacaat tacttttctc attttatttt catcattgtc 180

ttcctattta ccgattacac tcgatttcct tggatgttcc tggagatttc cctggttacc 240

tggaccctca ttattgttgt tgtttcaccc agcgagctgt cccaattgct tattatttgc 300

gcttacaact tcgtcctaat atttttctgg ttgatcgggt tgattgagct cccgggctat 360

cctgccattc aac 373

<210> 49

<211> 280

<212> DNA

<213> Hepacivirus K

<400> 49

gggaacaatg gtccgtccgc ggaacgactc tagccatgag tctagtacga gtgcgtgcca 60

cccattagca caaaaaccac tgactgagcc acacccctcc cggaatcctg agtacaggac 120

attcgctcgg acgacgcatg agcctccatg ccgagaaaat tgggtatacc cacgggtaag 180

gggtggccac ccagcgggaa tctgggggct ggtcactgac tatggtacag cctgataggg 240

tgctgccgca gcgtcagtgg tatgcggctg ttcatggaac 280

<210> 50

<211> 384

<212> DNA

<213> Hepacivirus A

<400> 50

acctccgtgc taggcacggt gcgttgtcag cgttttgcgc ttgcatgcgc tacacgcgtc 60

gtccaacgcg gagggaactt cacatcacca tgtgtcactc cccctatgga gggttccacc 120

ccgcttacac ggaaatgggt taaccatacc caaagtacgg gtatgcgggt cctcctaggg 180

ccccccggc aggtcgaggg agctggaatt cgtgaattcg tgagtacacg aaaatcgcgg 240

cttgaacgtc tttgaccttc ggagccgaaa tttgggcgtg ccccacgaag gaaggcgggg 300

gcggtgttgg gccgccgccc cctttatccc acggtctgat aggatgcttg cgagggcacc 360

tgccggtctc gtagaccata ggac 384

<210> 51

<211> 384

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bovine viral diarrhea virus 1 sequence"

<400> 51

gtatacgaga atttgcctag gacctcgttt acaatatggg caatctaaaa ttataattag 60

gcctaaggga caaatcctcc tcagcgaagg ccgaaaagag gctagccatg cccttagtag 120

gactagcaaa ataaggggggg tagcaacagt ggtgagttcg ttggatggct gaagccctga 180

gtacagggta gtcgtcagtg gttcgacgct tcggaggaca agcctcgaga taccacgtgg 240

acgagggcat gcccacagca catcttaacc tggacggggg tcgttcaggt gaaaacggtt 300

taaccaaccg ctacgaatac agcctgatag ggtgctgcag aggcccactg tattgctact 360

gaaaatctct gctgtacatg gcac 384

<210> 52

<211> 372

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Border Disease virus sequence"

<400> 52

gtatacggga gtagctcatg cccgtataca aaattggata ttccaaaact cgattgggtt 60

agggagccct cctagcgacg gccgaaccgt gttaaccata cacgtagtag gactagcaga 120

cgggaggact agccatcgtg gtgagatccc tgagcagtct aaatcctgag tacaggatag 180

tcgtcagtag ttcaacgcag gcacggttct gccttgagat gctacgtgga cgagggcatg 240

cccaagactt gctttaatct cggcgggggt cgccgaggtg aaaacaccta acggtgttgg 300

ggttacagcc tgatagggtg ctgcagaggc ccacgaatag gctagtataa aaatctctgc 360

tgtacatggc ac 372

<210> 53

<211> 388

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bovine viral diarrhea virus 2 sequence"

<400> 53

gtatacgaga ttagctaaag tactcgtata tggattggac gtcaacaaat ttttaattgg 60

caacgtaggg aaccttcccc tcagcgaagg ccgaaaagag gctagccatg ccctttagta 120

ggactagcaa aagtaggggg actagcggta gcagtgagtt cgttggatgg ccgaacccct 180

gagtacaggg gagtcgtcaa tggttcgaca ctccattagt cgaggagtct cgagatgcca 240

tgtggacgag ggcatgccca cggcacatct taacccatgc gggggttgca tgggtgaaag 300

cgctaatcgt ggcgttatgg acacagcctg atagggtgta gcagagacct gctattccgc 360

tagtaaaaaa ctctgctgta catggcac 388

<210> 54

<211> 373

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Classical swine fever virus sequence"

<400> 54

gtatacgagg ttagttcatt ctcgtatgca ttattggaca aatcaaaatt tcaatttggt 60

tcagggcctc cctccagcga cggccgaact gggctagcca tgcccatagt aggactagca 120

aacggaggga ctagccgtag tggcgagctc cctgggtgtt ctaagtcctg agtacaggac 180

agtcgtcagt agttcgacgt gagcagaagc ccacctcgag atgctatgtg gacgagggca 240

tgcccaagac gcaccttaac cctagcgggg gtcgctaggg tgaaatcaca ccacgtgatg 300

ggagtccgac ctgatagggt gctgcagagg ctcactatta ggctagtata aaaatctctg 360

ctgtacatgg cac 373

<210> 55

<211> 191

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human blood-associated dicistrovirus sequence"

<400> 55

aaaaccgacc ccagagatca gaaagtcgtt gacgcgatct tttattagag gacgttgcgc 60

tggcgcgagc tttaattagc agacgccaaa aataaacaac aaaatgctga tcgcgagact 120

taattgtcag acgattggcc aaatccgatg tgatctttgc tgctcccaga ttgccgaaat 180

aggagtagta g 191

<210> 56

<211> 438

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hubei Picorna-like virus sequence"

<400> 56

ccccaaaacc ccccccttaa actcaacact gtagtggatt cattttccgt tgcaaaacaa 60

aacattacta cccgcattta tgtaggctct gtgttttcta tgcgaccgtt acattaatct 120

ctactctgac ccactagttt ataaaaccga agacctgaat gaaacgattt tccttctttt 180

caacctctaa cgaacctctg acggcttgag aaacctgaag ttagtaatta tgtttaaaag 240

aaaggaaagt caaacgcgat gactcttaca tccctattcc ataccgttgc tccacaatgt 300

gagcgatgcg aggtcgggac tgcagtatta ggggaacgag ctacatggag agttaattat 360

ctctcccctc ctacgggagt ctcatgtgag ctgtagaaag cggttggcac ctctcgttac 420

ctcgcctgta catgatcc 438

<210> 57

<211> 193

<212> DNA

<213> Cricket paralysis virus

<400> 57

aaaagcaaaa atgtgatctt gcttgtaaat acaattttga gaggttaata aattacaagt 60

agtgctattt ttgtatttag gttagctatt tagctttacg ttccaggatg cctagtggca 120

gccccacaat atccaggaag ccctctctgc ggtttttcag attaggtagt cgaaaaacct 180

aagaaattta cct 193

<210> 58

<211> 752

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus A sequence"

<400> 58

tttcctcctt tcgaccgcct tacggcaggc gggtccgcgg acggcttcgg cctacccgcg 60

acaagaatgc cgtcatctgt ccttatcacc catattcttt cccttccccc gcaaccatca 120

cgcttactcg cgcacgtgtt gagtggcacg tgcgttgtcc aaacagttac actcacaccc 180

ttggggcggg tttgtcccgc cctcgggttc ctcgcggaac cctccctctt ctctctccct 240

ttctatccgc cttcactttc cataactaca gtgctttggt aggtaagcat cctgaccccc 300

cgcggaagct gccaacgtgg caactgtggg gatccaggca ggttatcaaa ggcacccggt 360

ctttccgcct tcaggagtat ccctgccggt gaattccgac agggctctgc ttggtgccaa 420

cctcccccaa atgcgcgctg cgggagtgct cttccccaac tcatcttagt aacctctcat 480

gtgtgtgctt ggtcagcata tctgaggcga cgttccgctg tcccagacca gtccagcaat 540

ggacgggcca gtgtgcgtag tcgctttccg gtttcccggc gcatgtttgg cgaaacgctg 600

aggtaaggtt ggtgtgccca atgcccgtaa tttggtgaca cctcaagacc acccaggaat 660

gccagggagg taccccactt cggtgggatc tgaccctggg ctaattgtct acggtggttc 720

ttcttgcttc cacttctctt ttttctggca tg 752

<210> 59

<211> 708

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus A sequence"

<400> 59

tatggcaggc gggcttgtgg acggcttcgg cccacccaca gcaagaatgc catcatctgt 60

cctcaccccc atgtttcccc tttctttccc tgcaaccgtt acgcttactc gcaggtgcat 120

ttgagtggtg cacgtgttga ataaacagct acactcacat gggggcgggt tttcccgccc 180

tgcggcctct cgcgaggccc acccctcccc ttcctcccat aactacagtg ctttggtagg 240

taagcatcct gatcccccgc ggaagctgct cacgtggcaa ctgtggggac ccagacaggt 300

tatcaaaggc acccggtctt tccgccttca ggagtatccc tgctagtgaa ttctagtagg 360

gctctgcttg gtgccaacct cccccaaatg cgcgctgcgg gagtgctctt ccccaactca 420

ccctagtatc ctctcatgtg tgtgcttggt cagcatatct gagacgatgt tccgctgtcc 480

cagaccagtc cagtaatgga cgggccagtg tgcgtagtcg tcttccggct tttccggcgc 540

atgtttggtg aaccggtggg gtaaggttgg tgtgcccaac gcccgtactt tggtgatacc 600

tcaagaccac ccaggaatgc cagggaggta ccccgcttca cagcgggatc tgaccctggg 660

ctaattgtct acggtggttc ttcttgcttc cacttctttc tactgttc 708

<210> 60

<211> 718

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus A sequence"

<400> 60

tttcgaccgc cttatggcag gcgggcttgt ggacggcttc ggcccaccca cagcaagaat 60

gccatcatct gtcctcaccc ccatttctcc cctccttccc ctgcaaccat tacgcttact 120

cgcatgtgca ttgagtggtg cacgtgttga acaaacagct acactcacgt gggggcgggt 180

tttcccgccc ttcggcctct cgcgaggccc acccttcccc ttcctcccat aactacagtg 240

ctttggtagg taagcatcct gatcccccgc ggaagctgct cgcgtggcaa ctgtggggac 300

ccagacaggt tatcaaaggc acccggtctt tccgcctcca ggagtatccc tgctagtgaa 360

ttctagtggg gctctgcttg gtgccaacct cccccaaatg cgcgctgcgg gagtgctctt 420

ccccaactca ccctagtatc ctctcatgtg tgtgcttggt cagcatatct gagacgatgt 480

tccgctgtcc cagaccagtc cagcaatgga cgggccagtg tgcgtagtcg tcttccggct 540

tgtccggcgc atgtttggtg aaccggtggg gtaaggttgg tgtgcccaac gcccgtactt 600

tggtgacaac tcaagaccac ccaggaatgc cagggaggta ccccgcctca cggcgggatc 660

tgaccctggg ctaattgtct acggtggttc ttcttgcttc catttctttc ttctgttc 718

<210> 61

<211> 699

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus A sequence"

<400> 61

tatggcaggc gggcttgtgg acggtttcgg cccacccaca gcaagaatgc catcatctgt 60

cctcaccccc aattttccct ttcttcccct gcaatcatca cgcttactcg catgtgcatt 120

gagtggtgca tgtgttgaac aaacagctac actcacatgg gggcgggttt tcccgcccta 180

cggcctctcg cgaggcccac ccttcccctc cccttataac tacagtgctt tggcaggtaa 240

gcatcctgat cccccgcgga agctgctcac gtggcaactg tggggaccca gacaggttat 300

caaaggcacc cggtctttcc gccttcagga gcatccccac tagtgaattc tagtggggct 360

ctgcttggtg ccaacctccc ccaaatgcgc gctgcgggag tgctcttccc caacccatcc 420

tagtatcctc tcatgtgtgt gcttggtcag catatctgag acgacgttcc gctgtcccag 480

accagtccag taatggacgg gccagtgtgc gtagtcgtct tccggcttgt ccggcgcatg 540

tttggtgaac cggtggggta aggttggtgt gcccaacgcc cgtactttgg tgacacctca 600

agaccaccca ggaatgccag ggaggtaccc cgcctcacgg cgggatctga ccctgggcta 660

attgtctacg gtggttcttc ttgcttccac ttctttctt 699

<210> 62

<211> 624

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus A sequence"

<400> 62

ttctcctgca accattacgc ttaatcgcat gtgcattgag tggtgcatgt gttgaacaaa 60

cagctacaat cacatggggg cgggttttcc cgccccacgg cttctcgcga ggcccatccc 120

tccctttct cccataacta cagtgctttg gtaggtaagc atcccgatct cccgcggaag 180

ctgctcacgt ggcaactgtg gggacccaga caggttatca aaggcacccg gtctttccgc 240

cttcaggagt atccctgcta gcgaattcta gtagggctct gcttggtgcc aacctctccc 300

aaatgcgcgc tgcgggagtg ctcttcccca aatcacccca gtatcctctc atgtgtgtgc 360

ctggtcagca tatctgagac gatgttccgc tgtcccagac cagtccagta atggacgggc 420

cagtgtgcgt agtcgtcctc cggcttgtcc ggcgcatgtt tggtgaaccg gtggggtaag 480

gttggtgtgc ccaacgcccg taatcagggg atacctcaag gcacccagga atgccaggga 540

ggtatcccgc ctcacagcgg gatctgaccc tggggtaaat gtctgcgggg ggtcctcttg 600

gcccaattct cagtaatttt cagg 624

<210> 63

<211> 650

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus A sequence"

<400> 63

tctgtcctca ccccatcttc ccttctttcc tgcaccgtta cgcttactcg catgtgcatt 60

gagtggtgca cgtgcttgaa caaacagcta cactcacatg ggggcgggtt ttcccgccct 120

gcggcctctc gcgaggccca cccctcccct tcctcccata actacagtgc tttggtaggt 180

aagcatcctg atcccccgcg gaagctgctc acgtggcaac tgtggggacc cagacaggtt 240

atcaaaggca cccggtcttt ccgccttcag gagtatccct gctagtgaat tctagtaggg 300

ctctgcttgg tgccaacctc ccccaaatgc gcgctgcggg agtgctcttc cccaactcac 360

cctagtatcc tctcatgtgt gtgcttggtc agcatatctg agacgatgtt ccgctgtccc 420

agaccagtcc agtaatggac gggccagtgt gcgtagtcgt cttccggctt gtccggcgca 480

tgtttggtga accggtgggg taaggttggt gtgcccaacg cccgtacttt ggtgatacct 540

caagaccacc caggaatgcc agggaggtac cccgcttcac agcgggatct gaccctgggc 600

taattgtcta cggtggttct tcttgcttcc acttctttct actgttcatg 650

<210> 64

<211> 730

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Salivirus FHB sequence"

<400> 64

acatgggggg tctgcggacg gcttcggccc acccgcgaca agaatgccgt catctgtcct 60

cattacccgt attccttccc ttcccccgca accaccacgc ttactcgcgc acgtgttgag 120

tggcacgtgc gttgtccaaa cagctacacc cacacccttc ggggcgggtt tgtcccgccc 180

tcgggttcct cgcggaaccc ccccctccct ctctctcttt ctatccgccc tcacttccca 240

taactacagt gctttggtag gtgagcaccc tgacccccg cggaagctgc taacgtggca 300

actgtgggga tccaggcagg ttatcaaagg cacccggtct ttccgccttc aggagtatct 360

ctgccggtga attccggtag ggctctgctt ggtgccaacc tcccccaaat gcgcgctgcg 420

ggagtgctct tccccaactc atcttagtaa cctctcatgt gtgtgcttgg tcagcatatc 480

tgaggcgacg ttccgctgtc ccagaccagt ccagcaatgg acgggccagt gtgcgtagtc 540

gctttccggt tttccggcgc atgtttggcg aaacgctgag gtaaggttgg tgtgcccaac 600

gcccgtaatt tggtgatacc tcaagaccac ccaggaatgc cagggaggta ccccacttcg 660

gtgggatctg accctgggct aattgtctac ggtggttctt cttgcttcca cttctctttt 720

ttctggcatg 730

<210> 65

<211> 741

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CVB3 sequence"

<400> 65

ttaaaacagc ctgtgggttg atcccaccca caggcccatt gggcgctagc actctggtat 60

cacggtacct ttgtgcgcct gttttatacc ccctccccca actgtaactt agaagtaaca 120

cacaccgatc aacagtcagc gtggcacacc agccacgttt tgatcaagca cttctgttac 180

cccggactga gtatcaatag actgctcacg cggttgaagg agaaagcgtt cgttatccgg 240

ccaactactt cgaaaaacct agtaacaccg tggaagttgc agagtgtttc gctcagcact 300

accccagtgt agatcaggtc gatgagtcac cgcattcccc acgggcgacc gtggcggtgg 360

ctgcgttggc ggcctgccca tggggaaacc catgggacgc tctaatacag acatggtgcg 420

aagagtctat tgagctagtt ggtagtcctc cggcccctga atgcggctaa tcctaactgc 480

ggagcacaca ccctcaagcc agagggcagt gtgtcgtaac gggcaactct gcagcggaac 540

cgactacttt gggtgtccgt gtttcatttt attcctatac tggctgctta tggtgacaat 600

tgagagatcg ttaccatata gctattggat tggccatccg gtgactaata gagctattat 660

atatcccttt gttgggttta taccacttag cttgaaagag gttaaaacat tacaattcat 720

tgttaagttg aatacagcaa a 741

<210> 66

<211> 741

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CVB1 sequence"

<400> 66

ttaaaacagc ctgtgggttg ttcccaccca caggcccatt gggcgctagc actctggtat 60

cacggtacct ttgtgcgcct gttttacatc ccctccccaa attgtaattt agaagtttca 120

cacaccgatc attagcaagc gtggcacacc agccatgttt tgatcaagca cttctgttac 180

cccggactga gtatcaatag accgctaacg cggttgaagg agaaaacgtt cgttacccgg 240

ccaactactt cgaaaaacct agtaacacca tggaagttgc ggagtgtttc gctcagcact 300

accccagtgt agatcaggtc gatgagtcac cgcgttcccc acgggcgacc gtggcggtgg 360

ctgcgttggc ggcctgccta cggggaaacc cgtaggacgc tctaatacag acatggtgcg 420

aagagtctat tgagctagtt ggtaatcctc cggcccctga atgcggctaa tcctaactgc 480

ggagcacata ccctcaaacc agggggcagt gtgtcgtaac gggcaactct gcagcggaac 540

cgactacttt gggtgtccgt gtttcatttt attcctatac tggctgctta tggtgacaat 600

tgacaggttg ttaccatata gttattggat tggccatccg gtgactaaca gagcaattat 660

atatctcttt gttgggttta taccacttag cttgaaagag gttaaaacac tacatctcat 720

cattaaacta aatacaacaa a 741

<210> 67

<211> 742

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Echovirus 7 sequence"

<400> 67

ttaaaacagc ctgtgggttg ttcccaccca cagggcccat tgggcgtcag caccctggta 60

tcacggtacc tttgtgcgcc tgttttatat cccttccccc aattgtaact tagaagaaac 120

acacaccgat caacagcaag cgtggcacac cagccatgtt ttggtcaagc acttctgtta 180

ccccggactg agtatcaata gactgctcac gcggttgaag gagaaagcgt ccgttatccg 240

gccagctact tcgagaaacc tagtaacacc atggaagttg cggagtgttt cgctcagcac 300

taccccagtg tagatcaggt cgatgagtca ccgctttccc cacgggcgac cgtggcggtg 360

gctgcgttgg cggcctgcct atgggggaac ccataggacg ctctaataca gacatggtgc 420

gaagagtcta ttgagctagc tggtattcct ccggcccctg aatgcggcta atcctaactg 480

tggagcacat gcccctaatc caaggggtag tgtgtcgtaa tgagcaattc cgcagcggaa 540

ccgactactt tgggtgtccg tgtttcctct tattcttgta ctggctgctt atggtgacaa 600

ttgagagatt gttaccatat agctattgga ttggccatcc ggtgactaat agagctattg 660

tgtatctctt tgttggattt gtaccactta atttgaaaga aatcaggaca ctacgctaca 720

ttttactatt gaacaccgca aa 742

<210> 68

<211> 743

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CVB5 sequence"

<400> 68

ttaaaacagc ctgtgggttg tacccaccca cagggcccac tgggcgctag cactctggta 60

tcacggtacc tttgtgcgcc tgttttatgc ccccttcccc caattgaaac ttagaagtta 120

cacacaccga tcaacagcgg gcgtggcata ccagccgcgt cttgatcaag cactcctgtt 180

tccccggacc gagtatcaat agactgctca cgcggttgaa ggagaaaacg ttcgttaccc 240

ggctaactac ttcgagaaac ctagtagcat catgaaagtt gcgaagcgtt tcgctcagca 300

catccccagt gtagatcagg tcgatgagtc accgcattcc ccacgggcga ccgtggcggt 360

ggctgcgttg gcggcctgcc tacggggcaa cccgtaggac gcttcaatac agacatggtg 420

cgaagagtcg attgagctag ttagtagtcc tccggcccct gaatccggct aatcctaact 480

gcggagcaca taccctcaac ccaggggcca ttgtgtcgta acgggtaact ctgcagcgga 540

accgactact ttgggtgtcc gtgtttcctt ttattcttat aatggctgct tatggtgaca 600

attgaaagat tgttaccata tagctattgg attggccatc cggtgtctaa cagagctatt 660

atatacctct ttgttggatt tgtaccactt gatctaaagg aagtcaagac actacaattc 720

atcatacaat tgaacacagc aaa 743

<210> 69

<211> 743

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Enterovirus A71 sequence"

<400> 69

ttaaaacagc ctgtgggttg cacccactca cagggcccac tgggcgcaag cactctggca 60

cttcggtacc tttgtgcgcc tgttttatat cccctccccc aatgaaattt agaagcagca 120

aaccccgatc aatagcaggc ataacgctcc agttatgtct tgatcaagca cttctgtttc 180

cccggactga gtatcaatag actgctcacg cggttgaagg agaaaacgtt cgttatccgg 240

ctaactactt cggaaagcct agtaacacca tggaagttgc ggagagtttc gttcagcact 300

tccccagtgt agatcaggtc gatgagtcac cgcattcccc acgggcgacc gtggcggtgg 360

ctgcgttggc ggcctgccca tggggtaacc catgggacgc tctaatacgg acatggtgtg 420

aagagtctac tgagctagtt agtagtcctc cggcccctga atgcggctaa tcccaactgc 480

ggagcacacg cccacaagcc agtgggtagt gtgtcgtaac gggcaactct gcagcggaac 540

cgactacttt gggtgtccgt gtttcctttt attcttatgt tggctgctta tggtgacaat 600

taaagagttg ttaccatata gctattggat tggccatccg gtgtgcaaca gagcgatcgt 660

ttacctattt attggttttg taccattgac actgaagtct gtgatcaccc ttaattttat 720

cttaaccctc aacacagcca aac 743

<210> 70

<211> 742

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CVA3 sequence"

<400> 70

ttaaaacagc ctgtgggttg tacccaccca cagggcccac tgggcgctag cacactggta 60

ttacggtacc tttgtgcgcc tgttttatac ccccccccaac ctcgaaactt agaagtaaag 120

caaacccgat caatagcagg tgcggcgcac cagtcgcatc ttgatcaagc acttctgtaa 180

ccccggaccg agtatcaata gactgctcac gcggttgaag gagaaaacgt tcgttacccg 240

gctaactact tcgagaaacc cagtagcatc atgaaagttg cagagtgttt cgctcagcac 300

taccccccgtg tagatcaggc cgatgagtca ccgcacttcc ccacgggcga ccgtggcggt 360

ggctgcgttg gcggcctgcc tatggggcaa cccataggac gctctaatac ggacatggtg 420

cgaagagtct attgagctag ttagtagtcc tccggcccct gaatgcggct aatcctaact 480

gcggagcaca tacccttaat ccaaagggca gtgtgtcgta acgggtaact ctgcagcgga 540

accgactact ttgggtgtcc gtgtttcctt ttaattttta ctggctgctt atggtgacaa 600

ttgaggaatt gttgccatat agctattgga ttggccatcc ggtgactaac agagctattg 660

tgttccaatt tgttggattt accccgctca cactcacagt cgtaagaacc cttcattacg 720

tgttatttct caactcaaga aa 742

<210> 71

<211> 745

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CVA12 sequence"

<400> 71

ttaaaacagc ctgtgggttg tacccaccca cagggcccac tgggcgctag cactctggta 60

ctacggtacc tttgtgtgcc tgttttaagc ccctaccccc cactcgtaac ttagaaggct 120

tctcacactc gatcaatagt aggtgtggca cgccagtcac accgtgatca agcacttctg 180

ttaccccggt ctgagtacca ataagctgct aacgcggctg aaggggaaaa cgatcgttat 240

ccggctaact acttcgagaa acccagtacc accatgaacg ttgcagggtg tttcgctcgg 300

cacaacccca gtgtagatca ggtcgatgag tcaccgtatt ccccacgggc gaccgtggcg 360

gtggctgcgt tggcggcctg cccatggggt gacccatggg acgctctaat actgacatgg 420

tgcgaagagt ctattgagct agttagtagt cctccggccc ctgaatgcgg ctaatcctaa 480

ctgcggagca catacccctta atccaaaggg cagtgtgtcg taacgggcaa ctctgcagcg 540

gaaccgacta ctttgggtgt ccgtgtttcc ttttattctt acattggctg cttatggtga 600

caattgaaaa gttgttacca tatagctatt ggattggcca tccggtgaca aatagagcta 660

ttgtatatct ttttgttggt tacgtacccc ttaattacaa agtggtttca actttgaaat 720

acatcctaac actaaattgt agaaa 745

<210> 72

<211> 742

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Echovirus E24 sequence"

<400> 72

ttaaaacagc ctgtgggttg cacccaccca cagggcccac agggcgctag cactctggta 60

tcacggtacc tttgtgcgcc tgttttatta ccccttcccc aattgaaaat tagaagcaat 120

gcacaccgat caacagcagg cgtggcgcac cagtcacgtc tcgatcaagc acttctgttt 180

ccccggaccg agtatcaata gactgctcac gcggttgaag gagaaagtgt tcgttatccg 240

gctaaccact tcgagaaacc cagtaacacc atgaaagttg cagggtgttt cgctcagcac 300

ttccccagtg tagatcaggt cgatgagtca ccgcgttccc cacgggcgac cgtggcggtg 360

gctgcgttgg cggcctgcct atgggttaac ccataggacg ctctaataca gacatggtgc 420

gaagagttta ttgagctggt tagtatccct ccggcccctg aatgcggcta atcctaactg 480

cggagcacgt gcctccaatc caggggggttg catgtcgtaa cgggtaactc tgcagcggaa 540

ccgactactt tgggtgtccg tgtttccttt tattcttata ctggctgctt atggtgacaa 600

tcgaggaatt gttaccatat agctattgga ttggccatcc ggtgtctaac agagcgatta 660

tatacctctt tgttggattt atgcagctca ataccaccaa ctttaacaca ttgaaatata 720

tcttaaagtt aaacacagca aa 742

<210> 73

<211> 240

<212> DNA

<213> Anabaena sp.

<400> 73

gaagaaattc tttaagtgga tgctctcaaa ctcagggaaa cctaaatcta gttatagaca 60

aggcaatcct gagccaagcc gaagtagtaa ttagtaagtt aacaatagat gacttacaac 120

taatcggaag gtgcagagac tcgacgggag ctaccctaac gtcaagacga gggtaaagag 180

agagtccaat tctcaaagcc aataggcagt agcgaaagct gcaagagaat gaaaatccgt 240

<210> 74

<211> 239

<212> DNA

<213> Anabaena sp.

<400> 74

aagaaattct ttaagtggat gctctcaaac tcagggaaac ctaaatctag ttatagacaa 60

ggcaatcctg agccaagccg aagtagtaat tagtaagtta acaatagatg acttacaact 120

aatcggaagg tgcagagact cgacgggagc taccctaacg tcaagacgag ggtaaagaga 180

gagtccaatt ctcaaagcca ataggcagta gcgaaagctg caagagaatg aaaatccgt 239

<210> 75

<211> 238

<212> DNA

<213> Anabaena sp.

<400> 75

agaaattctt taagtggatg ctctcaaact cagggaaacc taaatctagt tatagacaag 60

gcaatcctga gccaagccga agtagtaatt agtaagttaa caatagatga cttacaacta 120

atcggaaggt gcagagactc gacgggagct accctaacgt caagacgagg gtaaagagag 180

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgt 238

<210> 76

<211> 190

<212> DNA

<213> Anabaena sp.

<400> 76

gttatagaca aggcaatcct gagccaagcc gaagtagtaa ttagtaagtt aacaatagat 60

gacttacaac taatcggaag gtgcagagac tcgacgggag ctaccctaac gtcaagacga 120

gggtaaagag agagtccaat tctcaaagcc aataggcagt agcgaaagct gcaagagaat 180

gaaaatccgt 190

<210> 77

<211> 189

<212> DNA

<213> Anabaena sp.

<400> 77

ttatagacaa ggcaatcctg agccaagccg aagtagtaat tagtaagtta acaatagatg 60

acttacaact aatcggaagg tgcagagact cgacgggagc taccctaacg tcaagacgag 120

ggtaaagaga gagtccaatt ctcaaagcca ataggcagta gcgaaagctg caagagaatg 180

aaaatccgt 189

<210> 78

<211> 188

<212> DNA

<213> Anabaena sp.

<400>

78

tatagacaag gcaatcctga gccaagccga agtagtaatt agtaagttaa caatagatga 60

cttacaacta atcggaaggt gcagagactc gacgggagct accctaacgt caagacgagg 120

gtaaagagag agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga 180

aaatccgt 188

<210> 79

<211> 187

<212> DNA

<213> Anabaena sp.

<400> 79

atagacaagg caatcctgag ccaagccgaa gtagtaatta gtaagttaac aatagatgac 60

ttacaactaa tcggaaggtg cagagactcg acgggagcta ccctaacgtc aagacgaggg 120

taaagagaga gtccaattct caaagccaat aggcagtagc gaaagctgca agagaatgaa 180

aatccgt 187

<210> 80

<211> 186

<212> DNA

<213> Anabaena sp.

<400> 80

tagacaaggc aatcctgagc caagccgaag tagtaattag taagttaaca atagatgact 60

tacaactaat cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt 120

aaagagagag tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa 180

atccgt 186

<210> 81

<211> 139

<212> DNA

<213> Anabaena sp.

<400> 81

acaatagatg acttacaact aatcggaagg tgcagagact cgacgggagc taccctaacg 60

tcaagacgag ggtaaagaga gagtccaatt ctcaaagcca ataggcagta gcgaaagctg 120

caagagaatg aaaatccgt 139

<210> 82

<211> 138

<212> DNA

<213> Anabaena sp.

<400> 82

caatagatga cttacaacta atcggaaggt gcagagactc gacgggagct accctaacgt 60

caagacgagg gtaaagagag agtccaattc tcaaagccaa taggcagtag cgaaagctgc 120

aagagaatga aaatccgt 138

<210> 83

<211> 137

<212> DNA

<213> Anabaena sp.

<400> 83

aatagatgac ttacaactaa tcggaaggtg cagagactcg acgggagcta ccctaacgtc 60

aagacgaggg taaagagaga gtccaattct caaagccaat aggcagtagc gaaagctgca 120

agagaatgaa aatccgt 137

<210> 84

<211> 136

<212> DNA

<213> Anabaena sp.

<400> 84

atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac cctaacgtca 60

agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg aaagctgcaa 120

gagaatgaaa atccgt 136

<210> 85

<211> 135

<212> DNA

<213> Anabaena sp.

<400> 85

tagatgactt acaactaatc ggaaggtgca gagactcgac gggagctacc ctaacgtcaa 60

gacgagggta aagagagagt ccaattctca aagccaatag gcagtagcga aagctgcaag 120

agaatgaaaa tccgt 135

<210> 86

<211> 134

<212> DNA

<213> Anabaena sp.

<400>

86

agatgactta caactaatcg gaaggtgcag agactcgacg ggagctaccc taacgtcaag 60

acgagggtaa agagagagtc caattctcaa agccaatagg cagtagcgaa agctgcaaga 120

gaatgaaaat ccgt 134

<210> 87

<211> 133

<212> DNA

<213> Anabaena sp.

<400> 87

gatgacttac aactaatcgg aaggtgcaga gactcgacgg gagctaccct aacgtcaaga 60

cgagggtaaa gagagagtcc aattctcaaa gccaataggc agtagcgaaa gctgcaagag 120

aatgaaaatc cgt 133

<210> 88

<211> 132

<212> DNA

<213> Anabaena sp.

<400> 88

atgacttaca actaatcgga aggtgcagag actcgacggg agctacccta acgtcaagac 60

gagggtaaag agagagtcca attctcaaag ccaataggca gtagcgaaag ctgcaagaga 120

atgaaaatcc gt 132

<210> 89

<211> 131

<212> DNA

<213> Anabaena sp.

<400> 89

tgacttacaa ctaatcggaa ggtgcagaga ctcgacggga gctaccctaa cgtcaagacg 60

agggtaaaga gagagtccaa ttctcaaagc caataggcag tagcgaaagc tgcaagagaa 120

tgaaaatccg t 131

<210> 90

<211> 78

<212> DNA

<213> Anabaena sp.

<400> 90

caagacgagg gtaaagagag agtccaattc tcaaagccaa taggcagtag cgaaagctgc 60

aagagaatga aaatccgt 78

<210> 91

<211> 77

<212> DNA

<213> Anabaena sp.

<400> 91

aagacgaggg taaagagaga gtccaattct caaagccaat aggcagtagc gaaagctgca 60

agagaatgaa aatccgt 77

<210> 92

<211> 76

<212> DNA

<213> Anabaena sp.

<400> 92

agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg aaagctgcaa 60

gagaatgaaa atccgt 76

<210> 93

<211> 75

<212> DNA

<213> Anabaena sp.

<400> 93

gacgagggta aagagagagt ccaattctca aagccaatag gcagtagcga aagctgcaag 60

agaatgaaaa tccgt 75

<210> 94

<211> 74

<212> DNA

<213> Anabaena sp.

<400> 94

acgagggtaa agagagagtc caattctcaa agccaatagg cagtagcgaa agctgcaaga 60

gaatgaaaat ccgt 74

<210> 95

<211> 40

<212> DNA

<213> Anabaena sp.

<400> 95

aataggcagt agcgaaagct gcaagagaat gaaaatccgt 40

<210> 96

<211> 39

<212> DNA

<213> Anabaena sp.

<400> 96

ataggcagta gcgaaagctg caagagaatg aaaatccgt 39

<210> 97

<211> 38

<212> DNA

<213> Anabaena sp.

<400> 97

taggcagtag cgaaagctgc aagagaatga aaatccgt 38

<210> 98

<211> 37

<212> DNA

<213> Anabaena sp.

<400> 98

aggcagtagc gaaagctgca agagaatgaa aatccgt 37

<210> 99

<211> 36

<212> DNA

<213> Anabaena sp.

<400> 99

ggcagtagcg aaagctgcaa gagaatgaaa atccgt 36

<210> 100

<211> 27

<212> DNA

<213> Anabaena sp.

<400> 100

gaaagctgca agagaatgaa aatccgt 27

<210> 101

<211> 26

<212> DNA

<213> Anabaena sp.

<400> 101

aaagctgcaa gagaatgaaa atccgt 26

<210> 102

<211> 25

<212> DNA

<213> Anabaena sp.

<400> 102

aagctgcaag agaatgaaaa tccgt 25

<210> 103

<211> 24

<212> DNA

<213> Anabaena sp.

<400> 103

agctgcaaga gaatgaaaat ccgt 24

<210> 104

<211> 23

<212> DNA

<213> Anabaena sp.

<400> 104

gctgcaagag aatgaaaatc cgt 23

<210> 105

<211> 22

<212> DNA

<213> Anabaena sp.

<400> 105

ctgcaagaga atgaaaatcc gt 22

<210> 106

<211> 18

<212> DNA

<213> Anabaena sp.

<400> 106

aagagaatga aaatccgt 18

<210> 107

<211> 17

<212> DNA

<213> Anabaena sp.

<400> 107

agagaatgaa aatccgt 17

<210> 108

<211> 16

<212> DNA

<213> Anabaena sp.

<400> 108

gagaatgaaa atccgt 16

<210> 109

<211> 35

<212> DNA

<213> Anabaena sp.

<400> 109

gcagtagcga aagctgcaag agaatgaaaa tccgt 35

<210> 110

<211> 33

<212> DNA

<213> Anabaena sp.

<400> 110

agtagcgaaa gctgcaagag aatgaaaatc cgt 33

<210> 111

<211> 32

<212> DNA

<213> Anabaena sp.

<400> 111

gtagcgaaag ctgcaagaga atgaaaatcc gt 32

<210> 112

<211> 26

<212> DNA

<213> Anabaena sp.

<400> 112

acggacttaa ataattgagc cttaaa 26

<210> 113

<211> 27

<212> DNA

<213> Anabaena sp.

<400> 113

acggacttaa ataattgagc cttaaag 27

<210> 114

<211> 28

<212> DNA

<213> Anabaena sp.

<400> 114

acggacttaa ataattgagc cttaaaga 28

<210> 115

<211> 76

<212> DNA

<213> Anabaena sp.

<400> 115

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatcta 76

<210> 116

<211> 77

<212> DNA

<213> Anabaena sp.

<400> 116

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctag 77

<210> 117

<211> 78

<212> DNA

<213> Anabaena sp.

<400> 117

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagt 78

<210> 118

<211> 79

<212> DNA

<213> Anabaena sp.

<400> 118

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtt 79

<210> 119

<211> 80

<212> DNA

<213> Anabaena sp.

<400> 119

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta 80

<210> 120

<211> 127

<212> DNA

<213> Anabaena sp.

<400> 120

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagtta 127

<210> 121

<211> 128

<212> DNA

<213> Anabaena sp.

<400> 121

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaa 128

<210> 122

<211> 129

<212> DNA

<213> Anabaena sp.

<400> 122

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaac 129

<210> 123

<211> 130

<212> DNA

<213> Anabaena sp.

<400

> 123

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca 130

<210> 124

<211> 131

<212> DNA

<213> Anabaena sp.

<400> 124

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca a 131

<210> 125

<211> 132

<212> DNA

<213> Anabaena sp.

<400> 125

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca at 132

<210> 126

<211> 133

<212> DNA

<213> Anabaena sp.

<400> 126

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca ata 133

<210> 127

<211> 134

<212> DNA

<213> Anabaena sp.

<400> 127

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atag 134

<210> 128

<211> 135

<212> DNA

<213> Anabaena sp.

<400> 128

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca ataga 135

<210> 129

<211> 188

<212> DNA

<213> Anabaena sp.

<400> 129

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgt 188

<210> 130

<211> 189

<212> DNA

<213> Anabaena sp.

<400> 130

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtc 189

<210> 131

<211> 190

<212> DNA

<213> Anabaena sp.

<400> 131

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca 190

<210> 132

<211> 191

<212> DNA

<213> Anabaena sp.

<400> 132

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca a 191

<210> 133

<211> 192

<212> DNA

<213> Anabaena sp.

<400> 133

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca ag 192

<210> 134

<211> 226

<212> DNA

<213> Anabaena sp.

<400

> 134

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagcc 226

<210> 135

<211> 227

<212> DNA

<213> Anabaena sp.

<400> 135

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagcca 227

<210> 136

<211> 228

<212> DNA

<213> Anabaena sp.

<400> 136

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaa 228

<210> 137

<211> 229

<212> DNA

<213> Anabaena sp.

<400> 137

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaat 229

<210> 138

<211> 230

<212> DNA

<213> Anabaena sp.

<400> 138

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 230

<210> 139

<211> 239

<212> DNA

<213> Anabaena sp.

<400> 139

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagc 239

<210> 140

<211> 240

<212> DNA

<213> Anabaena sp.

<400> 140

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 240

<210> 141

<211> 241

<212> DNA

<213> Anabaena sp.

<400

> 141

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 240

a 241

<210> 142

<211> 242

<212> DNA

<213> Anabaena sp.

<400> 142

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 240

aa 242

<210> 143

<211> 243

<212> DNA

<213> Anabaena sp.

<400> 143

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 240

aaa 243

<210> 144

<211> 244

<212> DNA

<213> Anabaena sp.

<400> 144

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 240

aaag 244

<210> 145

<211> 248

<212> DNA

<213> Anabaena sp.

<400> 145

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 240

aaagctgc 248

<210> 146

<211> 249

<212> DNA

<213> Anabaena sp.

<400> 146

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 240

aaagctgca 249

<210> 147

<211> 250

<212> DNA

<213> Anabaena sp.

<400> 147

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 240

aaagctgcaa 250

<210> 148

<211> 231

<212> DNA

<213> Anabaena sp.

<400

> 148

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata g 231

<210> 149

<211> 233

<212> DNA

<213> Anabaena sp.

<400> 149

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggc 233

<210> 150

<211> 234

<212> DNA

<213> Anabaena sp.

<400> 150

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggca 234

<210> 151

<211> 104

<212> DNA

<213> Azoarcus sp.

<400> 151

tgcgccgatg aaggtgtaga gactagacgg cacccaccta aggcaaacgc tatggtgaag 60

gcatagtcca gggagtggcg aaagtcacac aaaccggaat ccgt 104

<210> 152

<211> 141

<212> DNA

<213> Azoarcus sp.

<400> 152

ccgggcgtat ggcaacgccg agccaagctt cggcgcctgc gccgatgaag gtgtagagac 60

tagacggcac ccacctaagg caaacgctat ggtgaaggca tagtccaggg agtggcgaaa 120

gtcacacaaa ccggaatccg t 141

<210> 153

<211> 78

<212> DNA

<213> Azoarcus sp.

<400> 153

acggcaccca cctaaggcaa acgctatggt gaaggcatag tccagggagt ggcgaaagtc 60

acacaaaccg gaatccgt 78

<210> 154

<211> 58

<212> DNA

<213> Azoarcus sp.

<400> 154

acgctatggt gaaggcatag tccagggagt ggcgaaagtc acacaaaccg gaatccgt 58

<210> 155

<211> 75

<212> DNA

<213> Thiomargarita sp.

<400> 155

attaaagtta tagaattatc agagaatgat atagtccaag ccttatggta acatgagggc 60

acttgaccct ggtag 75

<210> 156

<211> 195

<212> DNA

<213> Staphylococcus virus Twort

<400> 156

aagatgtagg caatcctgag ctaagctctt agtaataaga gaaagtgcaa cgactattcc 60

gataggaagt agggtcaagt gactcgaaat ggggattacc cttctagggt agtgatatag 120

tctgaacata tatggaaaca tatagaagga taggagtaac gaacctattc gtaacataat 180

tgaactttta gttat 195

<210> 157

<211> 163

<212> DNA

<213> Staphylococcus virus Twort

<400> 157

taataagaga aagtgcaacg actattccga taggaagtag ggtcaagtga ctcgaaatgg 60

ggattaccct tctagggtag tgatatagtc tgaacatata tggaaacata tagaaggata 120

ggagtaacga acctattcgt aacataattg aacttttagt tat 163

<210> 158

<211> 133

<212> DNA

<213> Staphylococcus virus Twort

<400> 158

taggaagtag ggtcaagtga ctcgaaatgg ggattaccct tctagggtag tgatatagtc 60

tgaacatata tggaaacata tagaaggata ggagtaacga acctattcgt aacataattg 120

aacttttagt tat 133

<210> 159

<211> 92

<212> DNA

<213> Staphylococcus virus Twort

<400> 159

ctagggtagt gatatagtct gaacatatat ggaaacatat agaaggatag gagtaacgaa 60

cctattcgta acataattga acttttagtt at 92

<210> 160

<211> 92

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

LSUp1 sequence"

<400> 160

agttaataaa gatgatgaaa tagtctgaac cattttgaga aaagtggaaa taaaagaaaa 60

tcttttatga taacataaat tgaacaggct aa 92

<210> 161

<211> 62

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bacteriophage Phi I sequence"

<400> 161

caaagactga tgatatagtc cgacactcct agtaatagga gaatacagaa aggatgaaat 60

cc 62

<210> 162

<211> 76

<212> DNA

<213> Nostoc sp.

<400> 162

agtcgagggt aaagggagag tccaattctc aaagcctatt ggcagtagcg aaagctgcgg 60

gagaatgaaa atccgt 76

<210> 163

<211> 76

<212> DNA

<213> Nostoc sp.

<400> 163

agccgagggt aaagggagag tccaattctc aaagccaata ggcagtagcg aaagctgcgg 60

gagaatgaaa atccgt 76

<210> 164

<211> 89

<212> DNA

<213> Nodularia sp.

<400> 164

agccgagggt aaagggagag tccaattctc aaagccgaag gttattaaaa cctggcagca 60

gtgaaagctg cgggagaatg aaaatccgt 89

<210> 165

<211> 79

<212> DNA

<213> Pleurocapsa sp.

<400> 165

agctgagggt aaagagagag tccaattctc aaagccagca gatggcagta gcgaaagctg 60

cgggagaatg aaaatccgt 79

<210> 166

<211> 76

<212> DNA

<213> Planktothrix sp.

<400> 166

agccgagggt aaagagagag tccaattctc aaagccaatt ggtagtagcg aaagctacgg 60

gagaatgaaa atccgt 76

<210> 167

<211> 118

<212> DNA

<213> Azoarcus sp.

<400> 167

gcggactcat atttcgatgt gccttgcgcc gggaaaccac gcaagggatg gtgtcaaatt 60

cggcgaaacc taagcgcccg cccgggcgta tggcaacgcc gagccaagct tcggcgcc 118

<210> 168

<211> 81

<212> DNA

<213> Azoarcus sp.

<400> 168

gcggactcat atttcgatgt gccttgcgcc gggaaaccac gcaagggatg gtgtcaaatt 60

cggcgaaacc taagcgcccg c 81

<210> 169

<211> 144

<212> DNA

<213> Azoarcus sp.

<400> 169

gcggactcat atttcgatgt gccttgcgcc gggaaaccac gcaagggatg gtgtcaaatt 60

cggcgaaacc taagcgcccg cccgggcgta tggcaacgcc gagccaagct tcggcgcctg 120

cgccgatgaa ggtgtagaga ctag 144

<210> 170

<211> 164

<212> DNA

<213> Azoarcus sp.

<400> 170

gcggactcat atttcgatgt gccttgcgcc gggaaaccac gcaagggatg gtgtcaaatt 60

cggcgaaacc taagcgcccg cccgggcgta tggcaacgcc gagccaagct tcggcgcctg 120

cgccgatgaa ggtgtagaga ctagacggca cccacctaag gcaa 164

<210> 171

<211> 179

<212> DNA

<213> Thiomargarita sp.

<400> 171

aggattagat actacactaa gtgtccccca gactggtgac agtctggtgt gcatccagct 60

atatcggtga aaccccattg gggtaatacc gagggaagct atattatata tatattaata 120

aatagccccg tagagactat gtaggtaagg agatagaaga tgataaaatc aaaatcatc 179

<210> 172

<211> 84

<212> DNA

<213> Staphylococcus virus Twort

<400> 172

actactgaaa gcataaataa ttgtgccttt atacagtaat gtatatcgaa aaatcctcta 60

attcagggaa cacctaaaca aact 84

<210> 173

<211> 116

<212> DNA

<213> Staphylococcus virus Twort

<400> 173

actactgaaa gcataaataa ttgtgccttt atacagtaat gtatatcgaa aaatcctcta 60

attcagggaa cacctaaaca aactaagatg taggcaatcc tgagctaagc tcttag 116

<210> 174

<211> 146

<212> DNA

<213> Staphylococcus virus Twort

<400> 174

actactgaaa gcataaataa ttgtgccttt atacagtaat gtatatcgaa aaatcctcta 60

attcagggaa cacctaaaca aactaagatg taggcaatcc tgagctaagc tcttagtaat 120

aagagaaagt gcaacgacta ttccga 146

<210> 175

<211> 187

<212> DNA

<213> Staphylococcus virus Twort

<400> 175

actactgaaa gcataaataa ttgtgccttt atacagtaat gtatatcgaa aaatcctcta 60

attcagggaa cacctaaaca aactaagatg taggcaatcc tgagctaagc tcttagtaat 120

aagagaaagt gcaacgacta ttccgatagg aagtagggtc aagtgactcg aaatggggat 180

taccctt 187

<210> 176

<211> 216

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

LSUp1 sequence"

<400> 176

cgctagggat ttataactgt gagtcctcca atattataaa atgttggtaa tatattgggt 60

aaatttcaaa gacaactttt ctccacgtca ggatatagtg tatttgaagc gaaacttatt 120

ttagcagtga aaaagcaaat aaggacgttc aacgactaaa aggtgagtat tgctaacaat 180

aatccttttt tttaatgccc aacatcttta ttaact 216

<210> 177

<211> 559

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bacteriophage Phi I sequence"

<400> 177

gtgggtgcat aaactatttc attgtgcaca ttaaatctgg tgaactcggt gaaaccctaa 60

tggggcaata ccgagccaag ccatagggag gatatatgag aggcaagaag ttaattcttg 120

aggccactga gactggctgt atcatcccta cgtcacacaa acttaatgcc gatggttatt 180

tcagaaagaa aaccaatggc gtcttagaga tgtatcacag aacggtgtgg aaggagcata 240

acggagacat acctgatggc ttcgagatag accataagtg tcgcaatagg gcttgctgta 300

atatagagca tttacagatg cttgagggta cagcccacac tgttaagacc aatcgtgaac 360

gctacgcaga cagaaaggaa acagctaggg aatactggct ggagactgga tgtaccggcc 420

tagcactcgg tgagaagttt ggtgtgtcgt tctcttctgc ttgtaagtgg attagagaat 480

ggaaggcgta gagactatcc gaaaggagta gggccgaggg tgagactccc tcgtaacccg 540

aagcgccaga cagtcaact 559

<210> 178

<211> 218

<212> DNA

<213> Nostoc sp.

<400> 178

acggacttaa gtaattgagc cttaaagaag aaattcttta agtggcagct ctcaaactca 60

gggaaaccta aatctgttca cagacaaggc aatcctgagc caagccgaaa gagtcatgag 120

tgctgagtag tgagtaaaat aaaagctcac aactcagagg ttgtaactct aagctagtcg 180

gaaggtgcag agactcgacg ggagctaccc taacgtaa 218

<210> 179

<211> 165

<212> DNA

<213> Nostoc sp.

<400> 179

acggacttaa actgaattga gccttagaga agaaattctt taagtgtcag ctctcaaact 60

cagggaaacc taaatctgtt gacagacaag gcaatcctga gccaagccga gaactctaag 120

ttattcggaa ggtgcagaga ctcgacggga gctaccctaa cgtca 165

<210> 180

<211> 181

<212> DNA

<213> Nodularia sp.

<400> 180

acggacttag aaaactgagc cttgatcgag aaatctttca agtggaagct ctcaaattca 60

gggaaaccta aatctgttta cagatatggc aatcctgagc caagccgaaa caagtcctga 120

gtgttaaagc tcataactca tcggaaggtg cagagactcg acgggagcta ccctaacgtt 180

a 181

<210> 181

<211> 188

<212> DNA

<213> Pleurocapsa sp.

<400> 181

acggacttaa aaaaattgag ccttggcaga gaaatctgtc atgcgaacgc tctcaaattc 60

agggaaacct aagtctggca acagatatgg caatcctgag ccaagcctta atcaaggaaa 120

aaaacatttt tacctttac cttgaaagga aggtgcagag actcaacggg agctacccta 180

acaggtca 188

<210> 182

<211> 200

<212> DNA

<213> Planktothrix sp.

<400> 182

acggacttaa agataaattg agccttgagg cgagaaatct ctcaagtgta agctgtcaaa 60

ttcagggaaa cctaaatctg taaattcaga caaggcaatc ctgagccaag cctaggggta 120

ttagaaatga gggagtttcc ccaatctaag atcaatacct aggaaggtgc agagactcga 180

cgggagctac cctaacgtta 200

<210> 183

<211> 166

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 183

agtatataag aaacaaacca ctagatgact tacaactaat cggaaggtgc agagactcga 60

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 120

ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct cgcagc 166

<210> 184

<211> 176

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 184

ctgaaattat acttatactc aaacaaacca ctagatgact tacaactaat cggaaggtgc 60

agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc 120

aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct cgcagc 176

<210> 185

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 185

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 186

<211> 196

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 186

catcaacaat atgaaattat acttatactc agtatatgac aaacaaacca ctagatgact 60

tacaactaat cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt 120

aaagagagag tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa 180

atccgtggct cgcagc 196

<210> 187

<211> 206

<212> DNA

<213> Artificial Sequence

<220><221> source

<223

> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 187

catcaacaat atgaaactat acttatactc agtatatgaa gcattatcgc aaacaaacca 60

ctagatgact tacaactaat cggaaggtgc agagactcga cgggagctac cctaacgtca 120

agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg aaagctgcaa 180

gagaatgaaa atccgtggct cgcagc 206

<210> 188

<211> 166

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 188

tagcgtcagc aaacaaacaa atagatgact tacaactaat cggaaggtgc agagactcga 60

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 120

ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct cgcagc 166

<210> 189

<211> 176

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 189

atactcatac tagcgtcagc aaacaaacaa atagatgact tacaactaat cggaaggtgc 60

agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc 120

aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct cgcagc 176

<210> 190

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 190

gtgtgaagct atactcatac tagcgtcagc aaacaaacaa atagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 191

<211> 196

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 191

cctcacctga gtgtgaagct atactcatac tagcgtcagc aaacaaacaa atagatgact 60

tacaactaat cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt 120

aaagagagag tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa 180

atccgtggct cgcagc 196

<210> 192

<211> 206

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 192

ccgaatgatg cctcacctga gtgtgaagct atactcatac tagcgtcagc aaacaaacaa 60

atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac cctaacgtca 120

agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg aaagctgcaa 180

gagaatgaaa atccgtggct cgcagc 206

<210> 193

<211> 166

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 193

cggtgcgagc aaacaaacaa atagatgact tacaactaat cggaaggtgc agagactcga 60

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 120

ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct cgcagc 166

<210> 194

<211> 176

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 194

cgctccgacc cagtgcgagc aaacaaacaa atagatgact tacaactaat cggaaggtgc 60

agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc 120

aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct cgcagc 176

<210> 195

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 195

ctgaaattat actaatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 196

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 196

ctgaaaatat actaatactc actatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 197

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 197

ctgataatat agtaatactc actatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 198

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 198

ctgataataa agtaatacac actataagac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 199

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 199

ctgaaattat acttatactc tctaagttac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 200

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 200

ctgaaattat gtgtgttaca tctaagttac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 201

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 201

gttgatcggt gtgtgttaca tctaagttac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 202

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223
> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 202

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtgatt 180

aaacag 186

<210> 203

<211> 196

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 203

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtgatt 180

cacaatataa attacg 196

<210> 204

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 204

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggat 180

catagc 186

<210> 205

<211> 196

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 205

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggat 180

cgcagcataa tatccg 196

<210> 206

<211> 196

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 206

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagcgcgc ctaccg 196

<210> 207

<211> 219

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 207

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagcgcgc ctaccgaaag ccggcgtcga cgttagcgc 219

<210> 208

<211> 219

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 208

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggat 180

cgcagcataa tatccgaaac gaggatacaa gtgacatgc 219

<210> 209

<211> 219

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 209

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtgatt 180

cacaatctaa attacgaaac gataaatgat aactctaac 219

<210> 210

<211> 156

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 210

aaacaaacca ctagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 60

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata ggcagtagcg 120

aaagctgcaa gagaatgaaa atccgtggct cgcagc 156

<210> 211

<211> 161

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 211

cgggcaaaca aacaaataga tgacttacaa ctaatcggaa ggtgcagaga ctcgacggga 60

gctaccctaa cgtcaagacg agggtaaaga gagagtccaa ttctcaaagc caataggcag 120

tagcgaaagc tgcaagagaa tgaaaatccg tggctcgcag c 161

<210> 212

<211> 186

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 212

cgctccgacg agcttccggc cagtgcgagc aaacaaacaa atagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgtggct 180

cgcagc 186

<210> 213

<211> 57

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 213

aaacaaacca cggcagtagc gaaagctgca agagaatgaa aatccgtggc tcgcagc 57

<210> 214

<211> 67

<

212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 214

agtatataag aaacaaacca cggcagtagc gaaagctgca agagaatgaa aatccgtggc 60

tcgcagc 67

<210> 215

<211> 77

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 215

ctgaaattat acttatactc aaacaaacca cggcagtagc gaaagctgca agagaatgaa 60

aatccgtggc tcgcagc 77

<210> 216

<211> 87

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 216

ctgaaattat acttatactc agtatatgac aaacaaacca cggcagtagc gaaagctgca 60

agagaatgaa aatccgtggc tcgcagc 87

<210> 217

<211> 67

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 217

tagcgtcagc aaacaaacaa aggcagtagc gaaagctgca agagaatgaa aatccgtggc 60

tcgcagc 67

<210> 218

<211> 77

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 218

atactcatac tagcgtcagc aaacaaacaa aggcagtagc gaaagctgca agagaatgaa 60

aatccgtggc tcgcagc 77

<210> 219

<211> 87

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 219

gtgtgaagct atactcatac tagcgtcagc aaacaaacaa aggcagtagc gaaagctgca 60

agagaatgaa aatccgtggc tcgcagc 87

<210> 220

<211> 67

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 220

cggtgcgagc aaacaaacaa aggcagtagc gaaagctgca agagaatgaa aatccgtggc 60

tcgcagc 67

<210> 221

<211> 77

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 221

cgctccgacc cagtgcgagc aaacaaacaa aggcagtagc gaaagctgca agagaatgaa 60

aatccgtggc tcgcagc 77

<210> 222

<211> 87

<

212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 222

cgctccgacg agcttccggc cagtgcgagc aaacaaacaa aggcagtagc gaaagctgca 60

agagaatgaa aatccgtggc tcgcagc 87

<210> 223

<211> 98

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 223

aaacaaacca caagacgagg gtaaagagag agtccaattc tcaaagccaa taggcagtag 60

cgaaagctgc aagagaatga aaatccgtgg ctcgcagc 98

<210> 224

<211> 108

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 224

agtatataag aaacaaacca caagacgagg gtaaagagag agtccaattc tcaaagccaa 60

taggcagtag cgaaagctgc aagagaatga aaatccgtgg ctcgcagc 108

<210> 225

<211> 118

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 225

ctgaaattat acttatactc aaacaaacca caagacgagg gtaaagagag agtccaattc 60

tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg ctcgcagc 118

<210> 226

<211> 128

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 226

ctgaaattat acttatactc agtatatgac aaacaaacca caagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg 120

ctcgcagc 128

<210> 227

<211> 108

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 227

tagcgtcagc aaacaaacaa aaagacgagg gtaaagagag agtccaattc tcaaagccaa 60

taggcagtag cgaaagctgc aagagaatga aaatccgtgg ctcgcagc 108

<210> 228

<211> 118

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 228

atactcatac tagcgtcagc aaacaaacaa aaagacgagg gtaaagagag agtccaattc 60

tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg ctcgcagc 118

<210> 229

<211> 128

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 229

gtgtgaagct atactcatac tagcgtcagc aaacaaacaa aaagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg 120

ctcgcagc 128

<210> 230

<211> 108

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 230

cggtgcgagc aaacaaacaa aaagacgagg gtaaagagag agtccaattc tcaaagccaa 60

taggcagtag cgaaagctgc aagagaatga aaatccgtgg ctcgcagc 108

<210> 231

<211> 118

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 231

cgctccgacc cagtgcgagc aaacaaacaa aaagacgagg gtaaagagag agtccaattc 60

tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg ctcgcagc 118

<210> 232

<211> 128

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 232

cgctccgacg agcttccggc cagtgcgagc aaacaaacaa aaagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg 120

ctcgcagc 128

<210> 233

<211> 176

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 233

ctgaaattat acttatactc agtatatgac aaacaaacca ctagatgact tacaactaat 60

cggaaggtgc agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag 120

tccaattctc aaagccaata ggcagtagcg aaagctgcaa gagaatgaaa atccgt 176

<210> 234

<211> 77

<

212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 234
ctgaaattat acttatactc agtatatgac aaacaaacca cggcagtagc gaaagctgca 60
agagaatgaa aatccgt 77
<210> 235
<211> 87
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 235
ctgaaattat acttatactc agtatatgac aaacaaacca cggcagtagc gaaagctgca 60

agagaatgaa aatccgtgat taaacag 87
<210> 236
<211> 97
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 236
ctgaaattat acttatactc agtatatgac aaacaaacca cggcagtagc gaaagctgca 60
agagaatgaa aatccgtgat tcacaatata aattacg 97
<210> 237
<211> 87
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"
<400> 237

ctgaaattat acttatactc agtatatgac aaacaaacca cggcagtagc gaaagctgca 60

agagaatgaa aatccgtgga tcatagc 87

<210> 238

<211> 97

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 238

ctgaaattat acttatactc agtatatgac aaacaaacca cggcagtagc gaaagctgca 60

agagaatgaa aatccgtgga tcgcagcata atatccg 97

<210> 239

<211> 97

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 239

ctgaaattat acttatactc agtatatgac aaacaaacca cggcagtagc gaaagctgca 60

agagaatgaa aatccgtggc tcgcagcgcg cctaccg 97

<210> 240

<211> 118

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 240

ctgaaattat acttatactc agtatatgac aaacaaacca caagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgt 118

<210> 241

<211> 128

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 241

ctgaaattat acttatactc agtatatgac aaacaaacca caagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtga 120

ttaaacag 128

<210> 242

<211> 138

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 242

ctgaaattat acttatactc agtatatgac aaacaaacca caagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtga 120

ttcacaatat aaattacg 138

<210> 243

<211> 128

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 243

ctgaaattat acttatactc agtatatgac aaacaaacca caagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg 120

atcatagc 128

<210> 244

<211> 138

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 244

ctgaaattat acttatactc agtatatgac aaacaaacca caagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg 120

atcgcagcat aatatccg 138

<210> 245

<211> 138

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 245

ctgaaattat acttatactc agtatatgac aaacaaacca caagacgagg gtaaagagag 60

agtccaattc tcaaagccaa taggcagtag cgaaagctgc aagagaatga aaatccgtgg 120

ctcgcagcgc gcctaccg 138

<210> 246

<211> 162

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 246

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aacttatata ct 162

<210> 247

<211> 172

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 247

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagagtataa gtataatttc ag 172

<210> 248

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 248

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc 180

ag 182

<210> 249

<211> 192

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 249

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc 180

atattgttga tg 192

<210> 250

<211> 202

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 250

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagcgataat gcttcatata ctgagtataa 180

gtatagtttc atattgttga tg 202

<210> 251

<211> 162

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 251

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagctgacgc ta 162

<210> 252

<211> 172

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 252

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagctgacgc tagtatgagt at 172

<210> 253

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 253

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagctgacgc tagtatgagt atagcttcac 180

ac 182

<210> 254

<211> 192

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 254

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagctgacgc tagtatgagt atagcttcac 180

actcaggtga gg 192

<210> 255

<211> 202

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 255

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagctgacgc tagtatgagt atagcttcac 180

actcaggtga ggcatcattc gg 202

<210> 256

<211> 162

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 256

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagctcgcac cg 162

<210> 257

<211> 172

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 257

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagctcgcac tgggtcggag cg 172

<210> 258

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 258

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc 180

ag 182

<210> 259

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 259

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc 180

ag 182

<210> 260

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223
> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 260

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc 180

ag 182

<210> 261

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 261

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc 180

ag 182

<210> 262

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 262

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtaactta gagagtataa gtataatttc 180

ag 182

<210> 263

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 263

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtaactta gatgtaacac acataatttc 180

ag 182

<210> 264

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 264

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtaactta gatgtaacac acaccgatca 180

ac 182

<210> 265

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 265

ctgtttaatc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc 180

ag 182

<210> 266

<211> 192

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 266

cgtaatttat attgtgaatc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa 180

gtataatttc ag 192

<210> 267

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 267

gctatgatcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc 180

ag 182

<210> 268

<211> 192

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 268

cggatattat gctgcgatcc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa 180

gtataatttc ag 192

<210> 269

<211> 192

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 269

cggtaggcgc gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca acacaaacac aagtcatata ctgagtataa 180

gtataatttc ag 192

<210> 270

<211> 215

<212> DNA

<213> Artificial Sequence

<220><221> source

<223

> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 270

gcgctaacgt cgacgccggc aaacggtagg cgcgctgcga gccacggact taaataattg 60

agccttaaag aagaaattct ttaagtggat gctctcaaac tcagggaaac ctaaatctag 120

ttatagacaa ggcaatcctg agccaagccg aagtagtaat tagtaagtta acaacacaaa 180

cacaagtcat atactgagta taagtataat ttcag 215

<210> 271

<211> 215

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 271

gcatgtcact tgtatcctcg aaacggatat tatgctgcga tccacggact taaataattg 60

agccttaaag aagaaattct ttaagtggat gctctcaaac tcagggaaac ctaaatctag 120

ttatagacaa ggcaatcctg agccaagccg aagtagtaat tagtaagtta acaacacaaa 180

cacaagtcat atactgagta taagtataat ttcag 215

<210> 272

<211> 215

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 272

gttagagtta tcatttatcg aaacgtaatt tagattgtga atcacggact taaataattg 60

agccttaaag aagaaattct ttaagtggat gctctcaaac tcagggaaac ctaaatctag 120

ttatagacaa ggcaatcctg agccaagccg aagtagtaat tagtaagtta acaacacaaa 180

cacaagtcat atactgagta taagtataat ttcag 215

<210> 273

<211> 152

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 273

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca acacaaacac aa 152

<210> 274

<211> 157

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 274

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagcccg 157

<210> 275

<211> 182

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 275

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca aaacaaaaac aagctcgcac tggccggaag ctcgtcggag 180

cg 182

<210> 276

<211> 251

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 276

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

cacaaacaca a 251

<210> 277

<211> 261

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 277

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

cacaaacaca acttatatac t 261

<210> 278

<211> 271

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 278

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

cacaaacaca agagtataag tataatttca g 271

<210> 279

<211> 281

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 279

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

cacaaacaca agtcatatac tgagtataag tataatttca g 281

<210> 280

<211> 261

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 280

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

aacaaaaaca agctgacgct a 261

<210> 281

<211> 271

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 281

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

aacaaaaaca agctgacgct agtatgagta t 271

<210> 282

<211> 281

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 282

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

aacaaaaaca agctgacgct agtatgagta tagcttcaca c 281

<210> 283

<211> 261

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 283

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

aacaaaaaca agctcgcacc g 261

<210> 284

<211> 271

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 284

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

aacaaaaaca agctcgcact gggtcggagc g 271

<210> 285

<211> 281

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 285

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

aacaaaaaca agctcgcact ggccggaagc tcgtcggagc g 281

<210> 286

<211> 210

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 286

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtcc acaaacacaa 210

<210> 287

<211> 220

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 287

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtcc acaaacacaa cttatatact 220

<210> 288

<211> 230

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 288

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtcc acaaacacaa gagtataagt ataatttcag 230

<210> 289

<211> 240

<212> DNA

<213> Artificial Sequence

<220><221> source

<223

> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 289

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtcc acaaacacaa gtcatatact gagtataagt ataatttcag 240

<210> 290

<211> 220

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 290

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca acaaaaacaa gctgacgcta 220

<210> 291

<211> 230

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 291

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca acaaaaacaa gctgacgcta gtatgagtat 230

<210> 292

<211> 240

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 292

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca acaaaaacaa gctgacgcta gtatgagtat agcttcacac 240

<210> 293

<211> 220

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 293

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca acaaaaacaa gctcgcaccg 220

<210> 294

<211> 230

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 294

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca acaaaaacaa gctcgcactg ggtcggagcg 230

<210> 295

<211> 240

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 295

gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca acaaaaacaa gctcgcactg gccggaagct cgtcggagcg 240

<210> 296

<211> 172

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 296

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca acacaaacac aagtcatata ctgagtataa gtataatttc ag 172

<210> 297

<211> 271

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 297

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata cacaaacaca 240

agtcatatac tgagtataag tataatttca g 271

<210> 298

<211> 281

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 298

ctgtttaatc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

cacaaacaca agtcatatac tgagtataag tataatttca g 281

<210> 299

<211> 291

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 299

cgtaatttat attgtgaatc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc 180

agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc 240

aaagccaata cacaaacaca agtcatatac tgagtataag tataatttca g 291

<210> 300

<211> 281

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 300

gctatgatcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc aaagccaata 240

cacaaacaca agtcatatac tgagtataag tataatttca g 281

<210> 301

<211> 291

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 301

cggatattat gctgcgatcc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc 180

agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc 240

aaagccaata cacaaacaca agtcatatac tgagtataag tataatttca g 291

<210> 302

<211> 291

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 302

cggtaggcgc gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc 180

agagactcga cgggagctac cctaacgtca agacgagggt aaagagagag tccaattctc 240

aaagccaata cacaaacaca agtcatatac tgagtataag tataatttca g 291

<210> 303

<211> 230

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 303

acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct ctcaaactca 60

gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag tagtaattag 120

taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga cgggagctac 180

cctaacgtcc acaaacacaa gtcatatact gagtataagt ataatttcag 230

<210> 304

<211> 240

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 304

ctgtttaatc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtcc acaaacacaa gtcatatact gagtataagt ataatttcag 240

<210> 305

<211> 250

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 305

cgtaatttat attgtgaatc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc 180

agagactcga cgggagctac cctaacgtcc acaaacacaa gtcatatact gagtataagt 240

ataatttcag 250

<210> 306

<211> 240

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 306

gctatgatcc acggacttaa ataattgagc cttaaagaag aaattcttta agtggatgct 60

ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc caagccgaag 120

tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc agagactcga 180

cgggagctac cctaacgtcc acaaacacaa gtcatatact gagtataagt ataatttcag 240

<210> 307

<211> 250

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 307

cggatattat gctgcgatcc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc 180

agagactcga cgggagctac cctaacgtcc acaaacacaa gtcatatact gagtataagt 240

ataatttcag 250

<210> 308

<211> 250

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 308

cggtaggcgc gctgcgagcc acggacttaa ataattgagc cttaaagaag aaattcttta 60

agtggatgct ctcaaactca gggaaaccta aatctagtta tagacaaggc aatcctgagc 120

caagccgaag tagtaattag taagttaaca atagatgact tacaactaat cggaaggtgc 180

agagactcga cgggagctac cctaacgtcc acaaacacaa gtcatatact gagtataagt 240

ataatttcag 250

<210> 309

<211> 1137

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 309

atgctgctgc tggtcacatc tctgctgctg tgcgagctgc cccatcctgc ctttctgctg 60

atccccgaca tccagatgac ccagaccaca agcagcctgt ctgccagcct gggcgataga 120

gtgaccatca gctgtagagc cagccaggac atcagcaagt acctgaactg gtatcagcaa 180

aagcccgacg gcaccgtgaa gctgctgatc taccacacca gcagactgca cagcggcgtg 240

ccaagcagat tttctggcag cggctctggc accgactaca gcctgacaat cagcaacctg 300

gaacaagagg atatcgctac ctacttctgc cagcaaggca acaccctgcc ttacaccttt 360

ggcggaggca ccaagctgga aatcaccggc tctacaagcg gcagcggcaa acctggatct 420

ggcgagggat ctaccaaggg cgaagtgaaa ctgcaagagt ctggccctgg actggtggcc 480

ccatctcagt ctctgagcgt gacctgtaca gtcagcggag tgtccctgcc tgattacggc 540

gtgtcctgga tcagacagcc tcctcggaaa ggcctggaat ggctgggagt gatctggggc 600

agcgagacaa cctactacaa cagcgccctg aagtcccggc tgaccatcat caaggacaac 660

tccaagagcc aggtgttcct gaagatgaac agcctgcaga ccgacgacac cgccatctac 720

tattgcgcca agcactacta ctacggcggc agctacgcca tggattattg gggccagggc 780

accagcgtga ccgtttcttc tgccgccgct atcgaagtga tgtaccctcc tccttacctg 840

gacaacgaga agtccaacgg caccatcatc cacgtgaagg gcaagcacct gtgtccttct 900

ccactgttcc ccggacctag caagcctttc tgggtgctcg ttgttgttgg cggcgtgctg 960

gcctgttaca gcctgctggt taccgtggcc ttcatcatct tttgggtcaa gagaggccgg 1020

aagaaacttc tttatatatt caagcagccc tttatgcgac ccgttcagac tacccaagag 1080

gaagatggat gcagttgccg ctttccagaa gaggaggagg gcgggtgcga actgtaa 1137

<210> 310

<211> 1134

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 310

atgctgctgc tggtcacatc tctgctgctg tgcgagctgc cccatcctgc ctttctgctg 60

atccccgaca tccagatgac ccagaccaca agcagcctgt ctgccagcct gggcgataga 120

gtgaccatca gctgtagagc cagccaggac atcagcaagt acctgaactg gtatcagcaa 180

aagcccgacg gcaccgtgaa gctgctgatc taccacacca gcagactgca cagcggcgtg 240

ccaagcagat tttctggcag cggctctggc accgactaca gcctgacaat cagcaacctg 300

gaacaagagg atatcgctac ctacttctgc cagcaaggca acaccctgcc ttacaccttt 360

ggcggaggca ccaagctgga aatcaccggc tctacaagcg gcagcggcaa acctggatct 420

ggcgagggat ctaccaaggg cgaagtgaaa ctgcaagagt ctggccctgg actggtggcc 480

ccatctcagt ctctgagcgt gacctgtaca gtcagcggag tgtccctgcc tgattacggc 540

gtgtcctgga tcagacagcc tcctcggaaa ggcctggaat ggctgggagt gatctggggc 600

agcgagacaa cctactacaa cagcgccctg aagtcccggc tgaccatcat caaggacaac 660

tccaagagcc aggtgttcct gaagatgaac agcctgcaga ccgacgacac cgccatctac 720

tattgcgcca agcactacta ctacggcggc agctacgcca tggattattg gggccagggc 780

accagcgtga ccgtttcttc tgccgccgct atcgaagtga tgtaccctcc tccttacctg 840

gacaacgaga agtccaacgg caccatcatc cacgtgaagg gcaagcacct gtgtccttct 900

ccactgttcc ccggacctag caagcctttc tgggtgctcg ttgttgttgg cggcgtgctg 960

gcctgttaca gcctgctggt taccgtggcc ttcatcatct tttgggtccg aagcaagcgg 1020

agccggctgc tgcactccga ctacatgaac atgaccccta gacggcccgg accaaccaga 1080

aagcactacc agccttacgc tcctcctaga gacttcgccg cctaccggtc ctaa 1134

<210> 311

<211> 1470

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 311

atgctgctgc tggtcacatc tctgctgctg tgcgagctgc cccatcctgc ctttctgctg 60

atccccgaca tccagatgac ccagaccaca agcagcctgt ctgccagcct gggcgataga 120

gtgaccatca gctgtagagc cagccaggac atcagcaagt acctgaactg gtatcagcaa 180

aagcccgacg gcaccgtgaa gctgctgatc taccacacca gcagactgca cagcggcgtg 240

ccaagcagat tttctggcag cggctctggc accgactaca gcctgacaat cagcaacctg 300

gaacaagagg atatcgctac ctacttctgc cagcaaggca acaccctgcc ttacaccttt 360

ggcggaggca ccaagctgga aatcaccggc tctacaagcg gcagcggcaa acctggatct 420

ggcgagggat ctaccaaggg cgaagtgaaa ctgcaagagt ctggccctgg actggtggcc 480

ccatctcagt ctctgagcgt gacctgtaca gtcagcggag tgtccctgcc tgattacggc 540

gtgtcctgga tcagacagcc tcctcggaaa ggcctggaat ggctgggagt gatctggggc 600

agcgagacaa cctactacaa cagcgccctg aagtcccggc tgaccatcat caaggacaac 660

tccaagagcc aggtgttcct gaagatgaac agcctgcaga ccgacgacac cgccatctac 720

tattgcgcca agcactacta ctacggcggc agctacgcca tggattattg gggccagggc 780

accagcgtga ccgtttcttc tgccgccgct atcgaagtga tgtaccctcc tccttacctg 840

gacaacgaga agtccaacgg caccatcatc cacgtgaagg gcaagcacct gtgtccttct 900

ccactgttcc ccggacctag caagcctttc tgggtgctcg ttgttgttgg cggcgtgctg 960

gcctgttaca gcctgctggt taccgtggcc ttcatcatct tttgggtccg aagcaagcgg 1020

agccggctgc tgcactccga ctacatgaac atgaccccta gacggcccgg accaaccaga 1080

aagcactacc agccttacgc tcctcctaga gacttcgccg cctaccggtc cagagtgaag 1140

ttcagcagat ccgccgatgc tcccgcctat cagcagggcc aaaaccagct gtacaacgag 1200

ctgaacctgg ggagaagaga agagtacgac gtgctggaca agcggagagg cagagatcct 1260

gaaatgggcg gcaagcccag acggaagaat cctcaagagg gcctgtataa tgagctgcag 1320

aaagacaaga tggccgaggc ctacagcgag atcggaatga agggcgagcg cagaagaggc 1380

aagggacacg atggactgta ccagggactg agcaccgcca ccaaggatac ctatgacgcc 1440

ctgcacatgc aggccctgcc tccaagataa 1470

<210> 312

<211> 1347

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 312

atgctgctgc tggtcacatc tctgctgctg tgcgagctgc cccatcctgc ctttctgctg 60

atccccgaca tccagatgac ccagaccaca agcagcctgt ctgccagcct gggcgataga 120

gtgaccatca gctgtagagc cagccaggac atcagcaagt acctgaactg gtatcagcaa 180

aagcccgacg gcaccgtgaa gctgctgatc taccacacca gcagactgca cagcggcgtg 240

ccaagcagat tttctggcag cggctctggc accgactaca gcctgacaat cagcaacctg 300

gaacaagagg atatcgctac ctacttctgc cagcaaggca acaccctgcc ttacaccttt 360

ggcggaggca ccaagctgga aatcaccggc tctacaagcg gcagcggcaa acctggatct 420

ggcgagggat ctaccaaggg cgaagtgaaa ctgcaagagt ctggccctgg actggtggcc 480

ccatctcagt ctctgagcgt gacctgtaca gtcagcggag tgtccctgcc tgattacggc 540

gtgtcctgga tcagacagcc tcctcggaaa ggcctggaat ggctgggagt gatctggggc 600

agcgagacaa cctactacaa cagcgccctg aagtcccggc tgaccatcat caaggacaac 660

tccaagagcc aggtgttcct gaagatgaac agcctgcaga ccgacgacac cgccatctac 720

tattgcgcca agcactacta ctacggcggc agctacgcca tggattattg gggccagggc 780

accagcgtga ccgtttcttc tgccgccgct atcgaagtga tgtaccctcc tccttacctg 840

gacaacgaga agtccaacgg caccatcatc cacgtgaagg gcaagcacct gtgtccttct 900

ccactgttcc ccggacctag caagcctttc tgggtgctcg ttgttgttgg cggcgtgctg 960

gcctgttaca gcctgctggt taccgtggcc ttcatcatct tttgggtcag agtgaagttc 1020

agcagatccg ccgatgctcc cgcctatcag cagggccaaa accagctgta caacgagctg 1080

aacctgggga gaagagaaga gtacgacgtg ctggacaagc ggagaggcag agatcctgaa 1140

atgggcggca agcccagacg gaagaatcct caagagggcc tgtataatga gctgcagaaa 1200

gacaagatgg ccgaggccta cagcgagatc ggaatgaagg gcgagcgcag aagaggcaag 1260

ggacacgatg gactgtacca gggactgagc accgccacca aggataccta tgacgccctg 1320

cacatgcagg ccctgcctcc aagataa 1347

<210> 313

<211> 1461

<212> DNA

<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"
<400> 313

atggctctcc cggtcacagc ccttctcctg cccctggcac tcttgctgca tgcggcacga 60

cccgacatcc agatgaccca gaccacaagc agcctgtctg ccagcctggg cgatagagtg 120

accatcagct gtagagccag ccaggacatc agcaagtacc tgaactggta tcagcaaaag 180

cccgacggca ccgtgaagct gctgatctac cacaccagca gactgcacag cggcgtgcca 240

agcagatttt ctggcagcgg ctctggcacc gactacagcc tgacaatcag caacctggaa 300

caagaggata tcgctaccta cttctgccag caaggcaaca ccctgcctta cacctttggc 360

ggaggcacca agctggaaat caccggtgga ggtggttctg gcggaggggg atctggtgga 420

ggcggttcag aagtgaaact gcaagagtct ggccctggac tggtggcccc atctcagtct 480

ctgagcgtga cctgtacagt cagcggagtg tccctgcctg attacggcgt gtcctggatc 540

agacagcctc ctcggaaagg cctggaatgg ctgggagtga tctggggcag cgagacaacc 600

tactacaaca gcgccctgaa gtcccggctg accatcatca aggacaactc caagagccag 660

gtgttcctga agatgaacag cctgcagacc gacgacaccg ccatctacta ttgcgccaag 720

cactactact acggcggcag ctacgccatg gattattggg gccagggcac cagcgtgacc 780

gtttcttcta ccacaacgcc cgccccgcga ccgcctactc ccgctcccac aattgcatca 840

caacccctgt ctttgagacc cgaagcttgt cgaccagctg ccggtggcgc ggttcacacg 900

cgggggctcg atttcgcctg tgatatatat atatgggccc cattggctgg aacatgcgga 960

gtattgcttc tgagcctggt gattaccctc tactgtaaga gaggccggaa gaaacttctt 1020

tatatattca agcagccctt tatgcgaccc gttcagacta cccaagagga agatggatgc 1080

agttgccgct ttccagaaga ggaggagggc gggtgcgaac tgagagtgaa gttcagcaga 1140

tccgccgatg ctcccgccta tcagcagggc caaaaccagc tgtacaacga gctgaacctg 1200

gggagaagag aagagtacga cgtgctggac aagcggagag gcagagatcc tgaaatgggc 1260

ggcaagccca gacggaagaa tcctcaagag ggcctgtata atgagctgca gaaagacaag 1320

atggccgagg cctacagcga gatcggaatg aagggcgagc gcagaagagg caagggacac 1380

gatggactgt accagggact gagcaccgcc accaaggata cctatgacgc cctgcacatg 1440

caggccctgc ctccaagata a 1461

<210> 314

<211> 1461

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 314

atggctctcc cggtcacagc ccttctcctg cccctggcac tcttgctgca tgcggcacga 60

cccgacatcc agatgaccca gaccacaagc agcctgtctg ccagcctggg cgatagagtg 120

accatcagct gtagagccag ccaggacatc agcaagtacc tgaactggta tcagcaaaag 180

cccgacggca ccgtgaagct gctgatctac cacaccagca gactgcacag cggcgtgcca 240

agcagatttt ctggcagcgg ctctggcacc gactacagcc tgacaatcag caacctggaa 300

caagaggata tcgctaccta cttctgccag caaggcaaca ccctgcctta cacctttggc 360

ggaggcacca agctggaaat caccggtgga ggtggttctg gcggaggggg atctggtgga 420

ggcggttcag aagtgaaact gcaagagtct ggccctggac tggtggcccc atctcagtct 480

ctgagcgtga cctgtacagt cagcggagtg tccctgcctg attacggcgt gtcctggatc 540

agacagcctc ctcggaaagg cctggaatgg ctgggagtga tctggggcag cgagacaacc 600

tactacaaca gcgccctgaa gtcccggctg accatcatca aggacaactc caagagccag 660

gtgttcctga agatgaacag cctgcagacc gacgacaccg ccatctacta ttgcgccaag 720

cactactact acggcggcag ctacgccatg gattattggg gccagggcac cagcgtgacc 780

gtttcttcta ccacaacgcc cgccccgcga ccgcctactc ccgctcccac aattgcatca 840

caacccctgt ctttgagacc cgaagcttgt cgaccagctg ccggtggcgc ggttcacacg 900

cgggggctcg atttcgcctg tgatatatat atatgggccc cattggctgg aacatgcgga 960

gtattgcttc tgagcctggt gattaccctc tactgtaaga gaggccggaa gaaacttctt 1020

tatatattca agcagccctt tatgcgaccc gttcagacta cccaagagga agatggatgc 1080

agttgccgct ttccagaaga ggaggagggc gggtgcgaac tgagagtgaa gttcagcaga 1140

tccgccgatg ctcccgccta taagcagggc caaaaccagc tgtacaacga gctgaacctg 1200

gggagaagag aagagtacga cgtgctggac aagcggagag gcagagatcc tgaaatgggc 1260

ggcaagccca gacggaagaa tcctcaagag ggcctgtata atgagctgca gaaagacaag 1320

atggccgagg cctacagcga gatcggaatg aagggcgagc gcagaagagg caagggacac 1380

gatggactgt accagggact gagcaccgcc accaaggata cctatgacgc cctgcacatg 1440

caggccctgc ctccaagata a 1461

<210> 315

<211> 1452

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 315

atgctgctgc tggtcacatc tctgctgctg tgcgagctgc cccatcctgc ctttctgctg 60

atcccccagg ttcaactcca gcagtctggt cccggcctcg ttaaaccaag ccagactttg 120

tctcttacct gtgctatcag tggcgatagc gtgtctagta attcagccgc atggaactgg 180

atccgacaat caccgagtag gggacttgaa tggctgggta gaacctatta ccggtccaaa 240

tggtacaatg actatgcagt gtctgtaaaa agcaggatca cgatcaaccc tgatacgtct 300

aaaaaccagt tttctctgca acttaatagt gtgacccctg aagacaccgc tgtgtattac 360

tgtgcacggg aggttaccgg tgatcttgaa gatgcttttg atatatgggg ccaaggtacg 420

atggtcacgg tgtctagtgg gggaggcggc agcgacatac agatgacgca gagcccatcc 480

agtctctccg cgtctgttgg tgacagagtg actattacat gtagggcgtc tcagaccatt 540

tggtcttacc tcaattggta tcaacagcga ccaggcaaag caccgaactt gctcatttac 600

gctgccagct cactccaaag tggtgtgccg tccagattta gtggtagggg cagtggcact 660

gatttcactc tgactatttc aagtcttcaa gctgaggatt ttgccacata ctactgccag 720

caaagttact caatacctca gacttttgga caggggacaa aattggagat taaatccgga 780

accacaacgc ccgccccgcg accgcctact cccgctccca caattgcatc acaacccctg 840

tctttgagac ccgaagcttg tcgaccagct gccggtggcg cggttcacac gcgggggctc 900

gatttcgcct gtgatatata tatgggccc ccattggctg gaacatgcgg agtattgctt 960

ctgagcctgg tgattaccct ctactgtaag agaggccgga agaaacttct ttatatattc 1020

aagcagccct ttatgcgacc cgttcagact acccaagagg aagatggatg cagttgccgc 1080

tttccagaag aggaggaggg cgggtgcgaa ctgagagtga agttcagcag atccgccgat 1140

gctcccgcct ataagcaggg ccaaaaccag ctgtacaacg agctgaacct ggggagaaga 1200

gaagagtacg acgtgctgga caagcggaga ggcagagatc ctgaaatggg cggcaagccc 1260

agacggaaga atcctcaaga gggcctgtat aatgagctgc agaaagacaa gatggccgag 1320

gcctacagcg agatcggaat gaagggcgag cgcagaagag gcaagggaca cgatggactg 1380

taccagggac tgagcaccgc caccaaggat acctatgacg ccctgcacat gcaggccctg 1440

cctccaagat aa 1452

<210> 316

<211> 2202

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 316

atgctgctgc tggtcacatc tctgctgctg tgcgagctgc cccatcctgc ctttctgctg 60

atccccgaca tccagatgac ccagaccaca agcagcctgt ctgccagcct gggcgataga 120

gtgaccatca gctgtagagc cagccaggac atcagcaagt acctgaactg gtatcagcaa 180

aagcccgacg gcaccgtgaa gctgctgatc taccacacca gcagactgca cagcggcgtg 240

ccaagcagat tttctggcag cggctctggc accgactaca gcctgacaat cagcaacctg 300

gaacaagagg atatcgctac ctacttctgc cagcaaggca acaccctgcc ttacaccttt 360

ggcggaggca ccaagctgga aatcaccggc ggcggaggat cccaggttca actccagcag 420

tctggtcccg gcctcgttaa accaagccag actttgtctc ttacctgtgc tatcagtggc 480

gatagcgtgt ctagtaattc agccgcatgg aactggatcc gacaatcacc gagtagggga 540

cttgaatggc tgggtagaac ctattaccgg tccaaatggt acaatgacta tgcagtgtct 600

gtaaaaagca ggatcacgat caaccctgat acgtctaaaa accagttttc tctgcaactt 660

aatagtgtga cccctgaaga caccgctgtg tattactgtg cacgggaggt taccggtgat 720

cttgaagatg cttttgatat atggggccaa ggtacgatgg tcacggtgtc tagtggctct 780

acaagcggca gcggcaaacc tggatctggc gagggatcta ccaagggcga catacagatg 840

acgcagagcc catccagtct ctccgcgtct gttggtgaca gagtgactat tacatgtagg 900

gcgtctcaga ccatttggtc ttacctcaat tggtatcaac agcgaccagg caaagcaccg 960

aacttgctca tttacgctgc cagctcactc caaagtggtg tgccgtccag atttagtggt 1020

aggggcagtg gcactgattt cactctgact atttcaagtc ttcaagctga ggattttgcc 1080

acatactact gccagcaaag ttactcaata cctcagactt ttggacaggg gacaaaattg 1140

gagattaaag ggggaggcgg cagcgaagtg aaactgcaag agtctggccc tggactggtg 1200

gccccatctc agtctctgag cgtgacctgt acagtcagcg gagtgtccct gcctgattac 1260

ggcgtgtcct ggatcagaca gcctcctcgg aaaggcctgg aatggctggg agtgatctgg 1320

ggcagcgaga caacctacta caacagcgcc ctgaagtccc ggctgaccat catcaaggac 1380

aactccaaga gccaggtgtt cctgaagatg aacagcctgc agaccgacga caccgccatc 1440

tactattgcg ccaagcacta ctactacggc ggcagctacg ccatggatta ttggggccag 1500

ggcaccagcg tgaccgtttc ttcttccgga accacaacgc ccgccccgcg accgcctact 1560

cccgctccca caattgcatc acaacccctg tctttgagac ccgaagcttg tcgaccagct 1620

gccggtggcg cggttcacac gcgggggctc gatttcgcct gtgatatata tatatgggcc 1680

ccattggctg gaacatgcgg agtattgctt ctgagcctgg tgattaccct ctactgtaag 1740

agaggccgga agaaacttct ttatatattc aagcagccct ttatgcgacc cgttcagact 1800

acccaagagg aagatggatg cagttgccgc tttccagaag aggaggaggg cgggtgcgaa 1860

ctgagagtga agttcagcag atccgccgat gctcccgcct ataagcaggg ccaaaaccag 1920

ctgtacaacg agctgaacct ggggagaaga gaagagtacg acgtgctgga caagcggaga 1980

ggcagagatc ctgaaatggg cggcaagccc agacggaaga atcctcaaga gggcctgtat 2040

aatgagctgc agaaagacaa gatggccgag gcctacagcg agatcggaat gaagggcgag 2100

cgcagaagag gcaagggaca cgatggactg taccagggac tgagcaccgc caccaaggat 2160

acctatgacg ccctgcacat gcaggccctg cctccaagat aa 2202

<210> 317

<211> 2262

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 317

atgctgctgc tggtcacatc tctgctgctg tgcgagctgc cccatcctgc ctttctgctg 60

atcccccagg ttcaactcca gcagtctggt cccggcctcg ttaaaccaag ccagactttg 120

tctcttacct gtgctatcag tggcgatagc gtgtctagta attcagccgc atggaactgg 180

atccgacaat caccgagtag gggacttgaa tggctgggta gaacctatta ccggtccaaa 240

tggtacaatg actatgcagt gtctgtaaaa agcaggatca cgatcaaccc tgatacgtct 300

aaaaaccagt tttctctgca acttaatagt gtgacccctg aagacaccgc tgtgtattac 360

tgtgcacggg aggttaccgg tgatcttgaa gatgcttttg atatatgggg ccaaggtacg 420

atggtcacgg tgtctagtgg gggaggcggc agcgacatac agatgacgca gagcccatcc 480

agtctctccg cgtctgttgg tgacagagtg actattacat gtagggcgtc tcagaccatt 540

tggtcttacc tcaattggta tcaacagcga ccaggcaaag caccgaactt gctcatttac 600

gctgccagct cactccaaag tggtgtgccg tccagattta gtggtagggg cagtggcact 660

gatttcactc tgactatttc aagtcttcaa gctgaggatt ttgccacata ctactgccag 720

caaagttact caatacctca gacttttgga caggggacaa aattggagat taaaggggga 780

ggcggatccg gcggtggtgg ctccggcggt ggtggttctg gaggcggcgg aagcggtggg 840

ggtggtagcg acatccagat gacccagacc acaagcagcc tgtctgccag cctgggcgat 900

agagtgacca tcagctgtag agccagccag gacatcagca agtacctgaa ctggtatcag 960

caaaagcccg acggcaccgt gaagctgctg atctaccaca ccagcagact gcacagcggc 1020

gtgccaagca gattttctgg cagcggctct ggcaccgact acagcctgac aatcagcaac 1080

ctggaacaag aggatatcgc tacctacttc tgccagcaag gcaacaccct gccttacacc 1140

tttggcggag gcaccaagct ggaaatcacc ggctctacaa gcggcagcgg caaacctgga 1200

tctggcgagg gatctaccaa gggcgaagtg aaactgcaag agtctggccc tggactggtg 1260

gccccatctc agtctctgag cgtgacctgt acagtcagcg gagtgtccct gcctgattac 1320

ggcgtgtcct ggatcagaca gcctcctcgg aaaggcctgg aatggctggg agtgatctgg 1380

ggcagcgaga caacctacta caacagcgcc ctgaagtccc ggctgaccat catcaaggac 1440

aactccaaga gccaggtgtt cctgaagatg aacagcctgc agaccgacga caccgccatc 1500

tactattgcg ccaagcacta ctactacggc ggcagctacg ccatggatta ttggggccag 1560

ggcaccagcg tgaccgtttc ttcttccgga accacaacgc ccgccccgcg accgcctact 1620

cccgctccca caattgcatc acaacccctg tctttgagac ccgaagcttg tcgaccagct 1680

gccggtggcg cggttcacac gcgggggctc gatttcgcct gtgatatata tatatgggcc 1740

ccattggctg gaacatgcgg agtattgctt ctgagcctgg tgattaccct ctactgtaag 1800

agaggccgga agaaacttct ttatatattc aagcagccct ttatgcgacc cgttcagact 1860

acccaagagg aagatggatg cagttgccgc tttccagaag aggaggaggg cgggtgcgaa 1920

ctgagagtga agttcagcag atccgccgat gctcccgcct ataagcaggg ccaaaaccag 1980

ctgtacaacg agctgaacct ggggagaaga gaagagtacg acgtgctgga caagcggaga 2040

ggcagagatc ctgaaatggg cggcaagccc agacggaaga atcctcaaga gggcctgtat 2100

aatgagctgc agaaagacaa gatggccgag gcctacagcg agatcggaat gaagggcgag 2160

cgcagaagag gcaagggaca cgatggactg taccagggac tgagcaccgc caccaaggat 2220

acctatgacg ccctgcacat gcaggccctg cctccaagat aa 2262

<210> 318

<211> 489

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 318

Met Leu Leu Leu Val Thr Ser Leu Leu Leu Cys Glu Leu Pro His Pro
1 5 10 15

Ala Phe Leu Leu Ile Pro Asp Ile Gln Met Thr Gln Thr Thr Ser Ser
20 25 30

Leu Ser Ala Ser Leu Gly Asp Arg Val Thr Ile Ser Cys Arg Ala Ser
35 40 45

Gln Asp Ile Ser Lys Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Asp Gly
50 55 60

Thr Val Lys Leu Leu Ile Tyr His Thr Ser Arg Leu His Ser Gly Val
65 70 75 80

Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr
85 90 95

Ile Ser Asn Leu Glu Gln Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln
100 105 110

Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
115 120 125

Thr Gly Ser Thr Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser
130 135 140

Thr Lys Gly Glu Val Lys Leu Gln Glu Ser Gly Pro Gly Leu Val Ala
145 150 155 160

Pro Ser Gln Ser Leu Ser Val Thr Cys Thr Val Ser Gly Val Ser Leu
165 170 175

Pro Asp Tyr Gly Val Ser Trp Ile Arg Gln Pro Pro Arg Lys Gly Leu

180 185 190

Glu Trp Leu Gly Val Ile Trp Gly Ser Glu Thr Thr Tyr Tyr Asn Ser
195 200 205
Ala Leu Lys Ser Arg Leu Thr Ile Ile Lys Asp Asn Ser Lys Ser Gln
210 215 220
Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Ile Tyr
225 230 235 240
Tyr Cys Ala Lys His Tyr Tyr Tyr Gly Gly Ser Tyr Ala Met Asp Tyr
245 250 255

Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Ala Ala Ile Glu
260 265 270
Val Met Tyr Pro Pro Pro Tyr Leu Asp Asn Glu Lys Ser Asn Gly Thr
275 280 285
Ile Ile His Val Lys Gly Lys His Leu Cys Pro Ser Pro Leu Phe Pro
290 295 300
Gly Pro Ser Lys Pro Phe Trp Val Leu Val Val Val Gly Gly Val Leu
305 310 315 320

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

Ala Asp Ala Pro Ala Tyr Gln Gln Gly Gln Asn Gln Leu Tyr Asn Glu
385 390 395 400
Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg Arg
405 410 415
Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg Arg Lys Asn Pro Gln
420 425 430

Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met Ala Glu Ala Tyr
435 440 445

Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg Gly Lys Gly His Asp
450 455 460
Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp Ala
465 470 475 480
Leu His Met Gln Ala Leu Pro Pro Arg
485

<210> 319
<211> 486
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400>
319
Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
1 5 10 15
His Ala Ala Arg Pro Asp Ile Gln Met Thr Gln Thr Thr Ser Ser Leu
20 25 30
Ser Ala Ser Leu Gly Asp Arg Val Thr Ile Ser Cys Arg Ala Ser Gln
35 40 45
Asp Ile Ser Lys Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Asp Gly Thr
50 55 60

Val Lys Leu Leu Ile Tyr His Thr Ser Arg Leu His Ser Gly Val Pro
65 70 75 80
Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser Leu Thr Ile
85 90 95
Ser Asn Leu Glu Gln Glu Asp Ile Ala Thr Tyr Phe Cys Gln Gln Gly
100 105 110
Asn Thr Leu Pro Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Thr

115 120 125

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu
130 135 140
Val Lys Leu Gln Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln Ser
145 150 155 160
Leu Ser Val Thr Cys Thr Val Ser Gly Val Ser Leu Pro Asp Tyr Gly
165 170 175
Val Ser Trp Ile Arg Gln Pro Pro Arg Lys Gly Leu Glu Trp Leu Gly
180 185 190

Val Ile Trp Gly Ser Glu Thr Thr Tyr Tyr Asn Ser Ala Leu Lys Ser
195 200 205
Arg Leu Thr Ile Ile Lys Asp Asn Ser Lys Ser Gln Val Phe Leu Lys
210 215 220
Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Ile Tyr Tyr Cys Ala Lys
225 230 235 240
His Tyr Tyr Tyr Gly Gly Ser Tyr Ala Met Asp Tyr Trp Gly Gln Gly
245 250 255

Thr Ser Val Thr Val Ser Ser Thr Thr Thr Pro Ala Pro Arg Pro Pro
260 265 270
Thr Pro Ala Pro Thr Ile Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu
275 280 285
Ala Cys Arg Pro Ala Ala Gly Gly Ala Val His Thr Arg Gly Leu Asp
290 295 300
Phe Ala Cys Asp Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly
305 310 315 320

Val Leu Leu Leu Ser Leu Val Ile Thr Leu Tyr Cys Lys Arg Gly Arg
325 330 335
Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met Arg Pro Val Gln
340 345 350
Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe Pro Glu Glu Glu
355 360 365

Glu Gly Gly Cys Glu Leu Arg Val Lys Phe Ser Arg Ser Ala Asp Ala
370 375 380

Pro Ala Tyr Lys Gln Gly Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu
385 390 395 400
Gly Arg Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg Arg Gly Arg Asp
405 410 415
Pro Glu Met Gly Gly Lys Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu
420 425 430
Tyr Asn Glu Leu Gln Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile
435 440 445

Gly Met Lys Gly Glu Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr
450 455 460
Gln Gly Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Met
465 470 475 480
Gln Ala Leu Pro Pro Arg
485

<210> 320
<211> 493
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 320
Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

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

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

65 70 75 80

Thr Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe

85 90 95

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

100 105 110

Cys Leu Gln Ser Arg Ile Phe Pro Arg Thr Phe Gly Gln Gly Thr Lys

115 120 125

Leu Glu Ile Lys Gly Ser Thr Ser Gly Ser Gly Lys Pro Gly Ser Gly

130 135 140

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

145 150 155 160

Leu Lys Lys Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly

165 170 175

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

180 185 190

Gln Gly Leu Glu Trp Met Gly Trp Ile Asn Thr Glu Thr Arg Glu Pro

195 200 205

Ala Tyr Ala Tyr Asp Phe Arg Gly Arg Phe Val Phe Ser Leu Asp Thr

210 215 220

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

225 230 235 240

Thr Ala Val Tyr Tyr Cys Ala Arg Asp Tyr Ser Tyr Ala Met Asp Tyr

245 250 255

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ala Ala Thr Thr

260 265 270

Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln

275 280 285

Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala

290 295 300

Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp Ala

305 310 315 320

Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr

325 330 335

Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln

340 345 350

Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser

355 360 365

Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg Val Lys

370 375 380

Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Gln Gln Gly Gln Asn Gln

385 390 395 400

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

405 410 415

Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg Arg

420 425 430

Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met

435 440 445

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

450 455 460

Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp

465 470 475 480

Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg

485 490

<210> 321

<211> 475

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 321

```
Met Gly Val Pro Thr Gln Leu Leu Gly Leu Leu Leu Leu Trp Ile Thr

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

65                  70                  75                  80
Arg Phe Ser Gly Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys Ile Thr
                85                  90                  95
Ser Met Gln Thr Glu Asp Glu Gly Val Tyr Phe Cys Gln Gln Gly Leu
            100                 105                 110
Thr Tyr Pro Arg Thr Phe Gly Gly Gly Thr Lys Leu Glu Leu Lys Gly
        115                 120                 125
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Val

    130                 135                 140
Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Thr Ser Val
145                 150                 155                 160
Lys Leu Ser Cys Lys Val Ser Gly Asp Thr Ile Thr Phe Tyr Tyr Met
                165                 170                 175
His Phe Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly Arg
            180                 185                 190
Ile Asp Pro Glu Asp Glu Ser Thr Lys Tyr Ser Glu Lys Phe Lys Asn

        195                 200                 205
Lys Ala Thr Leu Thr Ala Asp Thr Ser Ser Asn Thr Ala Tyr Leu Lys
    210                 215                 220
Leu Ser Ser Leu Thr Ser Glu Asp Thr Ala Thr Tyr Phe Cys Ile Tyr
225                 230                 235                 240
Gly Gly Tyr Tyr Phe Asp Tyr Trp Gly Gln Gly Val Met Val Thr Val
```

245 250 255

Ser Ser Ile Glu Phe Met Tyr Pro Pro Pro Tyr Leu Asp Asn Glu Arg

260 265 270

Ser Asn Gly Thr Ile Ile His Ile Lys Glu Lys His Leu Cys His Thr

275 280 285

Gln Ser Ser Pro Lys Leu Phe Trp Ala Leu Val Val Val Ala Gly Val

290 295 300

Leu Phe Cys Tyr Gly Leu Leu Val Thr Val Ala Leu Cys Val Ile Trp

305 310 315 320

Thr Asn Ser Arg Arg Asn Arg Gly Gly Gln Ser Asp Tyr Met Asn Met

325 330 335

Thr Pro Arg Arg Pro Gly Leu Thr Arg Lys Pro Tyr Gln Pro Tyr Ala

340 345 350

Pro Ala Arg Asp Phe Ala Ala Tyr Arg Pro Arg Ala Lys Phe Ser Arg

355 360 365

Ser Ala Glu Thr Ala Ala Asn Leu Gln Asp Pro Asn Gln Leu Phe Asn

370 375 380

Glu Leu Asn Leu Gly Arg Arg Glu Glu Phe Asp Val Leu Glu Lys Lys

385 390 395 400

Arg Ala Arg Asp Pro Glu Met Gly Gly Lys Gln Gln Arg Arg Arg Asn

405 410 415

Pro Gln Glu Gly Val Tyr Asn Ala Leu Gln Lys Asp Lys Met Ala Glu

420 425 430

Ala Tyr Ser Glu Ile Gly Thr Lys Gly Glu Arg Arg Arg Gly Lys Gly

435 440 445

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

450 455 460

Asp Ala Leu His Met Gln Thr Leu Ala Pro Arg

465 470 475

<210> 322

<211> 475

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 322

Met Gly Val Pro Thr Gln Leu Leu Gly Leu Leu Leu Leu Trp Ile Thr
1 5 10 15

Asp Ala Ile Cys Asp Ile Gln Met Thr Gln Ser Pro Ala Ser Leu Ser
20 25 30

Thr Ser Leu Gly Glu Thr Val Thr Ile Gln Cys Gln Ala Ser Glu Asp
35 40 45

Ile Tyr Ser Gly Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ser Pro
50 55 60

Gln Leu Leu Ile Tyr Gly Ala Ser Asp Leu Gln Asp Gly Val Pro Ser
65 70 75 80

Arg Phe Ser Gly Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys Ile Thr
85 90 95

Ser Met Gln Thr Glu Asp Glu Gly Val Tyr Phe Cys Gln Gln Gly Leu
100 105 110

Thr Tyr Pro Arg Thr Phe Gly Gly Gly Thr Lys Leu Glu Leu Lys Gly
115 120 125

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Val
130 135 140

Gln Leu Gln Gln Ser Gly Ala Glu Leu Val Arg Pro Gly Thr Ser Val
145 150 155 160

Lys Leu Ser Cys Lys Val Ser Gly Asp Thr Ile Thr Phe Tyr Tyr Met
165 170 175

His Phe Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly Arg
180 185 190

Ile Asp Pro Glu Asp Glu Ser Thr Lys Tyr Ser Glu Lys Phe Lys Asn
195 200 205

Lys Ala Thr Leu Thr Ala Asp Thr Ser Ser Asn Thr Ala Tyr Leu Lys

210 215 220

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

225 230 235 240

Gly Gly Tyr Tyr Phe Asp Tyr Trp Gly Gln Gly Val Met Val Thr Val

245 250 255

Ser Ser Ile Glu Phe Met Tyr Pro Pro Pro Tyr Leu Asp Asn Glu Arg

260 265 270

Ser Asn Gly Thr Ile Ile His Ile Lys Glu Lys His Leu Cys His Thr

275 280 285

Gln Ser Ser Pro Lys Leu Phe Trp Ala Leu Val Val Val Ala Gly Val

290 295 300

Leu Phe Cys Tyr Gly Leu Leu Val Thr Val Ala Leu Cys Val Ile Trp

305 310 315 320

Thr Asn Ser Arg Arg Asn Arg Gly Gly Gln Ser Asp Tyr Met Asn Met

325 330 335

Thr Pro Arg Arg Pro Gly Leu Thr Arg Lys Pro Tyr Gln Pro Tyr Ala

340 345 350

Pro Ala Arg Asp Phe Ala Ala Tyr Arg Pro Arg Ala Lys Phe Ser Arg

355 360 365

Ser Ala Glu Thr Ala Ala Asn Leu Gln Asp Pro Asn Gln Leu Tyr Asn

370 375 380

Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu Glu Lys Lys

385 390 395 400

Arg Ala Arg Asp Pro Glu Met Gly Gly Lys Gln Gln Arg Arg Arg Asn

405 410 415

Pro Gln Glu Gly Val Tyr Asn Ala Leu Gln Lys Asp Lys Met Ala Glu

420 425 430

Ala Tyr Ser Glu Ile Gly Thr Lys Gly Glu Arg Arg Arg Gly Lys Gly

435 440 445

His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp Thr Tyr

450 455 460

Asp Ala Leu His Met Gln Thr Leu Ala Pro Arg

465 470 475

<210> 323

<211> 447

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 323

Gln Val Gln Leu Val Gln Ser Gly Val Glu Val Lys Lys Pro Gly Ala

1 5 10 15

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

20 25 30

Tyr Met Tyr Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Gly Ile Asn Pro Ser Asn Gly Gly Thr Asn Phe Asn Glu Lys Phe

50 55 60

Lys Asn Arg Val Thr Leu Thr Thr Asp Ser Ser Thr Thr Thr Ala Tyr

65 70 75 80

Met Glu Leu Lys Ser Leu Gln Phe Asp Asp Thr Ala Val Tyr Tyr Cys

85 90 95

Ala Arg Arg Asp Tyr Arg Phe Asp Met Gly Phe Asp Tyr Trp Gly Gln

100 105 110

Gly Thr Thr Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val

115 120 125

Phe Pro Leu Ala Pro Cys Ser Arg Ser Thr Ser Glu Ser Thr Ala Ala

130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser

145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val

165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro
180 185 190
Ser Ser Ser Leu Gly Thr Lys Thr Tyr Thr Cys Asn Val Asp His Lys
195 200 205
Pro Ser Asn Thr Lys Val Asp Lys Arg Val Glu Ser Lys Tyr Gly Pro
210 215 220
Pro Cys Pro Pro Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val
225 230 235 240

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
245 250 255
Pro Glu Val Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro Glu
260 265 270
Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
275 280 285
Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val Val Ser
290 295 300

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
305 310 315 320
Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile
325 330 335
Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
340 345 350
Pro Ser Gln Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
355 360 365

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
370 375 380
Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
385 390 395 400
Asp Gly Ser Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg
405 410 415

Trp Gln Glu Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
420 425 430

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly Lys
435 440 445

<210> 324
<211> 9
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic peptide"
<220><221> VARIANT
<222> (5)..(5)
<223> /replace="Ser"
<220><221> SITE
<222> (1)..(9)
<223> /note="Variant residues given in the sequence have no preference with respect to those in the annotations for variant positions"

<400> 324
Gly Asp Val Glu Glu Asn Pro Gly Pro
1 5

<210> 325
<211> 440
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 325
Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
1 5 10 15
Ser Leu Arg Leu Asp Cys Lys Ala Ser Gly Ile Thr Phe Ser Asn Ser

20 25 30

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

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

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

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

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
275 280 285

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

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
355 360 365
Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
370 375 380
Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
385 390 395 400
Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe
405 410 415

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
420 425 430
Ser Leu Ser Leu Ser Leu Gly Lys
435 440

<210> 326
<211> 214
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 326
Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Tyr

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile

35 40 45

Tyr Asp Ala Ser Asn Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly

50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu Pro

65 70 75 80

Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Ser Ser Asn Trp Pro Arg

85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala

100 105 110

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

115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala

130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln

145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser

165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr

180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser

195 200 205

Phe Asn Arg Gly Glu Cys

210

<210> 327

<211> 448

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 327

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Asp Ser
20 25 30

Trp Ile His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Trp Ile Ser Pro Tyr Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Arg His Trp Pro Gly Gly Phe Asp Tyr Trp Gly Gln Gly Thr
100 105 110

Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro
115 120 125

Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
130 135 140

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn
145 150 155 160

Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln
165 170 175

Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser
180 185 190

Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser
195 200 205

Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr
210 215 220

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser

225 230 235 240

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
245 250 255
Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
260 265 270
Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
275 280 285
Lys Thr Lys Pro Arg Glu Glu Gln Tyr Ala Ser Thr Tyr Arg Val Val
290 295 300

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
305 310 315 320
Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
325 330 335
Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
340 345 350
Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
355 360 365

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
370 375 380
Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
385 390 395 400
Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
405 410 415
Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
420 425 430

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
435 440 445

<210> 328

<211> 214

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 328

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Val Ser Thr Ala
20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Leu Tyr His Pro Ala
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
195 200 205

Phe Asn Arg Gly Glu Cys
210

<210> 329

<211> 450

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 329

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
20 25 30

Ile Met Met Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ser Ile Tyr Pro Ser Gly Gly Ile Thr Phe Tyr Ala Asp Thr Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Ile Lys Leu Gly Thr Val Thr Thr Val Asp Tyr Trp Gly Gln
100 105 110

Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val
115 120 125

Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala
130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser
145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro

180 185 190

Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys

195 200 205

Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp

210 215 220

Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly

225 230 235 240

Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile

245 250 255

Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu

260 265 270

Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His

275 280 285

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg

290 295 300

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys

305 310 315 320

Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu

325 330 335

Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr

340 345 350

Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu

355 360 365

Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp

370 375 380

Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val

385 390 395 400

Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp

405 410 415

Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His

420 425 430

Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro
435 440 445

Gly Lys
450

<210> 330

<211> 216

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 330

Gln Ser Ala Leu Thr Gln Pro Ala Ser Val Ser Gly Ser Pro Gly Gln
1 5 10 15

Ser Ile Thr Ile Ser Cys Thr Gly Thr Ser Ser Asp Val Gly Gly Tyr
20 25 30

Asn Tyr Val Ser Trp Tyr Gln Gln His Pro Gly Lys Ala Pro Lys Leu
35 40 45

Met Ile Tyr Asp Val Ser Asn Arg Pro Ser Gly Val Ser Asn Arg Phe
50 55 60

Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu
65 70 75 80

Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Ser Ser Tyr Thr Ser Ser
85 90 95

Ser Thr Arg Val Phe Gly Thr Gly Thr Lys Val Thr Val Leu Gly Gln
100 105 110

Pro Lys Ala Asn Pro Thr Val Thr Leu Phe Pro Pro Ser Ser Glu Glu
115 120 125

Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu Ile Ser Asp Phe Tyr
130 135 140

Pro Gly Ala Val Thr Val Ala Trp Lys Ala Asp Gly Ser Pro Val Lys

145 150 155 160

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

<210> 331
<211> 451
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 331
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Arg Tyr
20 25 30
Trp Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Asn Ile Lys Gln Asp Gly Ser Glu Lys Tyr Tyr Val Asp Ser Val
50 55 60
Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
65 70 75 80
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Ala Arg Glu Gly Gly Trp Phe Gly Glu Leu Ala Phe Asp Tyr Trp Gly
100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
115 120 125

```
Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
    130                 135                 140
Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
145                 150                 155                 160
Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
                165                 170                 175

Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
            180                 185                 190
Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His
        195                 200                 205
Lys Pro Ser Asn Thr Lys Val Asp Lys Arg Val Glu Pro Lys Ser Cys
    210                 215                 220
Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Phe Glu Gly
225                 230                 235                 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
                245                 250                 255
Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
            260                 265                 270
Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
        275                 280                 285
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
    290                 295                 300

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
305                 310                 315                 320
Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Ser Ile
                325                 330                 335
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
            340                 345                 350
Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser
        355                 360                 365

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
```

370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro

385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val

405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met

420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser

435 440 445

Pro Gly Lys

450

<210> 332

<211> 215

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 332

Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly

1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Arg Val Ser Ser Ser

20 25 30

Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu

35 40 45

Ile Tyr Asp Ala Ser Ser Arg Ala Thr Gly Ile Pro Asp Arg Phe Ser

50 55 60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu

65 70 75 80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Leu Pro

85 90 95

Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala

100 105 110

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

115 120 125

Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu

130 135 140

Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser

145 150 155 160

Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu

165 170 175

Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val

180 185 190

Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys

195 200 205

Ser Phe Asn Arg Gly Glu Cys

210 215

<210> 333

<211> 133

<212> PRT

<213> Homo sapiens

<400> 333

Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu Gln Leu Glu His

1 5 10 15

Leu Leu Leu Asp Leu Gln Met Ile Leu Asn Gly Ile Asn Asn Tyr Lys

20 25 30

Asn Pro Lys Leu Thr Arg Met Leu Thr Phe Lys Phe Tyr Met Pro Lys

35 40 45

Lys Ala Thr Glu Leu Lys His Leu Gln Cys Leu Glu Glu Glu Leu Lys

50 55 60

Pro Leu Glu Glu Val Leu Asn Leu Ala Gln Ser Lys Asn Phe His Leu

65 70 75 80

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

85 90 95

Lys Gly Ser Glu Thr Thr Phe Met Cys Glu Tyr Ala Asp Glu Thr Ala

100 105 110

Thr Ile Val Glu Phe Leu Asn Arg Trp Ile Thr Phe Cys Gln Ser Ile

115 120 125

Ile Ser Thr Leu Thr

130

<210> 334

<211> 197

<212> PRT

<213> Homo sapiens

<400> 334

Arg Asn Leu Pro Val Ala Thr Pro Asp Pro Gly Met Phe Pro Cys Leu

1 5 10 15

His His Ser Gln Asn Leu Leu Arg Ala Val Ser Asn Met Leu Gln Lys

20 25 30

Ala Arg Gln Thr Leu Glu Phe Tyr Pro Cys Thr Ser Glu Glu Ile Asp

35 40 45

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

50 55 60

Pro Leu Glu Leu Thr Lys Asn Glu Ser Cys Leu Asn Ser Arg Glu Thr

65 70 75 80

Ser Phe Ile Thr Asn Gly Ser Cys Leu Ala Ser Arg Lys Thr Ser Phe

85 90 95

Met Met Ala Leu Cys Leu Ser Ser Ile Tyr Glu Asp Leu Lys Met Tyr

100 105 110

Gln Val Glu Phe Lys Thr Met Asn Ala Lys Leu Leu Met Asp Pro Lys

115 120 125

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

130 135 140

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

145 150 155 160

Leu Glu Glu Pro Asp Phe Tyr Lys Thr Lys Ile Lys Leu Cys Ile Leu

165 170 175

Leu His Ala Phe Arg Ile Arg Ala Val Thr Ile Asp Arg Val Met Ser

180 185 190

Tyr Leu Asn Ala Ser

195

<210> 335

<211> 306

<212> PRT

<213> Homo sapiens

<400> 335

Ile Trp Glu Leu Lys Lys Asp Val Tyr Val Val Glu Leu Asp Trp Tyr

1 5 10 15

Pro Asp Ala Pro Gly Glu Met Val Val Leu Thr Cys Asp Thr Pro Glu

20 25 30

Glu Asp Gly Ile Thr Trp Thr Leu Asp Gln Ser Ser Glu Val Leu Gly

35 40 45

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

50 55 60

Gln Tyr Thr Cys His Lys Gly Gly Glu Val Leu Ser His Ser Leu Leu

65 70 75 80

Leu Leu His Lys Lys Glu Asp Gly Ile Trp Ser Thr Asp Ile Leu Lys

85 90 95

Asp Gln Lys Glu Pro Lys Asn Lys Thr Phe Leu Arg Cys Glu Ala Lys

100 105 110

Asn Tyr Ser Gly Arg Phe Thr Cys Trp Trp Leu Thr Thr Ile Ser Thr

115 120 125

Asp Leu Thr Phe Ser Val Lys Ser Ser Arg Gly Ser Ser Asp Pro Gln

130 135 140

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

145 150 155 160

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

210 215 220
Lys Asn Ser Arg Gln Val Glu Val Ser Trp Glu Tyr Pro Asp Thr Trp
225 230 235 240
Ser Thr Pro His Ser Tyr Phe Ser Leu Thr Phe Cys Val Gln Val Gln
245 250 255
Gly Lys Ser Lys Arg Glu Lys Lys Asp Arg Val Phe Thr Asp Lys Thr
260 265 270
Ser Ala Thr Val Ile Cys Arg Lys Asn Ala Ser Ile Ser Val Arg Ala

275 280 285
Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser Val Pro
290 295 300
Cys Ser
305

<210> 336
<211> 152
<212> PRT
<213> Homo sapiens
<400> 336
Asp Cys Asp Ile Glu Gly Lys Asp Gly Lys Gln Tyr Glu Ser Val Leu
1 5 10 15
Met Val Ser Ile Asp Gln Leu Leu Asp Ser Met Lys Glu Ile Gly Ser
20 25 30

Asn Cys Leu Asn Asn Glu Phe Asn Phe Phe Lys Arg His Ile Cys Asp
35 40 45
Ala Asn Lys Glu Gly Met Phe Leu Phe Arg Ala Ala Arg Lys Leu Arg

50 55 60

Gln Phe Leu Lys Met Asn Ser Thr Gly Asp Phe Asp Leu His Leu Leu
65 70 75 80

Lys Val Ser Glu Gly Thr Thr Ile Leu Leu Asn Cys Thr Gly Gln Val
85 90 95

Lys Gly Arg Lys Pro Ala Ala Leu Gly Glu Ala Gln Pro Thr Lys Ser
100 105 110

Leu Glu Glu Asn Lys Ser Leu Lys Glu Gln Lys Lys Leu Asn Asp Leu
115 120 125

Cys Phe Leu Lys Arg Leu Leu Gln Glu Ile Lys Thr Cys Trp Asn Lys
130 135 140

Ile Leu Met Gly Thr Lys Glu His
145 150

<210> 337

<211> 160

<212> PRT

<213> Homo sapiens

<400> 337

Ser Pro Gly Gln Gly Thr Gln Ser Glu Asn Ser Cys Thr His Phe Pro
1 5 10 15

Gly Asn Leu Pro Asn Met Leu Arg Asp Leu Arg Asp Ala Phe Ser Arg
20 25 30

Val Lys Thr Phe Phe Gln Met Lys Asp Gln Leu Asp Asn Leu Leu Leu
35 40 45

Lys Glu Ser Leu Leu Glu Asp Phe Lys Gly Tyr Leu Gly Cys Gln Ala
50 55 60

Leu Ser Glu Met Ile Gln Phe Tyr Leu Glu Glu Val Met Pro Gln Ala
65 70 75 80

Glu Asn Gln Asp Pro Asp Ile Lys Ala His Val Asn Ser Leu Gly Glu
85 90 95

Asn Leu Lys Thr Leu Arg Leu Arg Leu Arg Arg Cys His Arg Phe Leu
100 105 110

Pro Cys Glu Asn Lys Ser Lys Ala Val Glu Gln Val Lys Asn Ala Phe
115 120 125

Asn Lys Leu Gln Glu Lys Gly Ile Tyr Lys Ala Met Ser Glu Phe Asp
130 135 140
Ile Phe Ile Asn Tyr Ile Glu Ala Tyr Met Thr Met Lys Ile Arg Asn
145 150 155 160
<210> 338
<211> 114
<212> PRT
<213> Homo sapiens
<400> 338
Asn Trp Val Asn Val Ile Ser Asp Leu Lys Lys Ile Glu Asp Leu Ile
1 5 10 15
Gln Ser Met His Ile Asp Ala Thr Leu Tyr Thr Glu Ser Asp Val His

20 25 30
Pro Ser Cys Lys Val Thr Ala Met Lys Cys Phe Leu Leu Glu Leu Gln
35 40 45
Val Ile Ser Leu Glu Ser Gly Asp Ala Ser Ile His Asp Thr Val Glu
50 55 60
Asn Leu Ile Ile Leu Ala Asn Asn Ser Leu Ser Ser Asn Gly Asn Val
65 70 75 80
Thr Glu Ser Gly Cys Lys Glu Cys Glu Glu Leu Glu Glu Lys Asn Ile

85 90 95
Lys Glu Phe Leu Gln Ser Phe Val His Ile Val Gln Met Phe Ile Asn
100 105 110
Thr Ser

<210> 339
<211> 157
<212> PRT
<213> Homo sapiens
<400> 339

Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn
1 5 10 15
Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp
20 25 30

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

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

<210>
340
<211> 209
<212> PRT
<213> Homo sapiens
<400> 340
Arg Lys Gly Pro Pro Ala Ala Leu Thr Leu Pro Arg Val Gln Cys Arg
1 5 10 15
Ala Ser Arg Tyr Pro Ile Ala Val Asp Cys Ser Trp Thr Leu Pro Pro
20 25 30
Ala Pro Asn Ser Thr Ser Pro Val Ser Phe Ile Ala Thr Tyr Arg Leu
35 40 45

Gly Met Ala Ala Arg Gly His Ser Trp Pro Cys Leu Gln Gln Thr Pro

50 55 60

Thr Ser Thr Ser Cys Thr Ile Thr Asp Val Gln Leu Phe Ser Met Ala
65 70 75 80
Pro Tyr Val Leu Asn Val Thr Ala Val His Pro Trp Gly Ser Ser Ser
85 90 95
Ser Phe Val Pro Phe Ile Thr Glu His Ile Ile Lys Pro Asp Pro Pro
100 105 110
Glu Gly Val Arg Leu Ser Pro Leu Ala Glu Arg Gln Leu Gln Val Gln

115 120 125

Trp Glu Pro Pro Gly Ser Trp Pro Phe Pro Glu Ile Phe Ser Leu Lys
130 135 140
Tyr Trp Ile Arg Tyr Lys Arg Gln Gly Ala Ala Arg Phe His Arg Val
145 150 155 160
Gly Pro Ile Glu Ala Thr Ser Phe Ile Leu Arg Ala Val Arg Pro Arg
165 170 175
Ala Arg Tyr Tyr Val Gln Val Ala Ala Gln Asp Leu Thr Asp Tyr Gly

180 185 190

Glu Leu Ser Asp Trp Ser Leu Pro Ala Thr Ala Thr Met Ser Leu Gly
195 200 205
Lys

<210> 341
<211> 138
<212> PRT
<213> Homo sapiens
<400> 341
Gln Asp Pro Tyr Val Lys Glu Ala Glu Asn Leu Lys Lys Tyr Phe Asn
1 5 10 15
Ala Gly His Ser Asp Val Ala Asp Asn Gly Thr Leu Phe Leu Gly Ile
20 25 30

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

Asn Tyr Ser Val Thr Asp Leu Asn Val Gln Arg Lys Ala Ile His Glu
100 105 110
Leu Ile Gln Val Met Ala Glu Leu Ser Pro Ala Ala Lys Thr Gly Lys
115 120 125
Arg Lys Arg Ser Gln Met Leu Phe Arg Gly
130 135

<210> 342
<211> 112
<212> PRT
<213> Homo sapiens
<400> 342

Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys
1 5 10 15

Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp
20 25 30
Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly Pro Cys
35 40 45
Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu
50 55 60
Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys Val Pro
65 70 75 80

Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro
85 90 95
Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser

100 105 110

<210> 343

<211> 431

<212> PRT

<213> Homo sapiens

<400> 343

Met Pro Asn Pro Arg Pro Gly Lys Pro Ser Ala Pro Ser Leu Ala Leu
1 5 10 15
Gly Pro Ser Pro Gly Ala Ser Pro Ser Trp Arg Ala Ala Pro Lys Ala

20 25 30
Ser Asp Leu Leu Gly Ala Arg Gly Pro Gly Gly Thr Phe Gln Gly Arg
35 40 45
Asp Leu Arg Gly Gly Ala His Ala Ser Ser Ser Ser Leu Asn Pro Met
50 55 60
Pro Pro Ser Gln Leu Gln Leu Pro Thr Leu Pro Leu Val Met Val Ala
65 70 75 80
Pro Ser Gly Ala Arg Leu Gly Pro Leu Pro His Leu Gln Ala Leu Leu

85 90 95
Gln Asp Arg Pro His Phe Met His Gln Leu Ser Thr Val Asp Ala His
100 105 110
Ala Arg Thr Pro Val Leu Gln Val His Pro Leu Glu Ser Pro Ala Met
115 120 125
Ile Ser Leu Thr Pro Pro Thr Thr Ala Thr Gly Val Phe Ser Leu Lys
130 135 140
Ala Arg Pro Gly Leu Pro Pro Gly Ile Asn Val Ala Ser Leu Glu Trp

145 150 155 160
Val Ser Arg Glu Pro Ala Leu Leu Cys Thr Phe Pro Asn Pro Ser Ala
165 170 175
Pro Arg Lys Asp Ser Thr Leu Ser Ala Val Pro Gln Ser Ser Tyr Pro
180 185 190
Leu Leu Ala Asn Gly Val Cys Lys Trp Pro Gly Cys Glu Lys Val Phe
195 200 205

Glu Glu Pro Glu Asp Phe Leu Lys His Cys Gln Ala Asp His Leu Leu

210 215 220

Asp Glu Lys Gly Arg Ala Gln Cys Leu Leu Gln Arg Glu Met Val Gln
225 230 235 240
Ser Leu Glu Gln Gln Leu Val Leu Glu Lys Glu Lys Leu Ser Ala Met
245 250 255
Gln Ala His Leu Ala Gly Lys Met Ala Leu Thr Lys Ala Ser Ser Val
260 265 270
Ala Ser Ser Asp Lys Gly Ser Cys Cys Ile Val Ala Ala Gly Ser Gln

275 280 285

Gly Pro Val Val Pro Ala Trp Ser Gly Pro Arg Glu Ala Pro Asp Ser
290 295 300
Leu Phe Ala Val Arg Arg His Leu Trp Gly Ser His Gly Asn Ser Thr
305 310 315 320
Phe Pro Glu Phe Leu His Asn Met Asp Tyr Phe Lys Phe His Asn Met
325 330 335
Arg Pro Pro Phe Thr Tyr Ala Thr Leu Ile Arg Trp Ala Ile Leu Glu

340 345 350

Ala Pro Glu Lys Gln Arg Thr Leu Asn Glu Ile Tyr His Trp Phe Thr
355 360 365
Arg Met Phe Ala Phe Phe Arg Asn His Pro Ala Thr Trp Lys Asn Ala
370 375 380
Ile Arg His Asn Leu Ser Leu His Lys Cys Phe Val Arg Val Glu Ser
385 390 395 400
Glu Lys Gly Ala Val Trp Thr Val Asp Glu Leu Glu Phe Arg Lys Lys

405 410 415

Arg Ser Gln Arg Pro Ser Arg Cys Ser Asn Pro Thr Pro Gly Pro

420 425 430

<210> 344

<211> 417

<212> PRT

<213> Homo sapiens

<400> 344

Met Pro Asn Pro Arg Pro Gly Lys Pro Ser Ala Pro Ser Leu Ala Leu
1 5 10 15

Gly Pro Ser Pro Gly Ala Ser Pro Ser Trp Arg Ala Ala Pro Lys Ala
20 25 30

Ser Asp Leu Leu Gly Ala Arg Gly Pro Gly Gly Thr Phe Gln Gly Arg
35 40 45

Asp Leu Arg Gly Gly Ala His Ala Ser Ser Ser Ser Leu Asn Pro Met
50 55 60

Pro Pro Ser Gln Leu Gln Leu Pro Thr Leu Pro Leu Val Met Val Ala
65 70 75 80

Pro Ser Gly Ala Arg Leu Gly Pro Leu Pro His Leu Gln Ala Leu Leu
85 90 95

Gln Asp Arg Pro His Phe Met His Gln Leu Ser Thr Val Asp Ala His
100 105 110

Ala Arg Thr Pro Val Leu Gln Val His Pro Leu Glu Ser Pro Ala Met
115 120 125

Ile Ser Leu Thr Pro Pro Thr Thr Ala Thr Gly Val Phe Ser Leu Lys
130 135 140

Ala Arg Pro Gly Leu Pro Pro Gly Ile Asn Val Ala Ser Leu Glu Trp
145 150 155 160

Val Ser Arg Glu Pro Ala Leu Leu Cys Thr Phe Pro Asn Pro Ser Ala
165 170 175

Pro Arg Lys Asp Ser Thr Leu Ser Ala Val Pro Gln Ser Ser Tyr Pro
180 185 190

Leu Leu Ala Asn Gly Val Cys Lys Trp Pro Gly Cys Glu Lys Val Phe
195 200 205

Glu Glu Pro Glu Asp Phe Leu Lys His Cys Gln Ala Asp His Leu Leu
210 215 220

Asp Glu Lys Gly Arg Ala Gln Cys Leu Leu Gln Arg Glu Met Val Gln

225 230 235 240

Ser Leu Glu Gln Gln Leu Val Leu Glu Lys Glu Lys Leu Ser Ala Met

245 250 255

Gln Ala His Leu Ala Gly Lys Met Ala Leu Thr Lys Ala Ser Ser Val

260 265 270

Ala Ser Ser Asp Lys Gly Ser Cys Cys Ile Val Ala Ala Gly Ser Gln

275 280 285

Gly Pro Val Val Pro Ala Trp Ser Gly Pro Arg Glu Ala Pro Asp Ser

290 295 300

Leu Phe Ala Val Arg Arg His Leu Trp Gly Ser His Gly Asn Ser Thr

305 310 315 320

Phe Pro Glu Phe Leu His Asn Met Asp Tyr Phe Lys Phe His Asn Met

325 330 335

Arg Pro Pro Phe Thr Tyr Ala Thr Leu Ile Arg Trp Ala Ile Leu Glu

340 345 350

Ala Pro Glu Lys Gln Arg Thr Leu Asn Glu Ile Tyr His Trp Phe Thr

355 360 365

Arg Met Phe Ala Phe Phe Arg Asn His Pro Ala Thr Trp Lys Asn Ala

370 375 380

Ile Arg His Asn Leu Ser Leu His Lys Cys Phe Val Arg Val Glu Ser

385 390 395 400

Glu Lys Gly Ala Val Trp Thr Val Asp Glu Leu Glu Phe Arg Lys Lys

405 410 415

Arg

<210> 345

<211> 366

<212> PRT

<213> Homo sapiens

<400> 345

Gly Gly Ala His Ala Ser Ser Ser Ser Leu Asn Pro Met Pro Pro Ser

1 5 10 15

Gln Leu Gln Leu Pro Thr Leu Pro Leu Val Met Val Ala Pro Ser Gly
20 25 30
Ala Arg Leu Gly Pro Leu Pro His Leu Gln Ala Leu Leu Gln Asp Arg
35 40 45
Pro His Phe Met His Gln Leu Ser Thr Val Asp Ala His Ala Arg Thr

50 55 60
Pro Val Leu Gln Val His Pro Leu Glu Ser Pro Ala Met Ile Ser Leu
65 70 75 80
Thr Pro Pro Thr Thr Ala Thr Gly Val Phe Ser Leu Lys Ala Arg Pro
85 90 95
Gly Leu Pro Pro Gly Ile Asn Val Ala Ser Leu Glu Trp Val Ser Arg
100 105 110
Glu Pro Ala Leu Leu Cys Thr Phe Pro Asn Pro Ser Ala Pro Arg Lys

115 120 125
Asp Ser Thr Leu Ser Ala Val Pro Gln Ser Ser Tyr Pro Leu Leu Ala
130 135 140
Asn Gly Val Cys Lys Trp Pro Gly Cys Glu Lys Val Phe Glu Glu Pro
145 150 155 160
Glu Asp Phe Leu Lys His Cys Gln Ala Asp His Leu Leu Asp Glu Lys
165 170 175
Gly Arg Ala Gln Cys Leu Leu Gln Arg Glu Met Val Gln Ser Leu Glu

180 185 190
Gln Gln Leu Val Leu Glu Lys Glu Lys Leu Ser Ala Met Gln Ala His
195 200 205
Leu Ala Gly Lys Met Ala Leu Thr Lys Ala Ser Ser Val Ala Ser Ser
210 215 220
Asp Lys Gly Ser Cys Cys Ile Val Ala Ala Gly Ser Gln Gly Pro Val
225 230 235 240
Val Pro Ala Trp Ser Gly Pro Arg Glu Ala Pro Asp Ser Leu Phe Ala

245 250 255
Val Arg Arg His Leu Trp Gly Ser His Gly Asn Ser Thr Phe Pro Glu

260 265 270

Phe Leu His Asn Met Asp Tyr Phe Lys Phe His Asn Met Arg Pro Pro

275 280 285

Phe Thr Tyr Ala Thr Leu Ile Arg Trp Ala Ile Leu Glu Ala Pro Glu

290 295 300

Lys Gln Arg Thr Leu Asn Glu Ile Tyr His Trp Phe Thr Arg Met Phe

305 310 315 320

Ala Phe Phe Arg Asn His Pro Ala Thr Trp Lys Asn Ala Ile Arg His

325 330 335

Asn Leu Ser Leu His Lys Cys Phe Val Arg Val Glu Ser Glu Lys Gly

340 345 350

Ala Val Trp Thr Val Asp Glu Leu Glu Phe Arg Lys Lys Arg

355 360 365

<210> 346

<211> 787

<212> PRT

<213> Homo sapiens

<400> 346

Met Ala Val Trp Ile Gln Ala Gln Gln Leu Gln Gly Glu Ala Leu His

1 5 10 15

Gln Met Gln Ala Leu Tyr Gly Gln His Phe Pro Ile Glu Val Arg His

20 25 30

Tyr Leu Ser Gln Trp Ile Glu Ser Gln Ala Trp Asp Ser Val Asp Leu

35 40 45

Asp Asn Pro Gln Glu Asn Ile Lys Ala Thr Gln Leu Leu Glu Gly Leu

50 55 60

Val Gln Glu Leu Gln Lys Lys Ala Glu His Gln Val Gly Glu Asp Gly

65 70 75 80

Phe Leu Leu Lys Ile Lys Leu Gly His Tyr Ala Thr Gln Leu Gln Asn

85 90 95

Thr Tyr Asp Arg Cys Pro Met Glu Leu Val Arg Cys Ile Arg His Ile

100 105 110

```
Leu Tyr Asn Glu Gln Arg Leu Val Arg Glu Ala Asn Asn Gly Ser Ser
        115                 120                 125

Pro Ala Gly Ser Leu Ala Asp Ala Met Ser Gln Lys His Leu Gln Ile
    130                 135                 140
Asn Gln Thr Phe Glu Glu Leu Arg Leu Val Thr Gln Asp Thr Glu Asn
145                 150                 155                 160
Glu Leu Lys Lys Leu Gln Gln Thr Gln Glu Tyr Phe Ile Ile Gln Tyr
                165                 170                 175
Gln Glu Ser Leu Arg Ile Gln Ala Gln Phe Gly Pro Leu Ala Gln Leu
            180                 185                 190

Ser Pro Gln Glu Arg Leu Ser Arg Glu Thr Ala Leu Gln Gln Lys Gln
        195                 200                 205
Val Ser Leu Glu Ala Trp Leu Gln Arg Glu Ala Gln Thr Leu Gln Gln
    210                 215                 220
Tyr Arg Val Glu Leu Ala Glu Lys His Gln Lys Thr Leu Gln Leu Leu
225                 230                 235                 240
Arg Lys Gln Gln Thr Ile Ile Leu Asp Asp Glu Leu Ile Gln Trp Lys
                245                 250                 255

Arg Arg Gln Gln Leu Ala Gly Asn Gly Gly Pro Pro Glu Gly Ser Leu
            260                 265                 270
Asp Val Leu Gln Ser Trp Cys Glu Lys Leu Ala Glu Ile Ile Trp Gln
        275                 280                 285
Asn Arg Gln Gln Ile Arg Arg Ala Glu His Leu Cys Gln Gln Leu Pro
    290                 295                 300
Ile Pro Gly Pro Val Glu Glu Met Leu Ala Glu Val Asn Ala Thr Ile
305                 310                 315                 320

Thr Asp Ile Ile Ser Ala Leu Val Thr Ser Thr Phe Ile Ile Glu Lys
                325                 330                 335
Gln Pro Pro Gln Val Leu Lys Thr Gln Thr Lys Phe Ala Ala Thr Val
            340                 345                 350
Arg Leu Leu Val Gly Gly Lys Leu Asn Val His Met Asn Pro Pro Gln
```

355 360 365

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

370 375 380

Asn Glu Asn Thr Arg Asn Asp Tyr Ser Gly Glu Ile Leu Asn Asn Cys

385 390 395 400

Cys Val Met Glu Tyr His Gln Ala Thr Gly Thr Leu Ser Ala His Phe

405 410 415

Arg Asn Met Ser Leu Lys Arg Ile Lys Arg Ser Asp Arg Arg Gly Ala

420 425 430

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

435 440 445

Ser Val Gly Gly Asn Glu Leu Val Phe Gln Val Lys Thr Leu Ser Leu

450 455 460

Pro Val Val Val Ile Val His Gly Ser Gln Asp Asn Asn Ala Thr Ala

465 470 475 480

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

485 490 495

Ala Val Pro Asp Lys Val Leu Trp Pro Gln Leu Cys Glu Ala Leu Asn

500 505 510

Met Lys Phe Lys Ala Glu Val Gln Ser Asn Arg Gly Leu Thr Lys Glu

515 520 525

Asn Leu Val Phe Leu Ala Gln Lys Leu Phe Asn Asn Ser Ser Ser His

530 535 540

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

545 550 555 560

Glu Asn Leu Pro Gly Arg Asn Tyr Thr Phe Trp Gln Trp Phe Asp Gly

565 570 575

Val Met Glu Val Leu Lys Lys His Leu Lys Pro His Trp Asn Asp Gly

580 585 590

Ala Ile Leu Gly Phe Val Asn Lys Gln Gln Ala His Asp Leu Leu Ile

595 600 605

Asn Lys Pro Asp Gly Thr Phe Leu Leu Arg Phe Ser Asp Ser Glu Ile
610 615 620

Gly Gly Ile Thr Ile Ala Trp Lys Phe Asp Ser Gln Glu Arg Met Phe
625 630 635 640

Trp Asn Leu Met Pro Phe Thr Thr Arg Asp Phe Ser Ile Arg Ser Leu
645 650 655

Ala Asp Arg Leu Gly Asp Leu Asn Tyr Leu Ile Tyr Val Phe Pro Asp
660 665 670

Arg Pro Lys Asp Glu Val Tyr Ser Lys Tyr Tyr Thr Pro Val Pro Cys
675 680 685

Glu Ser Ala Thr Ala Lys Ala Val Asp Gly Tyr Val Lys Pro Gln Ile
690 695 700

Lys Gln Val Val Pro Glu Phe Val Asn Ala Ser Ala Asp Ala Gly Gly
705 710 715 720

Gly Ser Ala Thr Tyr Met Asp Gln Ala Pro Ser Pro Ala Val Cys Pro
725 730 735

Gln Ala His Tyr Asn Met Tyr Pro Gln Asn Pro Asp Ser Val Leu Asp
740 745 750

Thr Asp Gly Asp Phe Asp Leu Glu Asp Thr Met Asp Val Ala Arg Arg
755 760 765

Val Glu Glu Leu Leu Gly Arg Pro Met Asp Ser Gln Trp Ile Pro His
770 775 780

Ala Gln Ser
785

<210> 347
<211> 526
<212> PRT
<213> Homo sapiens
<400> 347

Met Glu Thr Glu Ala Ile Asp Gly Tyr Ile Thr Cys Asp Asn Glu Leu
1 5 10 15

Ser Pro Glu Arg Glu His Ser Asn Met Ala Ile Asp Leu Thr Ser Ser

20 25 30

Thr Pro Asn Gly Gln His Ala Ser Pro Ser His Met Thr Ser Thr Asn

35 40 45

Ser Val Lys Leu Glu Met Gln Ser Asp Glu Glu Cys Asp Arg Lys Pro

50 55 60

Leu Ser Arg Glu Asp Glu Ile Arg Gly His Asp Glu Gly Ser Ser Leu

65 70 75 80

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

85 90 95

Gln Glu Leu Gln Gly Glu Gly Gly Ile Arg Leu Pro Asn Gly Lys Leu

100 105 110

Lys Cys Asp Val Cys Gly Met Val Cys Ile Gly Pro Asn Val Leu Met

115 120 125

Val His Lys Arg Ser His Thr Gly Glu Arg Pro Phe His Cys Asn Gln

130 135 140

Cys Gly Ala Ser Phe Thr Gln Lys Gly Asn Leu Leu Arg His Ile Lys

145 150 155 160

Leu His Ser Gly Glu Lys Pro Phe Lys Cys Pro Phe Cys Ser Tyr Ala

165 170 175

Cys Arg Arg Arg Asp Ala Leu Thr Gly His Leu Arg Thr His Ser Val

180 185 190

Gly Lys Pro His Lys Cys Asn Tyr Cys Gly Arg Ser Tyr Lys Gln Arg

195 200 205

Ser Ser Leu Glu Glu His Lys Glu Arg Cys His Asn Tyr Leu Gln Asn

210 215 220

Val Ser Met Glu Ala Ala Gly Gln Val Met Ser His His Val Pro Pro

225 230 235 240

Met Glu Asp Cys Lys Glu Gln Glu Pro Ile Met Asp Asn Asn Ile Ser

245 250 255

Leu Val Pro Phe Glu Arg Pro Ala Val Ile Glu Lys Leu Thr Gly Asn

260 265 270

```
Met Gly Lys Arg Lys Ser Ser Thr Pro Gln Lys Phe Val Gly Glu Lys
        275                 280                 285
Leu Met Arg Phe Ser Tyr Pro Asp Ile His Phe Asp Met Asn Leu Thr

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

        355                 360                 365
Ser Arg Glu Thr Ala Asp Ser His Glu Asn Asn Met Asp Gly Pro Ile
    370                 375                 380
Ser Leu Ile Arg Pro Lys Ser Arg Pro Gln Glu Arg Glu Ala Ser Pro
385                 390                 395                 400
Ser Asn Ser Cys Leu Asp Ser Thr Asp Ser Glu Ser Ser His Asp Asp
                405                 410                 415
His Gln Ser Tyr Gln Gly His Pro Ala Leu Asn Pro Lys Arg Lys Gln

            420                 425                 430
Ser Pro Ala Tyr Met Lys Glu Asp Val Lys Ala Leu Asp Thr Thr Lys
        435                 440                 445
Ala Pro Lys Gly Ser Leu Lys Asp Ile Tyr Lys Val Phe Asn Gly Glu
    450                 455                 460
Gly Glu Gln Ile Arg Ala Phe Lys Cys Glu His Cys Arg Val Leu Phe
465                 470                 475                 480
Leu Asp His Val Met Tyr Thr Ile His Met Gly Cys His Gly Tyr Arg

                485                 490                 495
Asp Pro Leu Glu Cys Asn Ile Cys Gly Tyr Arg Ser Gln Asp Arg Tyr
            500                 505                 510
Glu Phe Ser Ser His Ile Val Arg Gly Glu His Thr Phe His
```

515 520 525

<210> 348

<211> 625

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

AP1.0 sequence"

<400> 348

attctcgggc tacggccctg gagccactcc ggctcctaaa gatttagaag tttgagcaca 60

cccgcccact agggcccccc atccaggggg gcaacgggca agcacttctg tttccccggt 120

atgatctgat aggctgtaac cacggctgaa acagagatta tcgttatccg cttcactact 180

tcgagaagcc tagtaatgat gggtgaaatt gaatccgttg atccggtgtc tcccccacac 240

cagaaactca tgatgagggt tgccatcccg gctacggcga cgtagcgggc atccctgcgc 300

tggcatgagg cctcttagga ggacggatga tatggatctt gtcgtgaaga gcctattgag 360

ctagtgtcga ctcctccgcc cccgtgaatg cggctaatcc taaccccgga gcaggtgggt 420

ccaatccagg gcctggcctg tcgtaatgcg taagtctggg acggaaccga ctactttcgg 480

gaaggcgtgt ttccatttgt tcattatttg tgtgtttatg gtgacaactc tgggtaaacg 540

ttctattgcg tttattgaga gattcccaac aattgaacaa acgagaacta cctgttttat 600

taaatttaca cagagaagaa ttaca 625

<210> 349

<211> 549

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CK1.0 sequence"

<400> 349

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactcta ctgctgacaa ctcactgact atccacttgc 540

tctgtcacg 549

<210> 350

<211> 630

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

PV1.0 sequence"

<400> 350

aacaaaaggc tacaccactt gggctacggc ccgcgccacc ttgtggcgca aagacattag 60

aagaatagca taccgcccac tagggccctg cagccagcag ggtaacgggc aagcacttct 120

gtctccccgg tagaacggta taggctgtac ccacggccga aaactgaact atcgttaccc 180

gactccgtac ttcgcaaagc ttagtaggaa actggaaagt tcgagttatt gacccggagt 240

gttccccca ctccagaaac gcgtgatgag ggttgccacc ccgaccatgg cgacatggtg 300

ggcatccctg cgctggcacg cggcctctaa gaggataact cgctcctact ggtaaccgaa 360

gagccccgtg agctacggtt tattcctccg cctccctgaa tgcggctaat cctaacccat 420

gagcagttgc catagatcca tatggtggac tgtcgtaacg cgtaagttgt gggcggaacc 480

gactactttg ggatggcgtg tttccttgtt ttctccattt gttgttgtat ggtgacaagt 540

tatagatctc gatctatagc gtttcttgag agatttccaa acatttattc aagtcgtaca 600

attcttgtgt ttaagcagta cagtgtaacc 630

<210> 351

<211> 653

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

SV1.0 sequence"

<400> 351

tctgtcctca ccccatcttc ccttctttcc tgcaccgtta cgcttactcg catgtgcatt 60

gagtggtgca cgtgcttgaa caaacagcta cactcacatg ggggcgggtt ttcccgccct 120

gcggcctctc gcgaggccca cccctcccct tcctcccata actacagtgc tttggtaggt 180

aagcatcctg atcccccgcg gaagctgctc acgtggcaac tgtggggacc cagacaggtt 240

atcaaaggca cccggtcttt ccgccttcag gagtatccct gctagtgaat tctagtaggg 300

ctctgcttgg tgccaacctc ccccaaatgc gcgctgcggg agtgctcttc cccaactcac 360

cctagtatcc tctcatgtgt gtgcttggtc agcatatctg agacgatgtt ccgctgtccc 420

agaccagtcc agtaatggac gggccagtgt gcgtagtcgt cttccggctt gtccggcgca 480

tgtttggtga accggtgggg taaggttggt gtgcccaacg cccgtacttt ggtgatacct 540

caagaccacc caggaatgcc agggaggtac cccgcttcac agcgggatct gaccctgggc 600

taattgtcta cggtggttct tcttgcttcc acttctttct actgttcgcc acc 653

<210> 352

<211> 635

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 352

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactcta ctgctgacaa ctcactgact atccacttgc 540

tctcttgtgc ctttctgctc tggttcaagt tccttgattg tttttgactg cttttcactg 600

cttttcttct cacaatcctt gctcagttca aagtc 635

<210> 353

<211> 513

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 353

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactaaa gtc 513

<210> 354

<211> 549

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 354

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactcta ctgctgacaa ctcactgact atccacttgc 540

tctaaagtc 549

<210> 355

<211> 513

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 355

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactgcc acc 513

<210> 356

<211> 549

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 356

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactttc actgcttttc ttctcacaat ccttgctcag 540

ttcaaagtc 549

<210> 357

<211> 678

<212> DNA

<213> Parabovirus sp.

<400> 357

tgaaccgtta cgcaccactc agttggtgtt tggtggcacc aatgatggaa caaaaggcta 60

caccacttgg gctacggccc gcgccacctt gtggcgcaaa gacattagaa gaatagcata 120

ccgcccacta gggccctgca gccagcaggg taacgggcaa gcacttctgt ctccccggta 180

gaacggtata ggctgtaccc acggccgaaa actgaactat cgttacccga ctccgtactt 240

cgcaaagctt agtaggaaac tggaaagttc gagttattga cccggagtgt tccccccact 300

ccagaaacgc gtgatgaggg ttgccacccc gaccatggcg acatggtggg catccctgcg 360

ctggcacgcg gcctctaaga ggataactcg ctcctactgg taaccgaaga gccccgtgag 420

ctacggttta ttcctccgcc tccctgaatg cggctaatcc taacccatga gcagttgcca 480

tagatccata tggtggactg tcgtaacgcg taagttgtgg gcggaaccga ctactttggg 540

atggcgtgtt tccttgtttt ctccatttgt tgttgtatgg tgacaagtta tagatctcga 600

tctatagcgt ttcttgagag atttccaaac atttattcaa gtcgtacaat tcttgtgttt 660

aagcagtaca gtgtaagg 678

<210> 358

<211> 630

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 358

aacaaaaggc tacaccactt gggctacggc ccgcgccacc ttgtggcgca aagacattag 60

aagaatagca taccgcccac tagggccctg cagccagcag ggtaacgggc aagcacttct 120

gtctccccgg tagaacggta taggctgtac ccacggccga aaactgaact atcgttaccc 180

gactccgtac ttcgcaaagc ttagtaggaa actggaaagt tcgagttatt gacccggagt 240

gttccccca ctccagaaac gcgtgatgag ggttgccacc ccgaccatgg cgacatggtg 300

ggcatccctg cgctggcacg cggcctctaa gaggataact cgctcctact ggtaaccgaa 360

gagccccgtg agctacggtt tattcctccg cctccctgaa tgcggctaat cctaacccat 420

gagcagttgc catagatcca tatggtggac tgtcgtaacg cgtaagttgt gggcggaacc 480

gactactttg ggatggcgtg tttccttgtt ttctccattt gttgttgtat ggtgacaagt 540

tatagatctc gatctatagc gtttcttgag agatttccaa acatttattc aagtcgtaca 600

attcttgtgt ttaagcagta cagtgtaagg 630

<210> 359

<211> 611

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 359

tgggctacgg cccgcgccac cttgtggcgc aaagacatta gaagaatagc ataccgccca 60

ctagggccct gcagccagca gggtaacggg caagcacttc tgtctccccg gtagaacggt 120

ataggctgta cccacggccg aaaactgaac tatcgttacc cgactccgta cttcgcaaag 180

cttagtagga aactggaaag ttcgagttat tgacccggag tgttcccccc actccagaaa 240

cgcgtgatga gggttgccac cccgaccatg gcgacatggt gggcatccct gcgctggcac 300

gcggcctcta agaggataac tcgctcctac tggtaaccga agagccccgt gagctacggt 360

ttattcctcc gcctccctga atgcggctaa tcctaaccca tgagcagttg ccatagatcc 420

atatggtgga ctgtcgtaac gcgtaagttg tgggcggaac cgactacttt gggatggcgt 480

gtttccttgt tttctccatt tgttgttgta tggtgacaag ttatagatct cgatctatag 540

cgtttcttga gagatttcca aacatttatt caagtcgtac aattcttgtg tttaagcagt 600

acagtgtaag g 611

<210> 360

<211> 618

<212> DNA

<213> Artificial Sequence

<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"
<400> 360
tgaaccgtta cgcaccactc agttggtgtt tggtggcacc aatgatggaa caaaaggcta 60
caccacttgg gctacggccc gcgccacctt gtggcgcaaa gacattagaa gaatagcata 120
ccgcccacta gggccctgca gccagcaggg taacgggcaa gcacttctgt ctccccggta 180
gaacggtata ggctgtaccc acggccgaaa actgaactat cgttacccga ctccgtactt 240
cgcaaagctt agtaggaaac tggaaagttc gagttattga cccggagtgt tccccccact 300
ccagaaacgc gtgatgaggg ttgccacccc gaccatggcg acatggtggg catccctgcg 360

ctggcacgcg gcctctaaga ggataactcg ctcctactgg taaccgaaga gccccgtgag 420
ctacggttta ttcctccgcc tccctgaatg cggctaatcc taacccatga gcagttgcca 480
tagatccata tggtggactg tcgtaacgcg taagttgtgg gcggaaccga ctactttggg 540
atggcgtgtt tccttgtttt ctccatttgt tgttgtatgg tgacaagtta tagatctcga 600
tctatagcgt ttgtaagg 618
<210> 361
<211> 826
<212> DNA
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
Apodemus Picornavirus sequence"

<400> 361
tttgaaaggg gtgcggatat catggcgttt ctcgccatga tatccgcaca ttgcaaaccc 60
atattgcata cccactgggt atgcattatg gggaggcccc tttcacccct ccccccccaa 120
ttacctttc cccctctagt aaccatacgc tttactcagc gtaactactc cgggttacgt 180
gatgaagaag aggctacgga gattctcggg ctacggccct ggagccactc cggctcctaa 240
agatttagaa gtttgagcac acccgcccac tagggccccc catccagggg ggcaacgggc 300
aagcacttct gtttccccgg tatgatctga taggctgtaa ccacggctga aacagagatt 360
atcgttatcc gcttcactac ttcgagaagc ctagtaatga tgggtgaaat tgaatccgtt 420

gatccggtgt ctcccccaca ccagaaactc atgatgaggg ttgccatccc ggctacggcg 480

acgtagcggg catccctgcg ctggcatgag gcctcttagg aggacggatg atatggatct 540

tgtcgtgaag agcctattga gctagtgtcg actcctccgc ccccgtgaat gcggctaatc 600

ctaaccccgg agcaggtggg tccaatccag ggcctggcct gtcgtaatgc gtaagtctgg 660

gacggaaccg actactttcg ggaaggcgtg tttccatttg ttcattattt gtgtgtttat 720

ggtgacaact ctgggtaaac gttctattgc gtttattgag agattcccaa caattgaaca 780

aacgagaact acctgtttta ttaaatttac acagagaaga attaca 826

<210> 362

<211> 721

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 362

cccctccccc cccaattacc ttttcccccc ctagtaacca tacgctttac tcagcgtaac 60

tactccgggt tacgtgatga agaagaggct acggagattc tcgggctacg gccctggagc 120

cactccggct cctaaagatt tagaagtttg agcacacccg cccactaggg ccccccatcc 180

agggggggcaa cgggcaagca cttctgtttc cccggtatga tctgataggc tgtaaccacg 240

gctgaaacag agattatcgt tatccgcttc actacttcga gaagcctagt aatgatgggt 300

gaaattgaat ccgttgatcc ggtgtctccc ccacaccaga aactcatgat gagggttgcc 360

atcccggcta cggcgacgta gcgggcatcc ctgcgctggc atgaggcctc ttaggaggac 420

ggatgatatg gatcttgtcg tgaagagcct attgagctag tgtcgactcc tccgcccccg 480

tgaatgcggc taatcctaac cccggagcag gtgggtccaa tccagggcct ggcctgtcgt 540

aatgcgtaag tctgggacgg aaccgactac tttcgggaag gcgtgtttcc atttgttcat 600

tatttgtgtg tttatggtga caactctggg taaacgttct attgcgttta ttgagagatt 660

cccaacaatt gaacaaacga gaactacctg ttttattaaa tttacacaga gaagaattac 720

a 721

<210> 363

<211> 625

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 363

attctcgggc tacggccctg gagccactcc ggctcctaaa gatttagaag tttgagcaca 60

cccgcccact agggcccccc atccaggggg gcaacgggca agcacttctg tttccccggt 120

atgatctgat aggctgtaac cacggctgaa acagagatta tcgttatccg cttcactact 180

tcgagaagcc tagtaatgat gggtgaaatt gaatccgttg atccggtgtc tcccccacac 240

cagaaactca tgatgagggt tgccatcccg gctacggcga cgtagcgggc atccctgcgc 300

tggcatgagg cctcttagga ggacggatga tatggatctt gtcgtgaaga gcctattgag 360

ctagtgtcga ctcctccgcc cccgtgaatg cggctaatcc taaccccgga gcaggtgggt 420

ccaatccagg gcctggcctg tcgtaatgcg taagtctggg acggaaccga ctactttcgg 480

gaaggcgtgt ttccatttgt tcattatttg tgtgtttatg gtgacaactc tgggtaaacg 540

ttctattgcg tttattgaga gattcccaac aattgaacaa acgagaacta cctgttttat 600

taaatttaca cagagaagaa ttaca 625

<210> 364

<211> 766

<212> DNA

<213> Kobuvirus sp.

<400> 364

tttcacaccc tcttttccgg tggtccggac ccagaccacc gttactccat tcagctactt 60

cggtacctgt tcggaggaat taaacgggca ccctacccaa gggttacatg ggaccatatt 120

cctcctcccc tgtaacttta agttttgtgc ccgtattctt gactccaggc ggatgttgtg 180

tcgcccgtcc tgtgaacaaa cagctagaca ctttcctccc ctccctctgg gctgctccgg 240

cagtccactc cctccccca gcgtaacatg ccccgctgga gtgatgcacc tggaagtcgt 300

ggacgtgggt tagtaacttc ggtgaaaacc cactataatg acaactggtt gacccccaca 360

ctcaaaggac tcgagtcttt ctcccttaag gctagcccgg ccacatgaat ttgcagctgg 420

caactagtga gtccaccatg tcccgcaacc tcggctgcgg agtgctgttc cccaagcgta 480

tgccttcctt ctgtaagagt gcgcctggca agcacatctg agaagtcgtt ccgctgcgtc 540

gtgccaacct ggcgacaggt gacccagtgt gcgtagactt cttccggatt cgtccggctc 600

ttctctagga aacatgcgtg taaggttcat gtgccaaagc cctgcgcgcg gtgttcttct 660

actgccctag gaatgtgccg caggtacccc tacttcggta gggatctgag cggtagctaa 720

ttgtctacgg gtagtttcat ttccatcttc tcttcaggtc gacatc 766

<210> 365

<211> 608

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 365

ttgactccag gcggatgttg tgtcgcccgt cctgtgaaca aacagctaga cactttcctc 60

ccctccctct gggctgctcc ggcagtccac tccctccccc cagcgtaaca tgccccgctg 120

gagtgatgca cctggaagtc gtggacgtgg gttagtaact tcggtgaaaa cccactataa 180

tgacaactgg ttgaccccca cactcaaagg actcgagtct ttctccctta aggctagccc 240

ggccacatga atttgcagct ggcaactagt gagtccacca tgtcccgcaa cctcggctgc 300

ggagtgctgt tccccaagcg tatgccttcc ttctgtaaga gtgcgcctgg caagcacatc 360

tgagaagtcg ttccgctgcg tcgtgccaac ctggcgacag gtgacccagt gtgcgtagac 420

ttcttccgga ttcgtccggc tcttctctag gaaacatgcg tgtaaggttc atgtgccaaa 480

gccctgcgcg cggtgttctt ctactgccct aggaatgtgc cgcaggtacc cctacttcgg 540

tagggatctg agcggtagct aattgtctac gggtagtttc atttccatct tctcttcagg 600

tcgacatc 608

<210> 366

<211> 690

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 366

gaattaaacg ggcaccctac ccaagggtta catgggacca tattcctcct cccctgtaac 60

tttaagtttt gtgcccgtat tcttgactcc aggcggatgt tgtgtcgccc gtcctgtgaa 120

caaacagcta gacactttcc tcccctccct ctgggctgct ccggcagtcc actccctccc 180

cccagcgtaa catgccccgc tggagtgatg cacctggaag tcgtggacgt gggttagtaa 240

cttcggtgaa aacccactat aatgacaact ggttgacccc cacactcaaa ggactcgagt 300

ctttctccct taaggctagc ccggccacat gaatttgcag ctggcaacta gtgagtccac 360

catgtcccgc aacctcggct gcggagtgct gttccccaag cgtatgcctt ccttctgtaa 420

gagtgcgcct ggcaagcaca tctgagaagt cgttccgctg cgtcgtgcca acctggcgac 480

aggtgaccca gtgtgcgtag acttcttccg gattcgtccg gctcttctct aggaaacatg 540

cgtgtaaggt tcatgtgcca aagccctgcg cgcggtgttc ttctactgcc ctaggaatgt 600

gccgcaggta cccctacttc ggtagggatc tgagcggtag ctaattgtct acgggtagtt 660

tcatttccat cttctcttca ggtcgacatc 690

<210> 367

<211> 729

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 367

tttcacaccc tcttttccgg tggtccggac ccagaccacc gttactccat tcagctactt 60

cggtacctgt tcggaggaat taaacgggca ccctacccaa gggttacatg ggaccatatt 120

cctcctcccc tgtaacttta agttttgtgc ccgtattctt gactccaggc ggatgttgtg 180

tcgcccgtcc tgtgaacaaa cagctagaca ctttcctccc ctccctctgg gctgctccgg 240

cagtccactc cctcccccca gcgtaacatg ccccgctgga gtgatgcacc tggaagtcgt 300

ggacgtgggt tagtaacttc ggtgaaaacc cactataatg acaactggtt gacccccaca 360

ctcaaaggac tcgagtcttt ctcccttaag gctagcccgg ccacatgaat ttgcagctgg 420

caactagtga gtccaccatg tcccgcaacc tcggctgcgg agtgctgttc cccaagcgta 480

tgccttcctt ctgtaagagt gcgcctggca agcacatctg agaagtcgtt ccgctgcgtc 540

gtgccaacct ggcgacaggt gacccagtgt gcgtagactt cttccggatt cgtccggctc 600

ttctctagga aacatgcgtg taaggttcat gtgccaaagc cctgcgcgcg gtgttcttct 660

actgccctag gaatgtgccg caggtacccc tacttcggta gggatctgag cggtagctaa 720

ttggacatc 729

<210> 368

<211> 650

<212> DNA

<213> Salivirus sp.

<400> 368

tctgtcctca ccccatcttc ccttctttcc tgcaccgtta cgcttactcg catgtgcatt 60

gagtggtgca cgtgcttgaa caaacagcta cactcacatg ggggcgggtt ttcccgccct 120

gcggcctctc gcgaggccca cccctcccct tcctcccata actacagtgc tttggtaggt 180

aagcatcctg atcccccgcg gaagctgctc acgtggcaac tgtggggacc cagacaggtt 240

atcaaaggca cccggtcttt ccgccttcag gagtatccct gctagtgaat tctagtaggg 300

ctctgcttgg tgccaacctc ccccaaatgc gcgctgcggg agtgctcttc cccaactcac 360

cctagtatcc tctcatgtgt gtgcttggtc agcatatctg agacgatgtt ccgctgtccc 420

agaccagtcc agtaatggac gggccagtgt gcgtagtcgt cttccggctt gtccggcgca 480

tgtttggtga accggtgggg taaggttggt gtgcccaacg cccgtacttt ggtgatacct 540

caagaccacc caggaatgcc agggaggtac cccgcttcac agcgggatct gaccctgggc 600

taattgtcta cggtggttct tcttgcttcc acttctttct actgttcatg 650

<210> 369

<211> 508

<212> DNA

<213> Crohivirus B

<400> 369

gtataagaga caggtgtttg ccttgtcttc ggactggcat cttgggacca acccccctttt 60

tccccagcca tgggttaaat ggcaataaag gacgtaacaa ctttgtaacc attaagcttt 120

gtaattttgt aaccactaag ctttgtgcac ataatgtaac catcaagctt gttagtccca 180

gcaggaggtt tgcatgcttg tagccgaaat ggggctcgac cccccatagt aggatacttg 240

attttgcatt ccattgtgga cctgcaaact ctacacatag aggctttgtc ttgcatctaa 300

acacctgagt acagtgtgta cctagaccct atagtacggg aggaccgttt gtttcctcaa 360

taaccctaca taataggcta ggtgggcatg cccaatttgc aagatcccag actgggggtc 420

ggtctgggca gggttagatc cctgttagct actgcctgat agggtggtgc tcaaccatgt 480

gtagtttaaa ttgagctgtt catatacc 508

<210> 370

<211> 457

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 370

ccccccttttt ccccagccat gggttaaatg gcaataaagg acgtaacaac tttgtaacca 60

ttaagctttg taattttgta accactaagc tttgtgcaca taatgtaacc atcaagcttg 120

ttagtcccag caggaggttt gcatgcttgt agccgaaatg gggctcgacc ccccatagta 180

ggatacttga ttttgcattc cattgtggac ctgcaaactc tacacataga ggctttgtct 240

tgcatctaaa cacctgagta cagtgtgtac ctagacccta tagtacggga ggaccgtttg 300

tttcctcaat aaccctacat aataggctag gtgggcatgc ccaatttgca agatcccaga 360

ctgggggtcg gtctgggcag ggttagatcc ctgttagcta ctgcctgata gggtggtgct 420

caaccatgtg tagtttaaat tgagctgttc atatacc 457

<210> 371

<211> 742

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CVB3 sequence"

<400> 371

ttaaaacagc ctgtgggttg atcccaccca cagggcccat tgggcgctag cactctggta 60

tcacggtacc tttgtgcgcc tgttttatac cccctccccc aactgtaact tagaagtaac 120

acacaccgat caacagtcag cgtggcacac cagccacgtt ttgatcaagc acttctgtta 180

ccccggactg agtatcaata gactgctcac gcggttgaag gagaaagcgt tcgttatccg 240

gccaactact tcgaaaaacc tagtaacacc gtggaagttg cagagtgttt cgctcagcac 300

taccccagtg tagatcaggt cgatgagtca ccgcattccc cacgggcgac cgtggcggtg 360

gctgcgttgg cggcctgccc atggggaaac ccatgggacg ctctaataca gacatggtgc 420

gaagagtcta ttgagctagt tggtagtcct ccggcccctg aatgcggcta atcctaactg 480

cggagcacac accctcaagc cagagggcag tgtgtcgtaa cgggcaactc tgcagcggaa 540

ccgactactt tgggtgtccg tgtttcattt tattcctata ctggctgctt atggtgacaa 600

ttgagagatt gttaccatat agctattgga ttggccatcc ggtgaccaat agagctatta 660

tatatctctt tgttgggttt ataccactta gcttgaaaga ggttaaaaca ttacaattca 720

ttgttaagtt gaatacagca aa 742

<210> 372

<211> 651

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 372

ttaaaacagc ctgtgggttg atcccaccca cagggcccat tgggcgctag cactctggta 60

tcacggtacc tttgtgcgcc tgttttatac cccctccccc aactgtaact tagaagtaac 120

acacaccgat caacagtcag cgtggcacac cagccacgtt ttgatcaagc acttctgtta 180

ccccggactg agtatcaata gactgctcac gcggttgaag gagaaagcgt tcgttatccg 240

gccaactact tcgaaaaacc tagtaacacc gtggaagttg cagagtgttt cgctcagcac 300

taccccagtg tagatcaggt cgatgagtca ccgcattccc cacgggcgac cgtggcggtg 360

gctgcgttgg cggcctgccc atggggaaac ccatgggacg ctctaataca gacatggtgc 420

gaagagtcta ttgagctagt tggtagtcct ccggcccctg aatgcggcta atcctaactg 480

cggagcacac accctcaagc cagagggcag tgtgtcgtaa cgggcaactc tgcagcggaa 540

ccgactactt tgggtgtccg tgtttcattt tattcctata ctggctgctt atggtgacaa 600

ttgagagatt gttaccatat agctattgga ttggccatcc ggtgaagcaa a 651

<210> 373

<211> 721

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

SAFV sequence"

<400> 373

cacttattta attcggcctt ttgtgacaag cccctcggtg aaagaacctc tctcttttcg 60

acgtggttgg aattgccatc atttccgacg aaagtgctat catgcctccc cgattatgtg 120

atgttttctg ccctgctggg cggagcattc tcgggttgag aaaccttgaa tctttttctt 180

tggaaccttg gttcccccgg tctaagccgc ttggaatatg acagggttat tttcttgatc 240

ttatttctac ttttgcgggt tctatccgta aaaagggtac gtgctgcccc ttccttctct 300

ggagaattca cacggcggtc tttccgtctc tcaacaagtg tgaatgcagc atgccggaaa 360

cggtgaagaa aacagttttc tgtggaaatt tagagtgcac atcgaaacag ctgtagcgac 420

ctcacagtag cagcggactc ccctcttggc gacaagagcc tctgcggcca aaagccccgt 480

ggataagatc cactgctgtg agcggtgcaa ccccagcacc ctggttcgat gatcattctc 540

tatggaacca gaaaatggtt ttctcaagcc ctccggtaga gaagccaaga atgtcctgaa 600

ggtaccccgc gtgcgggatc tgatcaggag accaattggc ggtgctttac actgtcactt 660

tggtttaaaa attgtcacag cttctccaaa ccaagtggtc ttggttttcc aattttgttg 720

a 721

<210> 374

<211> 675

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 374

cctctctctt ttcgacgtgg ttggaattgc catcatttcc gacgaaagtg ctatcatgcc 60

tccccgatta tgtgatgttt tctgccctgc tgggcggagc attctcgggt tgagaaacct 120

tgaatctttt tctttggaac cttggttccc ccggtctaag ccgcttggaa tatgacaggg 180

ttattttctt gatcttattt ctacttttgc gggttctatc cgtaaaaagg gtacgtgctg 240

ccccttcctt ctctggagaa ttcacacggc ggtctttccg tctctcaaca agtgtgaatg 300

cagcatgccg gaaacggtga agaaaacagt tttctgtgga aatttagagt gcacatcgaa 360

acagctgtag cgacctcaca gtagcagcgg actcccctct tggcgacaag agcctctgcg 420

gccaaaagcc ccgtggataa gatccactgc tgtgagcggt gcaaccccag caccctggtt 480

cgatgatcat tctctatgga accagaaaat ggttttctca agccctccgg tagagaagcc 540

aagaatgtcc tgaaggtacc ccgcgtgcgg gatctgatca ggagaccaat tggcggtgct 600

ttacactgtc actttggttt aaaaattgtc acagcttctc caaaccaagt ggtcttggtt 660

ttccaatttt gttga 675

<210> 375

<211> 628

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 375

gtgctatcat gcctccccga ttatgtgatg ttttctgccc tgctgggcgg agcattctcg 60

ggttgagaaa ccttgaatct ttttctttgg aaccttggtt cccccggtct aagccgcttg 120

gaatatgaca gggttatttt cttgatctta tttctacttt tgcgggttct atccgtaaaa 180

agggtacgtg ctgccccttc cttctctgga gaattcacac ggcggtcttt ccgtctctca 240

acaagtgtga atgcagcatg ccggaaacgg tgaagaaaac agttttctgt ggaaatttag 300

agtgcacatc gaaacagctg tagcgacctc acagtagcag cggactcccc tcttggcgac 360

aagagcctct gcggccaaaa gccccgtgga taagatccac tgctgtgagc ggtgcaaccc 420

cagcaccctg gttcgatgat cattctctat ggaaccagaa aatggttttc tcaagccctc 480

cggtagagaa gccaagaatg tcctgaaggt accccgcgtg cgggatctga tcaggagacc 540

aattggcggt gctttacact gtcactttgg tttaaaaatt gtcacagctt ctccaaacca 600

agtggtcttg gttttccaat tttgttga 628

<210> 376

<211> 674

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 376

cacttattta attcggcctt ttgtgacaag cccctcggtg aaagaacctc tctcttttcg 60

acgtggttgg aattgccatc atttccgacg aaagtgctat catgcctccc cgattatgtg 120

atgttttctg ccctgctggg cggagcattc tcgggttgag aaaccttgaa tctttttctt 180

tggaaccttg gttcccccgg tctaagccgc ttggaatatg acagggttat tttcttgatc 240

ttatttctac ttttgcgggt tctatccgta aaaagggtac gtgctgcccc ttccttctct 300

ggagaattca cacggcggtc tttccgtctc tcaacaagtg tgaatgcagc atgccggaaa 360

cggtgaagaa aacagttttc tgtggaaatt tagagtgcac atcgaaacag ctgtagcgac 420

ctcacagtag cagcggactc ccctcttggc gacaagagcc tctgcggcca aaagccccgt 480

ggataagatc cactgctgtg agcggtgcaa ccccagcacc ctggttcgat gatcattctc 540

tatggaacca gaaaatggtt ttctcaagcc ctccggtaga gaagccaaga atgtcctgaa 600

ggtaccccgc gtgcgggatc tgatcaggag accaattggc ggtgctttac actgtcactt 660

tggtttaatg ttga 674

<210> 377

<211> 726

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 377

cacttattta attcggcctt ttgtgacaag cccctcggtg aaagaacctc tctcttttcg 60

acgtggttgg aattgccatc atttccgacg aaagtgctat catgcctccc cgattatgtg 120

atgttttctg ccctgctggg cggagcattc tcgggttgag aaaccttgaa tctttttctt 180

tggaaccttg gttcccccgg tctaagccgc ttggaatatg acagggttat tttcttgatc 240

ttatttctac ttttgcgggt tctatccgta aaaagggtac gtgctgcccc ttccttctct 300

ggagaattca cacggcggtc tttccgtctc tcaacaagtg tgaatgcagc atgccggaaa 360

cggtgaagaa aacagtttc tgtggaaatt tagagtgcac atcgaaacag ctgtagcgac 420

ctcacagtag cagcggactc ccctcttggc gacaagagcc tctgcggcca aaagccccgt 480

ggataagatc cactgctgtg agcggtgcaa ccccagcacc ctggttcgat gatcattctc 540

tatggaacca gaaaatggtt ttctcaagcc ctccggtaga gaagccaaga atgtcctgaa 600

ggtaccccgc gtgcgggatc tgatcaggag accaattggc ggtgctttac actgtcactt 660

tggtttaaaa attgtcacag cttctccaaa ccaagtggtc ttggttttcc aattttgttg 720

accgcc 726

<210> 378

<211> 549

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GLuc CK dCTG1 sequence"

<400> 378

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactcta cgtctgacaa ctcactgact atccacttgc 540

tctaaagtc 549

<210> 379

<211> 549

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GLuc CK dCTG1_2 sequence"

<400> 379

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactcta cgtcgtacaa ctcactgact atccacttgc 540

tctaaagtc 549

<210> 380

<211> 549

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GLuc CK dCTG1_2_3 sequence"

<400> 380

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactcta cgtcgtacaa ctcacgtact atccacttgc 540

tctaaagtc 549

<210> 381

<211> 549

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GLuc CK dAll sequence"

<400> 381

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactcta cgtcgtacaa ctcacgtact actcactgtc 540

tctaaagtc 549

<210> 382

<211> 549

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CK SZ1-L1S sequence"

<400> 382

gggggtgggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tccccctttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg tcgtcctggc aagactatga 300

cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt tccccagtgt gtgtagtgcg 360

atcttccagg tcctcctggt tggcgttgtc cagaaactgc ttcaggtaag tggggtgtgc 420

ccaatcccta caaaggttga accacccagg aatgccaggg aggtaccccg cttcacagcg 480

ggatctgacc ctgggctaat tgtctacggt ggttcttctt gcttccactt ctttctactg 540

ttcgccacc 549

<210> 383

<211> 553

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CK Aichi Scan (AV-S) sequence"

<400> 383

gggggtgggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tccccctttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg tcgtcctggc aagactatga 300

cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt tccccagtgt gtgtagtgcg 360

atcttccagg tcctcctggt tggcgttgtc cagaaactgc ttcaggtaag tggggtgtgc 420

ccaatcccta caaaggttga ttctttcacc accttaggaa tgctccggag gtaccccagc 480

aacagctggg atctgaccgg aggctaattg tctacgggtg gtgtttcatt tccaatcctt 540

ttatgtcgga gtc 553

<210> 384

<211> 667

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CK Aichi Loop (AV-L1) sequence"

<400> 384

gggggtgggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tccccctttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg tcgtcctggc aagactatga 300

cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt tccccagtgt gtgtagtgcg 360

atcttccagg tcctcctggt tggcgttgtc cagaaactgc ttcaggtaag tggggtgtgc 420

ccaatcccta caaaggttga actgccctag gaatgccagg caggtacccc acctccgggt 480

gggatctgag cctgggctaa ttgtctacgg gtagttttcc tttttctttt cacacaactc 540

tactgctgac aactcactga ctatccactt gctctcttgt gcctttctgc tctggttcaa 600

gttccttgat tgtttttgac tgcttttcac tgcttttctt ctcacaatcc ttgctcagtt 660

caaagtc 667

<210> 385

<211> 693

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CK SZ1-L2 sequence"

<400> 385

ggggggtggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tcccccttttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgcc tgatcccccg cggaagctgc 300

tcacgtggca actgtgggga cccagacagg ttatcaaagg cacccggtct ttccgccttc 360

aggagtatcc ctgctagtga attctagtag ggctctgctt gcgttgtcca gaaactgctt 420

caggtaagtg gggtgtgccc aatccctaca aaggttgatt ctttcaccac cttaggaatg 480

ctccggaggt accccagcaa cagctgggat ctgaccggag gctaattgtc tacgggtggt 540

gtttcctttt tcttttcaca caactctact gctgacaact cactgactat ccacttgctc 600

tcttgtgcct ttctgctctg gttcaagttc cttgattgtt tttgactgct tttcactgct 660

tttcttctca caatccttgc tcagttcaaa gtc 693

<210> 386

<211> 671

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CK Aichi TriLoop (AV-L2) sequence"

<400> 386

ggggggtggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tcccccttttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgca tgtgcctggc aagcatatct 300

gagaaggtgt tccgctgtgg ctgccaacct ggtgacaggt gccccagtgt gcgtaacctt 360

cttccgtctc cggacggtgc gttgtccaga aactgcttca ggtaagtggg gtgtgcccaa 420

tccctacaaa ggttgattct ttcaccacct taggaatgct ccggaggtac cccagcaaca 480

gctgggatct gaccggaggc taattgtcta cgggtggtgt ttccttttc ttttcacaca 540

actctactgc tgacaactca ctgactatcc acttgctctc ttgtgccttt ctgctctggt 600

tcaagttcct tgattgtttt tgactgcttt tcactgcttt tcttctcaca atccttgctc 660

agttcaaagt c 671

<210> 387

<211> 523

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 387

gggggtgggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tccccctttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg tcgtcctggc aagactatga 300

cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt tccccagtgt gtgtagtgcg 360

atcttccagg tcctcctggt tggcgttgtc cagaaactgc ttcaggtaag tggggtgtgc 420

ccaatcccta caaaggttga ttctttcacc accttaggaa tgctccggag gtaccccagc 480

aacagctggg atctgaccgg aggctaattg tctacgggtg gtg 523

<210> 388

<211> 592

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 388

gggggtgggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tccccctttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg tcgtcctggc aagactatga 300

cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt tccccagtgt gtgtagtgcg 360

atcttccagg tcctcctggt tggcgttgtc cagaaactgc ttcaggtaag tggggtgtgc 420

ccaatcccta caaaggttga tttccttttt cttttcacac aactctactg ctgacaactc 480

actgactatc cacttgctct cttgtgcctt tctgctctgg ttcaagttcc ttgattgttt 540

ttgactgctt ttcactgctt ttcttctcac aatccttgct cagttcaaag tc 592

<210> 389

<211> 572

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 389

gggggtgggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tccccctttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg cgttgtccag aaactgcttc 300

aggtaagtgg ggtgtgccca atccctacaa aggttgattc tttcaccacc ttaggaatgc 360

tccggaggta ccccagcaac agctgggatc tgaccggagg ctaattgtct acgggtggtg 420

tttccttttt cttttcacac aactctactg ctgacaactc actgactatc cacttgctct 480

cttgtgcctt tctgctctgg ttcaagttcc ttgattgttt ttgactgctt ttcactgctt 540

ttcttctcac aatccttgct cagttcaaag tc 572

<210> 390

<211> 167

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 390

ccctgcagcc gtcaccgtaa gtttgaagtt accgcatatc agcctctgct tcccagcgcg 60

tccaattcct gttcttattg tttcccctcc aggcgttacg cgtgacgacg aactgtgtcg 120

cagctaccac attattccgg agccttcatt ctcgcggctc tgatcgt 167

<210> 391

<211> 43

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 391

ggagaccgcg gccacgccga gtaggatcga gggtacagtc tcc 43

<210> 392

<211> 167

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 392

gacaccagga tcactcttgc tctgacccgc cctgtgtaga atagactcat gcttccctaa 60

gacctggatt tcttcccagg cactttcacc cgcctgccct gctccttcag tggactgcac 120

ccagggaggc ggtctctgac tgtcctttac tttctattct ggattgc 167

<210> 393

<211> 30

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 393

aaaccccccct aagccgccgc cgccgccacc 30

<210> 394

<211> 18

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 394

ccccccccaac ccgtcacg 18

<210> 395

<211> 6

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 395

gtcacg 6

<210> 396

<211> 141

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 396

tctgcgcact cgtaatcagt actaaccccc ctttgtcgga cactatgcga taatcgatcc 60

gccttttttca ccgccttcgg aattttattt acctcaactg atcctggagt ctctcttggt 120

tttcacggag gcctccgccc a 141

<210> 397

<211> 43

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 397
ggagaccgcg gccacgccga gtaggatcga gggtacagtc tcc 43
<210> 398
<211> 141
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"
<400> 398
ccccttgaaa cccccgcccc aggttcagtc tctcttcatc cctctgtcct gcatggtgat 60
acaaagaccc tttgtggacc ctaagccatg tagttgctgc tccctccttc cagttgtgaa 120
tattggtttc tgttaatcac a 141
<210> 399
<211> 30
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 399
aaaccccccct aagccgccgc cgccgccacc 30

<210> 400
<211> 18
<212> DNA
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"
<400> 400
ccccccccaac ccgtcacg 18

<210> 401

<211> 6

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 401

gtcacg 6

<210> 402

<211> 6

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 402

gtcacg 6

<210> 403

<211> 43

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 403

aataagagag aaaagaagag taagaagaaa tataagagcc acc 43

<210> 404

<211> 49

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 404

cgaactagta ttcttctggt ccccacagac tcagagagaa cccgccacc 49

<210> 405

<211> 7

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 405

agccacc 7

<210> 406

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 406

tgatagctaa ctag 14

<210> 407

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 407

tagtagctaa ctag 14

<210> 408

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 408

tgatgactga gtga 14

<210> 409

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 409

tagtagctag gtag 14

<210> 410

<211> 3

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 410

taa 3

<210> 411

<211> 14

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 411

taatagctaa ctag 14

<210> 412

<211> 15

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 412

taactagcta actag 15

<210> 413

<211> 579

<212> DNA

<213> Rhopalosiphum padi virus

<400> 413

gataaaagaa cctataatcc cttcgcacac cgcgtcacac cgcgctatat gctgctcatt 60

aggaattacg gctccttttt tgtggataca atctcttgta tacgatatac ttattgttaa 120

tttcattgac ctttacgcaa tcctgcgtaa atgctggtat agggtgtact tcggatttcc 180

gagcctatat tggttttgaa aggaccttta agtccctact atactacatt gtactagcgt 240

aggccacgta ggcccgtaag atattataac tattttatta tattttattc acccccaca 300

ttaatcccag ttaaagcttt ataactataa gtaagccgtg ccgaaacgtt aatcggtcgc 360

tagttgcgta acaactgtta gtttaatttt ccaaaattta tttttcacaa tttttagtta 420

agattttagc ttgccttaag cagtctttat atcttctgta tattatttta aagtttatag 480

gagcaaagtt cgctttactc gcaatagcta ttttatttat tttaggaata ttatcacctc 540

gtaattattt aattataaca ttagctttat ctatttata 579

<210> 414

<211> 675

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 414

cttgattcta accttgccgt atggtgccct aacgggttca tttaatcatg cgatgagggt 60

tgctataccg catccattct aaggcgattc aatgcttcat ttaggaattt tgttgacgat 120

taaaaggtac ccccacaaaa acaaaaccaa tcttacttga ttttcgtttt aactgaccac 180

tgcgatccca aattttcgcc ttcttatcaa agtatgttgt gttctttggg tgtacaacct 240

gagaacttgt ctacaactac atattactcg aggaagaaat tcggtttaag ccgtgccttc 300

tcacgtttag tatatctatc tggacacacc ttcttcatct tctaatcccc atctagtctc 360

ctgatcagag acgtcgttat taacaaataa ccccccttgt taataagaga caaagtacaa 420

tcaagctaag ttctcttgga gttcctgtag gaacttagcc attgtgatag agtcataagt 480

ctatgtgcat agacagctct agctcaccat ttccttccca acccatcttt tcatcagctt 540

aactctatga atccgatgca aaaaccattc taacatctta tggtgctttc caagccaaat 600

gagagctcac tcttttgagc cgctatttaa tggacaataa acgttttata gtgtacatca 660

tattgtaaaa acaaa 675

<210> 415

<211> 636

<212> DNA

<213> Oscivirus sp.

<400> 415

cctcggtccc tctttccgtc gccgcccacg acgttaaatg cggtgttgtg gtgcttaggt 60

gccacaccac tgctatttgg gtcccccttt cccctatata tgtttgtttg tttatttcaa 120

tttcttgagg attggcacct ccttatgcca aatctaaatc gtggaggatc ccaggctttc 180

tggtctttaa cagaactcca cgtccaggtc atagaaactg gttggtaggc tgcctgagta 240

gtccatttgc tagtagtccc ttgtgaacag ggtggctccc gtttactgct ggtattcccg 300

gtgtaggtcg ccatggtggt aacaccatcc tgcattgtgt gtgaaccagt accgcaagga 360

tagcaaggta tgaacacttg tggacgaaat ggtaagtgat caattcactt tcatggccgg 420

aaggtcacgt ggcaatcatg ccacccaggt accctcctct gggaggatct gagggtgggc 480

taagcagacc ctgccatgtg gctgaacttt tcccttattg ttttactttg taacatttat 540

agttgtgtta gtgatttgtg tgttgtgccc ttgtgagcta tatccagtat aagttcgcag 600

ctagaagtta atccttcgac atcggctgta ttggaa 636

<210> 416

<211> 675

<212> DNA

<213> Cadicivirus B

<400> 416

caccaaccct tgacctgtaa tgtcagtgga cagagtgctc ctctgttccc ggttaccgtg 60

ttccaggaca cgattgtaat cctgcgcctc accagcgctg cgtgcacgtc tgcataagga 120

aacgtgcctt ccccatgtct ctatcaattc tttggtgagt gaccgcccta gttgctcatc 180

ctatgggatt cttctctcat gggttctttg tggcatgcga atgtcacctt aattggaggt 240

ctttaattag atatcctttc ttcatctttg atatgagtgt cggaatttga ttcctagtct 300

ctgcaaaaca accccacttg atgaattcaa cttttcaacc gcacaaacat aatcaggttt 360

ttaaattgaa tgtttctaaa ttctaaattt agtttattta agtagtttgc catcttgact 420

cgatgtaaaa ttgtcataca agtcttcttt tcttttcttt acactttgaa gtttgcactt 480

agcagtcgtt ctgcacagct ttcgagtttt gtttgatcga catcgcaact tccacccacc 540

tctctttttc tagtgttgaa tgcggctaat cctaacccga gagcaataaa cccaggttta 600

ttgtcgtaac gcgcaagtct tggacggaac cgactataca cacacctctt tacccttag 660

tacacccttg gtacg 675

<210> 417

<211> 570

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 417

gcaaaatggg tacgtagtta accactgcgt atcaggattg caggccacga agggtatttg 60

catatctttc tatgcggtat tacggcttaa aacccgttgt atcttgttgt ttgactgcct 120

gtatcactag tggccatttt atttaggtta gagacccctg atagtaggag agttacaaac 180

tctttaaaaa ttgttgaccc cggaaaagat ggtgacccct gtaagtagtt gatcaagaag 240

atctatgcgc tggcatagta atccagtgtt tcctgtttta ggatgacctc tgaaagtaga 300

tgaccgtgga aagtcacgta gtgccccaat aagcacgttt gggcagcgtg cgctatcaca 360

aggcttgatc tccgaggagc cccttgtttt agctggctgg aagccaatga tcttaagtag 420

ataagtgctg ttgcttgtag ttcaacagaa agctttgagt acgtctttct tgcgagaaag 480

aacacatgca ttcttatgct ctcaattcta ttatttttat tttgggcgaa aggaaagctc 540

tcacgcgagt acgaatagcc aaccctttat 570

<210> 418

<211> 145

<212> DNA

<213> Plautia stali intestine virus

<400> 418

gctgactatg tgatcttatt aaaattaggt taaatttcga ggttaaaaat agttttaata 60

ttgctatagt cttagaggtc ttgtatattt atacttacca cacaagatgg accggagcag 120

ccctccaata tctagtgtac cctcg 145

<210> 419

<211> 312

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

PV Mahoney sequence"

<400> 419

atgagtctgg acatccctca ccggtgacgg tggtccaggc tgcgttggcg gcctacctat 60

ggctaacgcc atgggacgct agttgtgaac aaggtgtgaa gagcctattg agctacataa 120

gaatcctccg gcccctgaat gcggctaatc ccaacctcgg agcaggtggt cacaaaccag 180

tgattggcct gtcgtaacgc gcaagtccgt ggcggaaccg actactttgg gtgtccgtgt 240

ttccttttat tttattgtgg ctgcttatgg tgacaatcac agattgttat cataaagcga 300

attggattgg cc 312

<210> 420

<211> 577

<212> DNA

<213> Reticuloendotheliosis virus

<400> 420

ggggtcgccg tcctacacat tgttgtgacg tgcggcccag attcgaatct gtaataaaag 60

ctttttcttc tatatcctca gattggcagt gagaggagat tttgttcgtg gtgttggctg 120

gcctactggg tggggtaggg atccggactg aatccgtagt atttcggtac aacatttggg 180

ggctcgtccg ggattcctcc ccatcggcag aggtgcctac tgtttcttcg aactccggcg 240

ccggtaagta agtacttgat tttggtacct cgcgagggtt tgggaggatc ggagtggcgg 300

gacgctgccg ggaagctcca cctccgctca gcaggggacg ccctggtctg agctctgtgg 360

tatctgattg ttgttgaacc gtctctaaga cggtgatact ataagtcgtg gtttgtgtgt 420

ttgtttgtta ccttgtgttt gttcgtcact tgtcgacagc gccctgcgaa ttggtgtacc 480

cacaccgcgc ggcttgcgaa taatactttg gagagtcttt tgcctccagt gtcttccgtt 540

tgtactcgtc ctcctctccc tctccggccg ggatggg 577

<210> 421

<211> 675

<212> DNA

<213> Tropivirus A

<400> 421

tgtcgcatgt tgccaacatc aaaattctgg gagagtcgcg aactccttaa cactgccttg 60

cctcgacgga gccgttgtta tagtgtcgac gggatacaaa cattaaacta aacccacttg 120

cctcgacgga accccttacc ttttatttta ttttatagta tgaaagtgaa tcttgtatga 180

atgttcatag aaaactgcaa atgagtacca cgtctaacat gagagaatga tactggagaa 240

atccaagttt agaagtcact acgaatccca gcggaaacaa gggaattctg agcttctaat 300

aggcgtttaa gactatttgc aaaattctgg tgcgtaagtg atattttcat tgcgtagaac 360

gctggtaacc actccggcta gtataagcat tgttagtcac ttattatgaa actccacact 420

atcctttctg gagaagcaca caaacttaca tggtaaagct agaccattat cttaagcggt 480

gagtacactg caaccttgta acaatgcttg tatgactact ttttgtatat cttgagcaat 540

attgttgagg tggacatgtc caaaggtaat gttgttggga atggaggggt ccattttccc 600

gtgcacgtag tgtactagta ttgggtgata gccttgcggc ggatcaacca tgtattttaa 660

tccgttgact ttcac 675

<210> 422

<211> 592

<212> DNA

<213> Symapivirus A

<400> 422

ttgggaaatc cccaatgctt ctttcaacac cgcctgacta tgcggtggcg cttcggctca 60

aacaactagt cacttccccc tcttaactac tacccaagac ttctaactac ccttacctac 120

ttatttgtct aaatttcaaa cttttattct cacgcgtctt ataaacatct tttctatttg 180

ttatggtatg ttttgtgatt tgtgtggtgt atttcattta atgggatcta gtggaccgtg 240

ccccggttgg gtatccgctc cctttaaatg tttgcaagca ctcttgacat tataacctat 300

catttagttt acttgtttgt atgatcgtat ttctgaatcg taacatttat gcaattcttt 360

ctcgccgaga cttgtctagg agataaagtt cctgcatatt tagtgttacg gttgtataat 420

ggagacttag atagcttcac actgaggacg ctttttcgct atccttttga cctgattcag 480

gccagtgtgg agttaatgat tgtatggatg ggccctacaa tttgtctaag acttggtgat 540

agcctcgcgg ccgctcgcca tttatacaac tgaatagcgg ttgaaactct ct 592

<210> 423

<211> 591

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 423

tcacgcgctt ttccggtggt cacccaccgt tagggagcgc cagcgttcgc gcttccgcta 60

ccaggtgaca cactcctttc ccctccccca ttcccgttcc catcctctgg actggtttct 120

cctcacgatt gaccagcagc tgggagctgt taccagacgt tggacagtaa gtcccggatg 180

cactataggg ctggtggcta gtgcttggta agcactcaac gccataccta atgtgtacct 240

cggcttgccc tcctggtcgt ggtgaccggc tgtttctctt cccttggctc cagacgggct 300

ggtgtcctac caccaccgtt gcatgcagac ctccccctgc gcactcgaac gccctgtccc 360

agcagggtta gtatgtgctg tgcagatctg catgtgacac cccatccact ggtagagcag 420

gaagttgccc tagctaacgc ggcaagtatt actttccgct acacgtcctt gagattcctc 480

ggacctctgg aactagggtg actgtgggct tgggaaaacc caccttggtc ctgtactgcc 540

tgatagggtc gcggctggcc gaccagtgga tgtagccagt tgttttggga t 591

<210> 424

<211> 538

<212> DNA

<213> Rosavirus C

<400> 424

cagggagatc tccatgaata atcttttcca ccctctttag cgtctatgct attgaggacg 60

ggttggagcc ccgttgaccc agcgtcagag tgtgtcggta gcaggctttc tgctctcgcc 120

ccatgccggc cacacctccc attagtgatg tgaaggttgt aagttacatg tgaaaaggtt 180

tctaataatt gagctgaatg tagcgattac ctaaggtgag cggattcccc cacgtggtaa 240

cacgtgcctc tcaggccaaa agccaaggtg ttaaaagcac cccttaggta ggccactacc 300

ccgtggcctc agttctctta gaagattcac ttagtagtgt gtgcactggc aactcttaag 360

cagagctagt gagtgggcta aggatgccct gaaggtaccc gcaggtaacg ttaagacact 420

gtggatctga tcaggggctc gagtgctgaa gctttacaga ggtagctcga gttaaaaaac 480

gtctatgccc ctcccccacg ggagtggggg ttcccccaca ccaattttag attgcact 538

<210> 425

<211> 675

<212> DNA

<213> Rosavirus sp.

<400> 425

ccaggcatgg cgttaaacat gcattccctt cccctagtaa cctcccttcg cccccttcccc 60

acgttgtacc ccctccgaga tggctgctaa ggcgcttgct gctacagcag tctcgtgttt 120

cgggtgttat aagtgctttc ttttccactc cactccctgc ctatggggag cggaacggcc 180

ttgtctcggt cgttgcttct tgcagatctt caccccctcca ggctttctgg actcgccagg 240

ggtggagtag taggcgcact gtctaagtga aggtagcagt gttgttggcg aagagttgtg 300

gacctacttt gagtttgtag cgatcatcca gagctagcgg atctccccac gcggtaacgc 360

gtgcctctag gcccaaaagg cacggtgttc acagcaccct ttggatggcg ggggtgcccc 420

cctccgcact taaagtagaa aaacagctta gtagtcaaat aacatggctt tcctcaagca 480

ttcagtgcta catgggactg aaggatgccc agaaggtacc cgcaggcaac gataagctca 540

ctgtggatct gatctggggc cctgggccag gtgctataca cctggttaaa accaaatctg 600

gtagtcaggg ttaaaaaacg tctaagtccc acccccccgg ggacgggggg ttcccttaaa 660

ccctcaactg acacc 675

<210> 426

<211> 656

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Rhimavirus A sequence"

<400> 426

cgaattccgg acatctcctt tcgggggcga gcgtcaccgt gcccctcatg gaggcaactg 60

tgcctctaat cggtgaccca ctgagaaaat tttctttcta cgtggctaaa caatgcaact 120

ttataataac acaaatttaa tgcttaatct taacaccaaa gatttgaaca tatgtttgga 180

aagtggcaca cttcaaacat tgcatagttg ctaggggtga agtcccttta aggggttgca 240

gaggatcttt cctctttatg agcggctagg agtatcttct tgatattatg tggtcgtgca 300

actcacttcc cagatgtatg acggtgtact aagcgattgg aactagtcat aacctctttg 360

aattttggta ttgcgagtct agcaggggga tatttaccgc taaagggtga cacactcgtg 420

agggtggcct ttggtgtgtg tatatttatt ccgcccatct tgcatggggt gctaaaattc 480

taatgctgtg aaataaccat tttctgaata cattctctac atttggagtc aaatatgagg 540

aatgccactc aggtaccctt gacatgatct tggatctgag agtgggctaa ttatctaatt 600

atttggcgac tttctaaaat cttctgtttt tagtggtgac aatttatggt tataaa 656

<210> 427

<211> 487

<212> DNA

<213> Rafivirus sp.

<400> 427

gtgtccggga agcgactcaa gcttttgact gagtctctac accttcatcc gtaacatctt 60

taagtttatg tgcctatgga cctctagtgc actgccatca ccgggggtgt attggactgg 120

tttttccaca atccattcat cctgaggaat tttggctttg ttactaggat ggtcccacca 180

cacgcttatc tgtgcctatt gtgtcaacca tgttcttaag tagttgtgcc cgtgggtgag 240

tagataacca caacaatccg ataaagcatc tcgcaaggat gtgagtaatg gagtgtatgt 300

gctacagaga cccacaacct gaaccaagag agacacagtg aggattgtaa agggggaact 360

ctttgaaagg gcatgtcccg caattcctac tgactgacac cgggggttgg tgtcggtgga 420

ttttagcaaa tcctgttact gggtgatagc cttgtgcact tcacttggtt cttgtataag 480

tgctgta 487

<210> 428

<211> 560

<212> DNA

<213> Rafivirus sp.

<400> 428

tgcgaattta ttcgcacagt ctcttttccc ccatcttgtg tgtgtgatgg ggtaagccgc 60

agagtaatac ctactctgct gcaaacacac tcactctttt ctatctactt tatatcatgt 120

aataataagt agggaacata ttcaattcat attgttcatc tcactgaacc cgcatgaagg 180

actgcattgc atatcctgga cgaagtgacg tggaatattt ggacatttat ggattggaca 240

ccattacgct ttgtgcctct acggagatgt aaccataatc ttaagtagta gtaccccagc 300

acaagaggat aaagtggcat acacgacaac gggtgttgct cgcaccttag taatgtggat 360

gttcaccctt ggagcgtgct gaaactctgt gggtaaagac acacattagt acaaatgtgg 420

gggaactcac tgaaagggca tgtcccgtgt actggtgtgc cggaaagtgg gggtcgcttt 480

ctggagaact tagtagttct tgttattggg tgatagcctt gcggcggatc aactcacagt 540

tttaatccgt tgttttgcat 560

<210> 429

<211> 675

<212> DNA

<213> Poecivirus sp.

<400> 429

actacacaat cgcaacacgc gcaagtttgt agtttgattg gcgtgcaaat gtcaaatcaa 60

gcatataaca caatttggtg gctgttggtg tttgttatag gaattttggt tgtgttgaaa 120

ttgtggatgt gtaggaaata tgcacaatta cgtcagcgtc aggagtttta taacctggcg 180

caacaccaaa atggtcttcg cgctttaaca tcaccagcga ggtgtaaaca aattgaagtt 240

gaattagatc gtgtataggc cagggaacca tccctcccaa cgccacatct tgtggggaag 300

ttgggataat ggtgggtcta tatgaattgg tctgtagacc cacagtgaag agtgaatagt 360

atgcttgcgg ttccatttgt taatggtcta gcatgggtgg gggcggcaac cccgtgaggg 420

gttccccact ggccaaaagc ccaggggtta gtcatttcaa ccaaggaagc tggtaacctg 480

gtgacctgaa cttgagtggt gagacccccc tgctagagtg tgtaaaccga ttgtaagcat 540

tttgtttgct tagtatctgt ggtataagca gtcaattttg tataggctca aggctgtggt 600

agttagtaga tgcccggaag gttattactg atccggggac cgtgactata cattaggtaa 660

accggtttaa aaacc 675

<210> 430

<211> 519

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Megirivirus A LY sequence"

<400> 430

ttcgggacac tggatgggcg acttggtggg gctgccactc tatcttgacc tttcgttact 60

gactttcgga tctctgactc ctccttgtct cttgcgtttg gtccacggac ggactaattg 120

gaatgtttac tggctaagcc tcgttctgaa ataccctagc caatgggttg tagtaggatc 180

ctggtgtttc cattaaacct cttccgacca tagtagctag agttatggct gtgtaggatg 240

tgggtaagac cgctttttgc gtatctccca caagacaccg gattatggat gtgtccgctg 300

gataaggctc gaaacctccc aactgaaggt ggtgctgaaa tattgcaagc ctaggttgtg 360

tagaggcaag tagatgcctg ccgcgacatt cgtcttccgc ccttttgggt tagtagtgta 420

cctacatgga cgtggggctg ggaatcccca ccttgcataa cactggttga tagacctgcg 480

gctggtcaag ttactatggt ataaccagtt gaaatggct 519

<210> 431

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Megirivirus E sequence"

<400> 431

gcttggcaac ctcatatcgt tactctgccg accagtctgg gtcgtgtggc cacacaatgg 60

gattcgttct gttgtgtaga gtcacatggc attactgggc tgatcggtgg ggatccgttg 120

ccacccctaa acccttacat ttactggact gcttttcttg gccccggaat gattcgctca 180

cccgcgatga ggactgttgt tcttattatg gcaggattac gcgtctggtc cgcgtaagga 240

ctaattccta tgtttatacg ttactacctt gttctgaacg gtgggcgcca ccccgcctag 300

taggatcctg gcttatcgtg tagacctcta gggaccacat tagctagagt gtaggctgct 360

atggatggag tagtgacccc tttttgggta tcactctcta agactccgga atgtgtcata 420

gtacgctgga aatccttact tgtttttcca tgagggggag gtggtgctga aatattgcaa 480

gccacccctc ggttaaaaca gtttggtgcc gcttatgcca tattaccgcc ccttgtagtt 540

gggctgtttt tgcagctccg ggttagtaga gtaccatagt ggacgcggtg ttgggaatca 600

ccgccttggc tgcacactgc ttgatagagc tgcggctggt caagctaatt gtggtataac 660

cagttgattt ggcat 675

<210> 432

<211> 639

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Megirivirus C sequence"

<400> 432

ttcccgaccg gtctggcaaa ccggacggtt atcctggtta gatgtctgat ggttgctgga 60

acgtggtggc tactgctgcc accttctggc ttcctttaat gggcatctag ctgggttctt 120

tgccacaatc catcttactc tcttacccat tttctattac ccagacttgt tgttaactgg 180

taaagttgac ctactggctt cgttttgaga ctattctggt gttggtggac actctttcca 240

caagtagatt gtatggagtt catgctcgtt ttgaaccggg aatggcacaa cccgtagtag 300

gatcttgcct ctgccatact aatctgcgcc tgttgctttt agactatggg ctgctaagga 360

tgacattgga acccctttt ggatattcca tgtcaagtca actgtttcat ctggtgtacg 420

ctggaaatcc ttgttccgag gtcttgtctg gaggtggtgc tgaaatattg caagccacag 480

gcagttcctt ggacttggtg ccgctatcag atgctacacc ctctatgggc aaatgttgaa 540

ccttagtgga cgcgtgagat gggaatccac gccggccata gactggctga taagctcgcg 600

gctgatcgag ttgcaacagt aatcagttga tttgccact 639

<210> 433

<211> 497

<212> DNA

<213> Ludopivirus sp.

<400> 433

tagaccccca cctagccctt ttccccgtca gtgggggggct tactcactgg gcatctgtta 60

atctggccta actagattga caccactccc ttggaacgta actccacgct aactcactgg 120

ctctacgcac agacacacgg tctttctgct atccccgggg aagataccag atggcgaccg 180

gctgtcccag cggcctagta gctactcggg ttgagtaccc accacggttt tgacgcctgc 240

taaaattcaa gagacagagg taggggtgct tagtgtgtgg gggaagttcc cacaagcgag 300

gcaaagcatt gctccctcgc gtcaccgggt gcaaggtaaa ttggctggac ttccgctcta 360

cccttgctac tcgccctctt cggagggttc gaagtgacac taggtatacg catggttggg 420

aaaccatgcc tggcctacta ctgggtgata gcctggcggc gggtccgtct cttggcttat 480

acccgttgat ttgggat 497

<210> 434

<211> 610

<212> DNA

<213> Livupivirus sp.

<400> 434

tatctacatg gggatccagg ctgtatggaa tgtctgtctt aacaagcact ataccagaaa 60

gatccaccca aagtggtggg actgggactg tgaggtgaga aatcccgaaa ccagccttct 120

caagcgtcgg acgatctttc tgttttagtg aacaccttgc cttttaaatg gatgacaaca 180

ccccttcagc aaatcgcaat ctgaaatccc aaaagactgt ttagccgaac tctggtaatc 240

actccggaga agtaggatac gcagcccctg tggactcttg atttcaggac tcaaggtagc 300

tagagctgga acttcatgga atgacaaagg aatatatgca cattgtgcgc tttcctggcc 360

ttgtagcccg tcgtgaggat atgtcgttgg gaatcgacat cttagtccag tactgcttga 420

tagagtgtcg gctggcacag ttacctgaga ataagtcagt tgtacttaac atgaacaaaa 480

aaaataacta ccacaactac cacaatctac caatacttga attatgctga atctcgtaca 540

gtaaaaacgt tccgtggaag gacaagtatt gaagtgcggt tacatcatcc gatacgcgct 600

ggatccctca 610

<210> 435

<211> 629

<212> DNA

<213> Aichivirus A

<400> 435

cacccataca cccccacccc cttttctgta actcaagtat gtgtgctcgt aatcttgact 60

cccacggaat ggatcgatcc gctggagaac aaactgctag atccacatcc tccctccct 120

tgggaggacc tcggtcctcc cacatcctcc ctccagcctg acgtatcaca ggctgtgtga 180

agcccccgcg aaagctgctc acgtggcaat tgtgggtccc cccttcatca agacaccagg 240

tctttcctcc ttaaggctag ccccgatgtg tgaattcaca ttgggcaact agtggtgtca 300

ctgtgcgctc ccaatctcgg ccgcggagtg ctgttcccca agccaaaccc ctggcccttc 360

actatgtgcc tggcaagcat atctgagaag gtgttccgct gtggctgcca gcctggtaac 420

aggtgcccca gtgtgcgtaa ccttcttccg tctccggacg gtagtgattg gttaagattt 480

ggtgtaaggt tcatgtgcca acgccctgtg cgggatgaaa cctctactgc cctaggaatg 540

ccaggcaggt accccacctt cgggtgggat ctgagcctgg gctaattgtc tacgggtagt 600

ttcatttcca attcttttat gctggagtc 629

<210> 436

<211> 675

<212> DNA

<213> Aichivirus sp.

<400> 436

tactccattc agcttcttcg gaacctgttc ggaggaatta aacgggcacc catactcccc 60

cccacccccc ttttgtaact aagtatgtgt gctcgtgacc ttgactccca cggaacggac 120

cgatccgttg gtgaacaaac agctaggtcc acatcctcct ttcccctggg agggtccccg 180

ccctcccaca tccccccccc agcctgacgt gtcacaggct gtgtgaagcc cccgcgaaag 240

ctgctcacgt ggcaattgtg ggtcccccct tcatcaagac accaggtctt tcctccttaa 300

ggctagcccc ggcgtgtgaa ctcacgttgg gcaactagtg gtgtcactgt gcgctcccaa 360

tctcggccgc ggagtgctgt tccccaagcc aaaccccctgg cccttcacta tgtgcctggc 420

aagcacacct gagaaggtgt tccgctgtgg ctgccagcct ggtaacaggt gccccagtgt 480

gcgtaacctt cttccgtctt cggacggtgg tgattggtta agatttggtg taaggttcat 540

gtgccaacgc cctgtgcggg atgaaacctc tactgcccta ggaatgccag gcaggtaccc 600

caccttcggg tgggatctga gcctgggcta attgtctacg ggtggtttca tttccaattc 660

tttcatgtcg gagtc 675

<210> 437

<211> 675

<212> DNA

<213> Aichivirus sp.

<400> 437

tactccattc agcttcttcg gaacctgttc ggaggaatta aacgggcacc catacacccc 60

cacccccttt tctgcaactt aagtatgtgt gctcgtaatc ttgactccca cggaacggat 120

cgatccgctg gagaacaaac tgctagatcc acatcctccc ttcccctggg aggaccccgg 180

tcctcccaca tcctcccccc agcctgacgt aacacaggct gtgtgaagtc cccgcgaaag 240

ctgctcacgt ggcaattgtg ggtcccccct tcaccaagac accaggtctt tcctccttaa 300

ggctagcccc gatgtgtgaa ttcacattgg gcaactagtg gtgtcactgt gcgctcccaa 360

tctcggccgc ggagtgctgt tccccaagcc aaaccccctgg cccttcacta tgtgcctggc 420

aagcatatct gagaaggtgt tccgctgtgg ctgccagcct ggtaacaggt gccccagtgt 480

gcgtaacctt cttccgtctc cggacggtag tgattggtta agatttggtg taaggttcat 540

gtgccaacgc cctgtgcggg atgaaacctc tactgcccta ggaatgccag gcaggtaccc 600

caccttcggg tgggatctga gcctgggcta attgtctacg ggtagtttca tttccaattc 660

ttttatgtcg gagtc 675

<210> 438

<211> 610

<212> DNA

<213> Kobuvirus sp.

<400> 438

gtaacttcaa gtgtgtgtgc tcgtaatctt gactcctgcc ggaatgccgc ccggttcagt 60

gaacaaacag ctaggcaagt ccctcccttc ccctgtggtc ggttctcacc ggccaccatc 120

cctcccccag cctgacgtgt tacaggctgt gcaaagcccc cgcgaaagct gctcacgtgg 180

caattgtggg tccccccttt gtcaagacac cgagtctttc tcccttaagg ctagcccggt 240

cccacgaacg tggaactggc aactagtggt gtcactacac gcctccgacc tcggacgcgg 300

agtgctgttc cccaagctgt aaccctgacc caagactgtg ctgcctggca agcaccgtct 360

gggaagatgt tccgctgtgg ctgccaaacc tggtaacagg tgccccagtg tgtgtagtct 420

tcctccagtc tccggactgg cagtcttgtg taaagatgca gtgtaaggtt caagtgccaa 480

atccctggaa ggagtgaccc tctactgccc taggaatgct gtgcaggtac ccccaacttc 540

ggttggggat ctgagcacag gctaattgtc tacgggtagt ttcatttccc atcctctctt 600

ttttggcatc 610

<210> 439

<211> 576

<212> DNA

<213> Kobuvirus sp.

<400> 439

tttgaaaagg gggtgggggg gcctcggccc cctcaccctc ttttccggtg gccacccgcc 60

cgggccaccg ttactccact ccactccttc gggactggtt tggaggaaca taacagggct 120

tcccatccct gtttaccctt actccactca cccctcccct tgaccaaccc tatccacacc 180

ccactgactg actcctttgg atcttgacct cggaatgcct acttgacctc ccacttgcct 240

ctccctttc ggattgccgg tggtgcctgg cggaaaaagc acaagtgtgt tgttggctac 300

caaactccta cccgacaaag gtgcgtgtcc gcgtgctgag taatgggata ggagatgcca 360

ataacaggct cgcccatgag tagagcatgg actgcggtgc atgtgacttc ggtcaccagg 420

ggcatagcat tgctcacccc tgaatcaagt catcgagatt tctctgacct ctgaagtgca 480

ctgtggttgc gtggctggga atccacgctt gaccatgtac tgcttgatag agtcgcggct 540

ggccgactca tgggttaaag tcagttgaca agacac 576

<210> 440

<211> 513

<212> DNA

<213> Kobuvirus sp.

<400> 440

ccaccgttac ttcactccac tccctcggga ctggtttgga ggagcataac agggcttccc 60

atccctgttc accctcaata ccacccaccc tttccctcaa ccatccctat ccacacccca 120

ctgactgatt cccttggatt ttgacctcag aacgcctact tgacctccca cttgcctttc 180

ccttctcgga ttgccggtgg tgcctggcgg aaaaagcaca agtgtgttgc aggctaccaa 240

actcctaccc gacaaaggta cgtgtccgcg tgctgagtaa tgggatagga gatgcctaca 300

acaggctcgc ccatgagtag agcatggact gcggtgcatg tgacttcggt caccacgggc 360

atagcattgc tcacccgtga atcaagtcat tgagattcct ctgacctctg aagtgcactg 420

tggttgcgtg gctgggaatc cacgcttgac catgtactgc ttgatagagt cgcggctggc 480

cgactcatgg gttaaagtca gttgataaga cac 513

<210> 441

<211> 675

<212> DNA

<213> Kobuvirus sp.

<400> 441

gggggtgggg ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc 60

gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt 120

tcccattcct tccccctttt cccaacccca accgccgtat ctggtggcgg caagacacac 180

gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg 240

ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg tcgtcctggc aagactatga 300

cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt tccccagtgt gtgtagtgcg 360

atcttccagg tcctcctggt tggcgttgtc cagaaactgc ttcaggtaag tggggtgtgc 420

ccaatcccta caaaggttga ttctttcacc accttaggaa tgctccggag gtaccccagc 480

aacagctggg atctgaccgg aggctaattg tctacgggtg gtgtttcctt tttcttttca 540

cacaactcta ctgctgacaa ctcactgact atccacttgc tctcttgtgc ctttctgctc 600

tggttcaagt tccttgattg tttttgactg cttttcactg cttttcttct cacaatcctt 660

gctcagttca aagtc 675

<210> 442

<211> 675

<212> DNA

<213> Kobuvirus sp.

<400> 442

gggctataaa tatgggcatt cctcttcccc cttccccttt tgaagatgag tgcgcatatt 60

cttgactccg cctggattgg ccgcccaagg cgtgaacaag cagctaggcc accatgacac 120

tgcggtggtg tccgaacccg cgggtgcctt cacgggcacc tgtggtatgt aggactccca 180

ccgtggtctt ccctttcccc ctcaatcttt ccccctggtt cgactaacgg gaccagtgct 240

ggaacctgtc cggtgaacgg tatagcaggc ccccccggca gaaacacccg gtgcttaccc 300

cttaaggcta gccccccttcc atgaatttgg ttggggcaac tagtgggtgt acagttggcg 360

tgaaccctcc ggtctaggag tgctcttgcc caatcctctg tgtgtgcctt gcagtaggga 420

ctggcaatcc ttcgcgtagg tgatccgctg tgccatgcca tcctggcgac aggaggccca 480

gtgtgcgcaa cctacgtccc ttctgggtgc tgcattgcat tacctttgga gtaagcttgg 540

tgtgccgaaa ccccagggtt tacgtaccac tcgtggtgtg aggaatgtgc cgcaggtacc 600

ccatccttga ggtgggatct gagcggtagc taattgtcta gcaccacttt cttccttttt 660

tctttgctgg tcacg 675

<210> 443

<211> 604

<212> DNA

<213> Aalivirus sp.

<400> 443

ttgaaagggg gtgctcaggg tagctccctg agctcttccc tccaccctct ttcaacgtct 60

ggcccacgat acgggccacc tttcaatctt aactaactat ccctttaatc tatttggatt 120

ttctggttta gaataatttg gaacacataa ttggattatc ttttaggatt gtggatagga 180

tttgttcggg atatcactcc cttcctgtgc taacacatat tctaattccc tcctttgtct 240

attatctctt ggaggtggtg ctgaaatatt gcaagccact tgagtgtata gatgaagtag 300

gctcaagatg aatgttgtgt tactcaaggc aagtgtagct atcactaaga tattggtaac 360

gtgaaacgga ttaccggtag tagcgtgatc ttccgtctta gtgctctagt gactagagga 420

caacgacatg gcatcacata tcttaaccct ccagttttgg catccgggac agaatgggct 480

ggatatccgc tttctttctg gggtatgtga tgggtggtat tggggtaacc accttgacca 540

tgacgctcga taagagtgac cgcctgatca ttgaaacctc tagtataaaa ttcaggctga 600

aatc 604

<210> 444

<211> 445

<212> DNA

<213> Grusopivirus A

<400> 444

tgcctgagta ggattgtgaa tttaggtatg agagggttag ccaacccatt ctgaaccata 60

atagatacgt caatctgaat ccatctaaat ctatctctta ggcagtggtg ctgaaatatt 120

gcaagctact agggatagac gtgatctgat tcaagaacct atctaatgtg gtgatgagaa 180

ggctaggttt atccatagta atcccttgtt ctgaacaggc aatgcacatg ctctagtagg 240

atctcgggct ctgcgattgg ctctaaaccg accaatccag gtagaggcac taagtgtagg 300

acttgccaaa atgtattaca tgctggtacc gactcactag tctggaaact ccacactgaa 360

agtgactggg gggggcccca tcacatttgt gctactgctt gatagagttg cggctggtca 420

acttggattg gtataaccag ttgaa 445

<210> 445

<211> 537

<212> DNA

<213> Grusopivirus B

<400> 445

gccatccgta ggttctggta aggttccatc aactgttggg gcgctagttg ctatgaccgc 60

attcacggac ggatgattta tagtatcacc caatccgggc acaacttctt tagccacttc 120

ttccacatta ctaagggctc tcttgccgag tttcaacgtc tagtccacga cacggacctt 180

cctacttttc tatcttctta ttttctctac taaattggta tctggtactg aagatatgcg 240

gattgtgatt ttgtgcctgt ctaaactaac cctattctag ggttaggtgg gtaccatata 300

ctaatggtga acaggattac ctatgtatcc attagtccct atggatctgg cgacccacaa 360

actcatgttc atagagaggc taagctgagt gctcgccgaa taagcattgc ttcaggtgcc 420

gactattgtc tggaaaccac tcagtgatag ctataggggg gggccccgta gcatctgcct 480

tactgcctga tagggtggcg gctggtccat gaacatgcag taaccagttg acttgac 537

<210> 446

<211> 562

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Yancheng osbecks grenadier anchovy picornavirus sequence"

<400> 446

ccttagggtc tggaatgcgt cctttctggg cacttccaca atcctaaggt aattttcaac 60

gccagcgatg gagcgatatc caaagaacct ttatgtttta gttcgtcttg tatgtttata 120

aaatataaaa ttgggattag caacccacaa aaacattttt gttattctca ccacatgaga 180

gggtggttaa acctcttcgt aacctatctt gcttgattgg ctacttgggc tagatttagg 240

acaacctctt tagcaagcct aaattactcc tcgtacttgc acctggtaac aggcgtacct 300

ggaggttacg gtggcgctaa cttggacttc tcgttaattc gtgcaagttt aaatgatgcc 360

tattttgaat acaagaaagt atgatagtaa cttagggcgt gaagttccgc ttaacataag 420

gcagtataag tactaagata aggtgtaaga cctaccttaa taactgttgt cttttctcat 480

ggtctttccg tgggagccct tgctaggggt aaagttaagt attctcaaat aatttttcat 540

tcaaactctt tctctctgtt tt 562

<210> 447

<211> 675

<212> DNA

<213> Gallivirus sp.

<400> 447

ccactcgcac ttcctggata gtgcgttaga tatgcccgac atatcgcttc caggaccaaa 60

gccccccttt tctctttccc aaccagcttc gccactcaag ctgtaattcc atgtccggtc 120

tttccggcct tagtatcatg gaaatgtggt cgtgctcaaa tgaaattgag ttgacattga 180

tcaatgaaag ttgcactgaa ctttgctaaa ctggctagcg ccacctggtg tgtgccgttg 240

gtctcctcac atggtaacat gtgccaacgg gcccgaaagg ctagtgggca attaccgctc 300

caagggaggg gtacccaccc cgacctgaac agcggtaatg aagctcacct cccaggctct 360

gaccccgaga agtttagtta tttagtaggt gtaattagta cttgtgattg gtcaatttga 420

tagtagtttg aaacgttatg gatgaatgag tagacccccc gaaggtaccc cattacatgg 480

gatctgatca gggccacatt ctgcgtgtct ccccgcactt gtggttaaaa ccatgaaagt 540

tcatcccaaa caatcttttc ctcttctttt tcttttagtg gtgacaacct actggattgg 600

tgattaccaa tctgtactag tgttgtatta agacttgttg tgtggagaaa atggactctt 660

tcaagaagat ttttg 675

<210> 448

<211> 673

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Falcovirus A1 sequence"

<400> 448

taaaagggga cgcggtgtgg cagctttggc tgtcatgccg tgttctcctt ttaccccaag 60

gactagcctt gggggttttc caaattcctt tccctgtagg ctttacttct ctttatctat 120

cttttctgta actaagtttt gcctattcta aaaatatttt agaatgtgtt tggatgtaac 180

taagtttgtg cctgccctaa aaatatttta gggcttgttt ggataacctc gtcccttgtg 240

ttcagtgccg cacaatttgc taggcactgt tcacttcctt tgtttgtcca ttatgtatgc 300

taaggtatga attccatcat atgcttagcc tctacatgca taatcttatt ccctccctgg 360

tgcaaactac gcccccaaca tatgtgaatc ttttaagcat attcctgacc ccacacatat 420

atatgtgttc tcgtgaattc ccccaccgtg aggtggtcac ttggacgtgg tgtgtgtcac 480

acagcatata tatgatgcag gatgttgttt ttaagataag catatgtcct tagtgctttg 540

catcatttcc tccacacccc gtgaatgcgg ctaatcttaa ccctgttggg tccgtgggta 600

aaccaaccca ttaaccacag gacggaaccg actactttcg ggagtgtgtg tttctttttc 660

ttcttttgtc act 673

<210> 449

<211> 665

<212> DNA

<213> Tremovirus B

<400> 449

ttcaaatggc ccctgggttg atacccagtg gtcatttgga cactttggta aggaggtgta 60

attatccttc ccatgtggaa cctagtgctt aggtttactt tatatgttct ttgtttgtcc 120

tttgtacttt ctatcgggca atcttgttgt tcaatacaat atgtatttga actgcctaag 180

ataaattcag ttttcaacca acccctctct tggggttgtg tctttctttc tttcttatat 240

cctcttaagc tgacttactt gctaatccga ctcctcgtca acgggagggt aaagcagtat 300

cactagggta ttgtgatgta ggagaaaaag taagtagaga tagtgcatgt aacgaaagtg 360

acttggtact ttaaactctc ttaatcccaa agtgtggtat tggtcatgtt ggagtaggct 420

acgggtgaaa ctccttcaca tttagtaatg tgttcacacg ctaacgctac ggtagatgac 480

agactaggtc ttattctcaa cgtaggggga cgggtgtatg ttcatgatta gccacatatt 540

aaggttttga ggggctgagt catataagta tgtgcattaa tttctggtac tggtccctgg 600

ggactggccc ttttctaggt tgattttagt ttccccaatt tttaaaaact aatgagattt 660

acgac 665

<210> 450

<211> 675

<212> DNA

<213> Hepatovirus A

<400> 450

tctttggtct ggggaactaa aataccagac ccgcgtttgc ctagcgatat aggctttaat 60

tgttgtttgt cattgtgcgt ttgatatgtg ttttaatgta aataataatt ctagcaggtt 120

ctagttcttg atcatgtcct ctttaaggca ctcatttcaa cttgctatct ttcttttctt 180

ccttggttct ccctacacca aatgcactgg ccgctgcgcc cggcggggtc aaccacatga 240

ttagcatgtg gctgtaggtg ttgaaggctg ggacatgaac atcaatggaa tagtgcgcat 300

gcttactggg gtccattgaa gtagtgggat ctttctattg gggtaggcta cgggtgaaac 360

cccttaggtt aatactcata ttgagagata ccttggatag gttaactgtg ctggatatgg 420

ttgagtttaa cgacaaaaag ccatcaacag ctgtggacag aacctcatcc ttagattgct 480

cactatggat atgtgctctg ggcgtgtttc ttgcatgatg gccattggtc aattcatgcc 540

tgggccaatg taggattagc cttaaattac tttttaaaag tagcctcatt tagctggact 600

aatggtgggg cgtatgatcc tgcatttggc ctctggggta atcaggggca tttaggtttc 660

cacataatag caaat 675

<210> 451

<211> 675

<212> DNA

<213> Hepatovirus sp.

<400> 451

gcaaggggtg gttttaacct tgcacgcgtt taccgtgcgt taacggtttt ccatgtttgt 60

atgtcttgtt tgtattatgt gttttgtaaa tattaattcc tgcaggttca gggttcttta 120

atcatgttgg gctgtaccca cactcaactt ttggccataa gtgagtttct taacgaacct 180

tttaacacag gatgttatta gggcccaata ttttccctga ggccttcttt ggcctctatt 240

ttttcccctt ttctatctcc ttgtattccg ggctcacgtg atgccaatgg actgacccat 300

gcgcccgtgg gggttaacta ctggagtagc cagtagctgt aggtgctaaa agtcacgtac 360

gtgtaagact ggacgagacc tctcagctat aactgaaagt agtaagtatg tctgaacttc 420

ttgaaggggt aggctacggg tgaaacccct taggttaata ctcatattga gagatacctc 480

tgataggtga aggtttccgg tagaggtgag tttaacgaca aagcctctca acggatgtgg 540

gcccacctca tcagcaagat gctttcatac ccaataccgt aggggctggg ttgttgagtt 600

cagtcccaag cgtccctccc gcaaggttgt aggggtactc aggggcattt aggtttccac 660

aattaaacaa ataca 675

<210> 452

<211> 675

<212> DNA

<213> Hepatovirus D

<400> 452

ctttggatgc ccatagtgcg ggggtataaa taccgcactc cctttagctg ttccgagggt 60

atcggaacct atatgtttgt tttctgtctg tctgtcagct ttatgtgtgc tcgtcccctt 120

tagggcactc atttcagctt gctttcattc ttttcttccc cggttctcac cttaccggag 180

gcactggccg ttgcgcccgg cggggtcaac ctagtgatta gcactaggct gtaggtgtct 240

aaagtggtga cattaagact tggtaactga tttcagcact gttaactgat gttggggatg 300

acttgattga tcttctggaa ggggtaggct acgggtgaaa ccccttatct taataccact 360

atgtagagat agattcagta ggttaagggc agtggataag gttgagttca ttttggacaa 420

taaaccttca acactggtgg acccaatctc actgaccaga tgctttcttg actgatcctt 480

cagaggggtg attcttctga ataggttgcc ttgacactga tgcctgagac ccattgggtc 540

gggccttaaa tcatggaact ccactggact ttcatggcct agcttctgcc ttagacagac 600

tctggggccc cacgaccctc tgggcccttc ggggtactca ggggcattta ggtttttcca 660

caattaaaag agtta 675

<210> 453

<211> 560

<212> DNA

<213> Hepatovirus sp.

<400> 453

gtcatgtttc tctttaagaa cactcaattt tggccataag tgagactctt gtcgaacctt 60

tcatgtcagg accatgttag ggccattatc cttttccctg gggcattctt cttgcccctg 120

tttcatcttt ctatcatctt tcttccgggc tctcacaatg ccaatggagc gaccgatgcg 180

cacgtcgggg ttaacccatg gattagccat gggctgtagc tgctaaaagt tgtgactcct 240

gaagcatact atcaatggta gtagatgtaa ctgaaacact gaagcttctc tgatcttgaa 300

agaagggtag gctacgggtg aaacccttca ggttaatact catattgaga gataccttg 360

gtaggttaac gttggcggat aatgttgagt ttaacgacaa taaacattca acgcctgtgg 420

gcgaacctca ccaatttcat gctttgaagt gaatgtgcgt agggtctcta tcggagatgc 480

tatgtggatg gtgccctccc tggaaacagg ttgtaggggt actcaggtgc acttaggttt 540

ccacatttta aagatttttc 560

<210> 454

<211> 675

<212> DNA

<213> Hepatovirus I

<400> 454

ggctgcctgt gtctcagggg taagtactgg ggccgcgttg accgtgcggt acggttatgc 60

ttttagatta ggatgtccgt ctgtccggca ctctcttttg cttaaaatgg ccttaaatcc 120

atgggaggcg taaccatggg ccctttgtta cctagacatg attgcattgg gggccgtcct 180

tggggcttag gccccagcca tttctcttga ctcgtctaag agtttacttc atccttttct 240

ttactttatt ttccaggctc tcagcatgcc gacggctctg accactgcgc ccggtggggt 300

taactgcatg attagcatgc agctgtagga gttaaaagtg ctgacaggcc aattctgacg 360

taagtccact ctatattaac ttgatcaagt aaggttgatt gatctttgtg agagggtagg 420

ctacgggtga aaccctctag gttaatactc atattgagag atacctccag aaggtgaagg 480

ttggcggata ttggtgagtt cttttaggac aaaaaccttt caacgcctgt gggcccacct 540

cactggcaca atgctttcat ccccaattgt gatgggtagt ttggactgaa atcaggagta 600

acctgcccta cgagtttagg ggtagttcag gggtatttag gcttccacat ttgatagagt 660

ttatgagagt gagcc 675

<210> 455

<211> 527

<212> DNA

<213> Hepatovirus C

<400> 455

ttcaaaagcc ccagcggggt ttcattaccc cgctgtggct tttggacttc cctaggatgg 60

ggaagtaaat taccatcctc gcgtttgccg tgcgttaacg gctacttttc ttctagctgt 120

agaagtaaaa ttcagcatgt tttatgtttg tttgtcttgt ttgttatata cattttaca 180

ctcctacaaa tgcacatgaa gaacagtttg tagagattaa caaacgctta gctgaaccta 240

ggtggtgaat ctagtagtaa gataagtaga ggaagctata ccttaagttg gttgggccct 300

cgtgtttgct ctataaacaa aaccaagtga gtagagtgga tgaacagtac taaatccctg 360

agtacaggga acctcacagg tgtgatacac ttatgtctat gtgacctggt tggaggttgg 420

gcgtgcccta tgatactgga gtgggagatc ttttggggaa cccacgtttt cacactgcct 480

gatagggtct tgccgagaga ctcacttgtt tcggctgtac ttgtaac 527

<210> 456

<211> 665

<212> DNA

<213> Fipivirus A

<400> 456

tgcgggtaaa ctcccgcatg tgtgaatgag gcgatgtccc aggaactaac tgccgatcct 60

ggttttaact acgatccgta tttgttacta atgcgatatc cccccattgt ttgcctccat 120

gttgttttca acgcttttgg ccttgagtgt tatcaagtgt tttagcgaca tagtgggaag 180

ctacggctgc gtccccattt ttgagtggcg acccagtttt agtggccact ctgtccctga 240

actgcgctat aatgtgaatt tatgttcaca aaaacggact gatgtaactg ttaatgacta 300

aggaatagta cctcactgaa gtatcaagac cccgttcgag cggtgtacat atatggatgg 360

aaaccttgtc tgagtcatct cgaatactaa tcaatgaggg atgtcgagta agcatatcat 420

gaaccacata gaatagtggg gtttcggggt tagaggctct ctgcagcaat gtatctctaa 480

caccatggcc gaaatgagag atagagacca cgatgtttgt gtgtaagtaa tgatgtgtgg 540

aaagaaaatt ctgaatgttg gtatgatatc agtctaaggg gagtggctca cctaagagct 600

acccaaacat ttcacagcag acaacataac gtactgagag tagttggaag gttccagaaa 660

tcagt 665

<210> 457

<211> 467

<212> DNA

<213> Fipivirus C

<400> 457

cgcggttaaa cccgcgcaac cttctttcag ccgcgtctga gtagcgcggt tagtcctgat 60

acacagtttc ctgttgggta ctgtgtcttc gggtgaatgc tcttgtgtga atgttttagg 120

ctgtttaagg gaagcgtttc cccgtgcgct gtgagggttt ctcacgctct ttcggggtgc 180

agtctcttct gttgttcatt aagatgtatg gatgcactgt tgtgaaggat ttgtgaactg 240

gggatcgaca ccccgtgagg ggtgccccag tgtccatagg agtttgctgg agaggtgtgt 300

tgctgtagtg actatccgtg acctggcatt ctaaggtgtt gaccccaacc tgtgagggtc 360

tggatcgcag tgttgaagtg ctttggaggg ttcaatgggg tttctgtagt ggatattatg 420

tgcttgacga ctactggtac gagtgtattg ggggtctaca tgtgtga 467

<210> 458

<211> 657

<212> DNA

<213> Fipivirus E

<400> 458

ctcttccgat cttgggggtt cgcccccatg tctcatttca actagccgtg tgtctagtta 60

acgcaccgcc tcaccctggt cgttatcggg tcggttcttg cgaccgttag atcgtgagcg 120

tttctgagga tcagttcgta taagttctcc ggtgtggcga ccgtaaaatc gtcacgtccc 180

atgcaataga tgacgttaaa ctcgtttgcc agttacataa aggaatgttg ttacttttaa 240

attgtctgtt acatttaaca tcttgccagt atgatgctac tgtacactac gggtgtagga 300

accttgtagt gtgacgtatc actcatatgt ggatgggtgc tccagacctt tatggaagct 360

ctcagttagt agtgatcctt gacttcattg agccctggta acagtggaag tcaagatgta 420

tatgttgctc aacacacttc ggtgctacga agctgtttgt ggaagtactg gcgaggttca 480

ttctgaatca tatgtttgtc acatagtcag ggagtgccgt cgcttacgac ggacccttt 540

tctttataat tacaaatctg tgtctcaagt gttgttggct ggttttcttc ttctgttttc 600

attgttcata tatatacgtc agagtgaaag actcggtata tacaaaactg atccaga 657

<210> 459

<211> 506

<212> DNA

<213> Aquamavirus sp.

<400> 459

ttcaaaggtg gcgggagagt tggcctcacg ctgtttagcg tgagagctgg ctctcctgcc 60

ccttcccctg agccggggat cttggctcat tcccctcttt tctatcctcc ctcattggac 120

tttacggatg acccggcata aacttgacaa ccgatgttgg atttcccttg tggctgtgat 180

ggaggacata ccctcgggtg tagttgtgtg cgtgtcgctc tgcgactcga gcttcaaagt 240

ggtgctgaaa tattgcaagc gtcgttgctc gattaacgga gtggtacaat cctatgaacc 300

caagtgcatt catgcgaaag ccccggaggg gtgagtagca tggactcgaa tcagaagagc 360

tggagctcgc ttggtacggc acgtagcatt gctttgccta aagaccaagg gggtatggct 420

ataggtgggg gcctatagct tgtccagtgc tggttgacag actcgtgcta cgcgtctggt 480

tcgagtataa gtagctgcaa ctcact 506

<210> 460

<211> 548

<212> DNA

<213> Avisivirus A

<400> 460

ttcactcgct ttcccccccc ctctataggg gcggtctttt aattcttatt aatttcctac 60

tttactatca aatttcttct aagtagggac tgaggtcact tagccctccc tctcctgggc 120

tttccagggt tatagaggtt ctaaagctaa gccatgtgtc ttgagctaca cttagtacaa 180

aggtttagta atgattgtac atgccagtaa ccttctagtg cccatggatt aaagagtggt 240

aacactctcc atggggcccg aaaggctagt gggcatagtt ggcatcaagg aaggggtccc 300

caccccaacc tgaattgctg gctagaagct caccttagaa gaagtgctgg gtgacaacgt 360

gtccaatcgt gaacgactga tggaaacgtg tggagatgga tatgtggggg ttcactgagt 420

agatgccctg aaggagaaat ctgatcaggg gcccgtgact atacgctagg taaaccgggt 480

ataaaaacca tgaaaggtgg cccaaaatct cttcctttta ttttatttct atgttggtga 540

cagtcaag 548

<210> 461

<211> 492

<212> DNA

<213> Avisivirus B

<400> 461

caccccctac tgccctaacc cccaaagtta gttatagggt ggctccctac ccttactcca 60

cggggtaagc cctaacccgg ttgaatctca agatcagcct tagcgaggac tattagtacc 120

gctcaaaccc tttgcctgta gtgcccaggg gtcacagagg ggtgaccctc tccctggggc 180

ccaaaaggct aggtggcaag acagggtcca agtgaggggc tactctaagt agccccaagc 240

tgaacatcct gtctgaagcc acccttgcag ggccaggttt gattggggaa actagacacc 300

agctttgtcc tgggattggg gggatatcga gttagtccag gaggtgcgag tagatgcccc 360

cgaaggtacc ccaggcacat ctgggatctg atcgggggcc cgtgactata caataggtaa 420

accgggttaa aaaacatgaa agcgcctctc tctttcctac ttcttttatt gactggtgac 480

aaaaatagca gt 492

<210> 462

<211> 517

<212> DNA

<213> Crohivirus A

<400> 462

gttgaagtcc atttcttgct tgcccccgat gaatcctgtt aaggcctcac ggccctaagg 60

gtgaaactcg gttatcccct cctgtacttc gagaagatta gtacaacact atgaaatcta 120

catcttgtga tccgggataa ccccaatccc agaaacctgt gatgggcgtc accacccctc 180

ttatggtaac ataagggtgt cgccgcgttg gcacaggacc ctttgggctg gatgttttta 240

gtaatggtgt cgaaggtcct attgagctac aggagtttcc tccgccctgg tgaatgcggc 300

taatcttatc cctgagccta aggttgcgat ccagcaactt gatggtcgta atgcgtaagt 360

tgggggcgga accgactact ttccagaagg cgtgtttctt tgttttgtct gttactatgg 420

tgcatgatat agatattgaa tatttgatct ttttgagctg tttcttatct tattgctaca 480

tcctttcagg tgttggattt acattttggt taataag 517

<210> 463

<211> 393

<212> DNA

<213> Kunsagivirus B

<400> 463

gattttctgg ttatcccttt tggacttggt aggggcccac gtgcccaccc acctctgtgt 60

gtgttgattt ctaatcgatg cctggcagtg gcggccacct ctccttactg gtaaacctcc 120

ggtgagtgaa gttgtcaagc tacaggtacc gtgcaggatg aaatgcgcac atgtgaacaa 180

actaggagtc atacaccggg tcaaactctg gaaacggagt ccgggactct gaccttggtt 240

gggtgagctc gaggcatcac attgatggac gcgattcgct atccttccct agtaggacct 300

tgtggtgtac ccctggttgg gaatccaggg ctggtcgggt gcagggtgac agcctgttct 360

ccacctcaac cattgtagga gaaatcaacc cct 393

<210> 464

<211> 675

<212> DNA

<213> Limnipivirus A

<400> 464

ttctttggat atccatttaa cgtgtaccct atacgataat tggggtggat tctggatgcc 60

tagttccagt gattggttaa gaactcgttt actacgtata gtatgattag caaagtgctc 120

gattgatcac gtaatgatct atgtggttaa aaacccagta gtatggtata tactcagtag 180

tgtacactgt gagtacaact cttggcgtag agagaacaat tcacccgaat ccgtggcgta 240

tccatggaaa taagtttacc taattgtatg ttacaaggca tatgagacat ttatgagata 300

tggtttattt tgactaaacg agtgtagagg tggtggagtc tatccaactt caagccatgc 360

aattgttgtg ttgattgata tcattgacca tttttgtgga ttgtgtacac atacaatttg 420

aaaattaacc ccctcaagaa taagacatgg gaccattcgt ggtagatacc gtgctcggat 480

gcttgagatt agatgggtta gactagtttt ggaatgagat tgccgagaaa gtcccgctag 540

acatgtttta caagtcgtgg tattccgcta gacttttcg cagacacatg gaagggtcca 600

tgtgttgtgc aattgcaggg tgacagccca actgcagagt tttccttact agaataaaaa 660

tctgttgtca atttt 675

<210> 465

<211> 501

<212> DNA

<213> Limnipivirus C

<400> 465

gtttctgagc actggtaaga gcttagacaa acgtttttaa aatttatttt ctctgcaact 60

tttgtttgtg tttattttta tttgttaatt ttgcgcctaa gcatttgttg cgaagtattt 120

gattcattag taatattact tattgtttat ttagatggta ttcaaagtgg tgggagtatc 180

gaacccaagc gtcgtatgct atctccttga acaattttta atcattgcga agtgatcatt 240

gaaaaggata ggtgtttaag aactcaaaga gtgttaataa tgttgggtga caggtgtccc 300

catagaattt attaacatga tttggactgg ttatctagta agaagaacca tcgaacgcac 360

gagcgagcat tgcttgcggg gcagttaccc tgcgtcgatg taagtgtgta ccgggggggtg 420

cacatgttga ttctttatgg cctgataggg tgcgtcattc gcgcctagat aattagtata 480

atgcgaatgg aataaattta c 501

<210> 466

<211> 675

<212> DNA

<213> Orivirus sp.

<400> 466

ggtcccaggc caatattctt cgtaaggctt ggttccaatt ttccaccact cgtgtttggg 60

ttctggccta tggtacccag aggggcggtt tgggggaatt aactcccccc cccctgtggt 120

cctataccac cccacacctc tgtgggcttt ctttactatc ttcttgtttt ccgactttta 180

aacactaggc aggcgcgcct agtcatacac cgcccggctg gtctttccag ctcttgtggg 240

cggtgcgcgc tggtccatcg tgcccagcga catagcacct tgtggacacc tccgaacgcc 300

ctcccctgta tggggtggtg cccaggggtt tcagtgtggt gacacactcc ctggggcccg 360

aaaggctagt gtgcaacagg tgaggtacag ccagctgccc ccgtggctgg agggaccaag 420

cttgtgaagc acacctcacc ttcttggggg tgggctagta agtggtgaaa gcatagtgtc 480

cgtgtcgctg gccaacactt tgggtcaagt ccagccactc agtgagtaga tgcccaggag 540

gtacccctag tggatctgac ttggggcctg ttacttaatg caggttaaaa actatgaaag 600

ctgagtagtg tagcccggct ggtggcttct cttccttatt cattctattt tatggtgaca 660

aacgcaactg aagcc 675

<210> 467

<211> 675

<212> DNA

<213> Hepatovirus A

<400> 467

cttgatacct caccgccgtt tgcctaggct ataggctaaa tttccctttc cctgtccttt 60

ccctatttcc ttttgttttg tttgtaaata ttaattcctg caggttcagg gttctttaat 120

ctgtttctct ataagaacac tcaattttca cgctttctgt ctcctttctt ccagggctct 180

ccccttgccc taggctctgg ccgttgcgcc cggcggggtc aactccatga ttagcatgga 240

gctgtaggag tctaaattgg ggacgcagat gtttgggacg tcgccttgca gtgttaactt 300

ggctttcatg aacctctttg atcttccaca aggggtaggc tacgggtgaa acctcttagg 360

ctaatacttc tatgaagaga tgccttggat agggtaacag cggcggatat tggtgagttg 420

ttaagacaaa aaccattcaa cgccgaagga ctggctctca tccagtggat gcattgaggg 480

aattgattgt cagggctgtc tctaggttta atctcagacc tctctgtgct tagggcaaac 540

actatttggc cttaaatggg atcctgtgag agggggtccc tccattgaca gctggactgt 600

tctttggggc cttatgtggt gtttgcctct gaggtactca ggggcattta ggtttttcct 660

cattcttaaa caata 675

<210> 468

<211> 675

<212> DNA

<213> Hepatovirus A

<400> 468

cgccgtttgc ctaggctata ggctaaattt tccctttccc ttttcccttt cctattccct 60

ttgttttgct tgtaaatatt gatttgtaaa tattgattcc tgcaggttca gggttcttaa 120

atctgtttct ctataagaac actcatttca cgctttctgt cttctttctt ccagggctct 180

ccccttgccc taggctctgg ccgttgcgcc cggcggggtc aactccatga ttagcatgga 240

gctgtaggag tctaaattgg ggacacagat gtttggaacg tcaccttgca gtgttaactt 300

ggctttcatg aatctctttg atcttccaca aggggtaggc tacgggtgaa acctcttagg 360

ctaatacttc tatgaagaga tgccttggat agggtaacag cggcggatat tggtgagttg 420

ttaagacaaa aaccattcaa cgccggagga ctgactctca tccagtggat gcattgagtg 480

gattgactgt cggggctgtc tttaggctta attccagacc tctctgtgct tggggcaaac 540

atcatttggc cttaaatggg attctgtgag aggggatccc tccattgcca gctggactgt 600

tctttggggc cttatgtggt gtttgccgct gaggtactca ggggcattta ggtttttcct 660

cattcttaaa taata 675

<210> 469

<211> 519

<212> DNA

<213> Parechovirus F

<400> 469

ggtcggggag atgtgttcat gatcggttaa caccatcatg gatcatctct ccccgacctc 60

tttttgaccc agctatgggt taaatagtac ttttcttttc tcttttgctt tcttttgtgt 120

tttgtttgtt ttgcaacata taacaagcat tttatcagta ttagtgtctg caactgtata 180

acaagcaagg tggagcaatc atgcgagtat atctcaattg aattgtgaca cacaagtgtg 240

cactatgtgg aataaatgcc attttggcca aacctggtta gccagaccag tagtaggaca 300

atttggcacc cttagtgggc gcgacctaga tgctagggat gagcaaacct atttcccctg 360

agtacagggg ctctccttca cctctacatt ttggacctct ttttgagtat cctcgataga 420

aggtgaagtg acggtgtacc ggatggttaa ttgatctcat tgctgggtga cagcccgcta 480

ggaccaggca gcatctttgt atggacctgt acatgtaac 519

<210> 470

<211> 343

<212> DNA

<213> Parechovirus D

<400> 470

acatggggca ggtgtgctgt gccaagagca acactacggt ggccgagccg atggttcgtc 60

accacgtagt aggactccgt agtgcttggt tacggcggac gtaagtcagt tgagtgatgt 120

ctaagtggca aaccatgagt acatggtaac cttgtgtgga ctcgcgggac ggaatttcct 180

atcccattga ctccttgtag caaggtgggt atacccaacc acaatggcag caccctgggt 240

gggaacccag gggcctggat tagtatccag tcacacagcc tgatagggtg gcggctcagc 300

cactgaccag cgtctctaaa taattgtgag ctgttcatgc acc 343

<210> 471

<211> 675

<212> DNA

<213> Parechovirus C

<400> 471

cggtcatccc cctttcccca cagccggtgt gggttctaat cggctcctac taaacaccta 60

agcatcactg cgcctctatc tctcctatcc acaggtctaa gacgcttgga ataagacatg 120

tgggtgcaat aggaagatta gctagtccaa tctctccttc cagctacgct tctcccttcg 180

atgagcgtag ggggggcccc cacctccctc atctctggat agggctcttg ctacggggct 240

ttcccgtctg gaccagcagg cccactggtg cgcttccatt caagtttagt gtgcattact 300

gtctgaaata ttgctttgct aggatctagt gtagcgacct gcatattgcc agcggacttc 360

cccacatggt aacatgtgcc tctgggccca aaaggcatgt ctttgaccgt atgcagtaca 420

accccagtat aggtcctttc tatggcagta tggatctcag tgatgagtct atacagaata 480

tggaagtggt tcggatatgt cagcccgaag gatgcccaga aggtacccgc agataacctt 540

aagagactgt ggatctgatc tggggcccac caccttcggg tgggtagaag ctaaccatgc 600

cttgggttaa aaaacgtcta agggctgacc agacccgggg gatccgggtt ttccctatct 660

tgacctactc taatc 675

<210> 472

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Ljungan Virus 87-012 sequence"

<400> 472

ctcattgccc acacctggtt ggttcccagg ttcatacaat aaccatcaat aaacttttaa 60

catctaagat agtattatcc catactagac tggacgaagc cgcttggaat aagtctagtc 120

ttatcttgta tgtgtcctgc actgaacttg tttctgtctc tggagtgctc tacacttcag 180

taggggctgt acccgggcgg tcccactctt cacaggaatc tgcacaggtg gctttcacct 240

ctggacagtg cattccacac ccgctccacg gtagaagatg atgtgtgtct ttgcttgtga 300

aaagcttgtg aaaatcgtgt gtaggcgtag cggctacttg agtgccagcg gattacccct 360

agtggtaaca ctagcctctg ggcccaaaag gcatgtcatt tgaccactca ggtacacaac 420

cccagtgatg cacacgctta gtaatggctt agtaacaaac attgattgat catttgaaag 480

ctgttaggag gtttaggtat gacgggctga aggatgccct gaaggtaccc ataggtaacc 540

ttaagcgact atggatctga tcaggggccc accatgtaac acatgggtag aagtcttcgg 600

accttgggtt aaaaaacgtc taggcccgcc ccccacaggg atgtggggtt tcccttataa 660

ccccaatatt gtata 675

<210> 473

<211> 675

<212> DNA

<213> Parechovirus sp.

<400> 473

gccgtcgggc cttacacccc gacttgctga gtttctctag gagagtccct ttcccagcca 60

gaggtggctg gtcaaacaat accaaacgta actaaacatc taagataaca tagccctatg 120

cctggtctcc accagttgaa ggcatcttgc aataaaatgg gtggattaag acgcttaaag 180

catggagtca attatctttt ctaactagtg atcttcactg ggtggcagat ggcgtgccat 240

aactctatta gtgggatacc acgctcgtgg atcttatgcc cacacagcca tcctctagta 300

agtttgcaag gtgtctgatg aggcgtggga acttattgga aataattact tgctgcgaag 360

catcctactg ccagcggatc aacacctggt aacaggtgcc cctggggcca aaagccacgg 420

tttaacagac cctttaggat tggttaaaac ctgagtaatt atggaagata cttagtacct 480

accaacttgg taacagtgca aacactagtt gtaaggccca cgaaggatgc ccagaaggta 540

cccgcaggta acaagagaca ctgtggatct gatctggggc cacctacctc tatcctggtg 600

aggtggttaa aaaacgtcta gtgggccaaa cccaggggggg atccctggtt tccttatttt 660

agtgtaaatg tcatt 675

<210> 474

<211> 631

<212> DNA

<213> Parechovirus sp.

<400> 474

agagtccttt tcccagccag aggtggctgg ttaaataata cctactgtaa caaaacatct 60

aagatgtaac aaccacacac ctggtctcca ctggccgaag gcaactagca ataaggcagg 120

tgggttcaga cgcttaaagt gtgttgtaca tattcttttc taacctgtgt tttacacagg 180

gtggcagatg gcgtgccata actctaacag tgagatacca cgcttgtgga ccttatgctc 240

acacagccat cctctagtaa gtttgtaaga tgtctgatga cgtgtgggaa cctgttggag 300

ataacagttt gctgcaaagc atcccactgc cagcggatct acatctggta acagatgcct 360

ctggggccaa aagccaaggt ttaacagacc ctttgggatt ggttcaaacc tgaactgtta 420

tggaagacat ttagtacctg ctgatttggt agtaatgcaa acactagttg taaggcccac 480

gaaggatgcc cagaaggtac ccgtaggtaa caagtggcac tatggatctg atctggggcc 540

agctacctct atcttggtga gttggttaaa aaacgtctag tgggccaaac ccaggggga 600

tcctggtttc tttttaattt aagtaatcac t 631

<210> 475

<211> 675

<212> DNA

<213> Parechovirus sp.

<400> 475

gggccttata ccccgacttg ctgagtttct ctaggagagt ccctttccca gccctgaggc 60

ggctggataa taaaggcctc acatgtaaca aacatctaag acaaaataat ttgccttgca 120

cctggtcccc actagttgaa ggcatctagc aataagatga gtggaacaag gacgcttaaa 180

gtgcaatgat agttatcttt tctaacccac tatttatagt ggggtggtgg atggcgcacc 240

ataattctaa tagtgagata ccacgcttgt ggaccttatg ctcacacagc catcctctag 300

taagtttgtg agacgtctgg tgacgtgtgg gaacttactg gaaacaatgc tttgccgtaa 360

ggctttcatt agccagcgga ccaccacctg gtaacaggtg cctctggggc caaaagccaa 420

ggtttaatag accctaatgg aatggttcaa acctggagca ttgtggaaag tacttagtac 480

ctgctgatct ggtagtaatg caaacactag ttgtacggcc cacgaaggat gcccagaagg 540

tacccgtagg taacaagtga cactatggat ctgatctggg gccaactacc tctatcttgg 600

tgagttggtt aaaaaacgtc tagtgggcca aacccagggg ggatccctgg tttcctttta 660

ttttactttg tcaat 675

<210> 476

<211> 433

<212> DNA

<213> Parechovirus sp.

<400> 476

ctctattagt gagataccac gcttgtggac cttatgctca cacagccatc ctctagtaag 60

tttgtaagac gtctggtgac gtgtgggaac ttgtgggaat caatattttg ctttaaagca 120

tccattagcc agcggataaa acacctggta acaggtgcct ctggggccaa aagccaaggt 180

ttaacagacc ctagtggatt ggtttcaaaa cctgaaatat tgtggaacac actcagtacc 240

tactgatctg gtagtaatgc aagcactagt tgtaaggccc acgaaggatg cccagaaggt 300

acctgtaggg aacaagagac actatagatc tgatctgggg ctggctacct ctattttggt 360

gagtcagtta aaaaacgtct agtgggccaa acccaggggg gaccctggtt tccatttatt 420

ttacaaaggc act 433

<210> 477

<211> 488

<212> DNA

<213> Potamipivirus A

<400> 477

cacatggaaa gcttttcgct tccatgttta cgcacacact ctctttgaca ccctgttgta 60

tggtgttaaa ctacaacatt tgtctgtcta taatcgttta ttttgtttac cctatatgta 120

cccaagtatt tgattgcttg actcacataa gcatcggtaa cccatactgt tttatgagct 180

actacctctg ctgtctacat acattttata tgaatggttt gagctctgcc tcaggatcaa 240

acatggtaac atgttccttt ggtcagttag aatcttattg tataatctaa ggtgtctatt 300

agtacgtaga aagttgtaac acatatgggg cctgatagcc gctatctctg atggatgtaa 360

ggtaaccttc tttaggtctg atacattctg cacaggatcc aattttcggt gccctgtacg 420

agtgcactct tatgcacgag gacgagatat gctacaaccc actgcaaatt taaacccaaa 480

ctttaaca 488

<210> 478

<211> 675

<212> DNA

<213> Potamipivirus B

<400> 478

tttcaacgtc gtggctgacg ttaaaaagcc acaattccac ttacctttta ccttttatgt 60

ttaatgtttg ttagttttgt gatctttaac aaatagatct aaataatttg ttggtaacca 120

atctcggatg tttcggctgc attgtagttt atttatttca ttttagttgt aggtggccac 180

tacgtcctgg aatcatacat ggtaacatgt acctcggcgg ttatccacta ttacgctaat 240

ctaagaatat ttaaatgaaa atgtaagtgt tacggctgac tttgggcctg atagttaaat 300

gctcgcactg acagatagta ccctccttta ggatcgattc tgttacatgg gatccatttt 360

ggtgccccac tgattcaacc tctttgttga aaaagagtta gcatactaca aattttccaa 420

acaaaaaccc tttttaatga ctacaactta tgattttatg aattttactg ctcttgaaaa 480

agatattttg acattgatcg ctgtactgtt tcagacattc attgcatcca tttttgttgg 540

ctactcctca caaactcaaa acttttccac acgagaaacc ttgtttattg aattttgcct 600

ttatattttg gaacttgttg ttggatttat tgtttgctta attattgacc tcacacctgt 660

tttaaacact acaat 675

<210> 479

<211> 546

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Beihai Conger Picornavirus sequence"

<400> 479

gggacaaccc cacagctggt acaaccattg tgggttggtc tccacccttt ttcaaccgtg 60

gcaacttcgg ttaaagttgc aaatcccccc tctccctatt ccacctccct tactttcact 120

ccccatatat ggtcccagat tttattctac ctctttatat ttttatttag tacagtggtg 180

gtgaattact cccagcataa actttgctgg atcagtgttc atcaagcata ctaattacta 240

atgtactgag ctatactatt atctggcatc tcacctggat aaccggtgtg accatatttc 300

ctaggttgcc tccctatgta ttttgtagca cctgtgcatc tgcacgttgg ggcgacaaat 360

tgtaggtttc ctggcacggg taagaattgt ggaaagctag tatgcagtta atgcaagggc 420

gcgtttttcg ctaccccgac actgctaaag tttttgggag gggtccctta aacatttcta 480

gtattgagtg atagctttgc ggcaggtcac cacaacctta ctataaataa acctgttgaa 540

tctcac 546

<210> 480

<211> 457

<212> DNA

<213> Sapelovirus sp.

<400> 480

tacgcatgta ttccacactc atttcccccc tccaccctta aggtggttgt atccccatac 60

cttaccctcc cttccacaat ggacggacaa atggatttga cctcacggca aacacatatg 120

gtatgatttc ggatacacct taacggcagt agcgtggcga gctatggaaa aatcgcaatt 180

gtcgatagcc atgttagtga cgcgcttcgg cgtgctcctt tggtgattcg gcgactggtt 240

acaggagagt aggcagtgag ctatgggcaa acctctacag tattacttag agggaatgtg 300

caattgagac ttgacgagcg tctcctcgga gatgtggcgc atgctcttgg cattaccata 360

gtgagcttcc aggttgggaa acctggactg ggcctatact acctgatagg gtcgcggctg 420

gccgcctgta actagtatag tcagttgaaa ccccccc 457

<210> 481

<211> 489

<212> DNA

<213> Sapelovirus sp.

<400> 481

ttgaaatggg tgtggggtac atgcgtatta cggtacgcat atattccaca ctcatttccc 60

cccctccacc cttaaggtgg ttgtatcccc ataccttacc ctcccttcta aaacagatgg 120

acaaatggat ttgaacttat ggcaagtgaa tatggtatga ctttggatac actttaacgg 180

cagtagcgtg gcgagctatg gaaaaatcgc aattgtcgat agccatgtta gtgacgcgct 240

tcggcgtgct cctttggtga ttcggcgact ggttacagga gagtaggcag tgagctatgg 300

gcaaacctct acagtattac ttagagggaa tgtgcaattg agacttgacg agcgtctctt 360

agagatgtgg cgcatgctct tggcattacc atagtgagct tccaggttgg gaaacctgga 420

ctgggcctat actacctgat agggtcgcgg ctggccgcct gtaactagta tagtcagttg 480

aaaccccccc 489

<210> 482

<211> 675

<212> DNA

<213> Sapelovirus sp.

<400> 482

ccaaggatct gttgcatagg cgttgtatcc cctaaccttt tacctaccca tcccaatagg 60

actggtattt cggttttgat tgagtaatgg atactgattc tatacctgtt acccattcag 120

gggaaaaatg gagtttcttt catggatctg acttgatatg accaagagtc aacactttgc 180

gtgttggccg tatggaatgc tttaaggttt attctttgga ttatgacttc agggttggcc 240

gcccaggata aaaggcaatt gtggtaagtg atgttagtca ttggtggttg aaacctgcct 300

aagacgtcct aggtctacgc tgtgcgggcc gaagtaagct taggaataac agggagtatg 360

ccattttctg ctttcaccca acacgaccgt acacgaaaga gctagaggca ctttggggca 420

aagggaaaag ctttgcttag cccgaatgtt catttgagtc cttgacgaat gcgtcccgtc 480

tgtcccgacg gtgaggcgta tggcgcatgc tcatggcatt acccaatggt gtatctgtga 540

ggggggggct cctcacactt agtctagtgc tacctgacag ggccgcggct ggtcgtttgt 600

gtatggtata accagtagta atcccccatg gattgcttta acttcccctc ctcccttacc 660

aagacattct ctaag 675

<210> 483

<211> 579

<212> DNA

<213> Sapelovirus sp.

<400> 483

ttttaacttg ttatgacatt caaggaaaaa atgtcttttt cattatggga ctgacctgtt 60

tatgaacatg agcagcggca ctgctccacg ggctatccgt gtaagaaata ttgattattc 120

ttatggatca tgatttcagg gttggccgcc cagtctaaaa ggcaattgtg gtaagctatg 180

taagtagttg gctgttgaaa ggagccaagt acatcctagg tctacgctgt gcgggccgaa 240

gtaagacttg gaacaactct gagtaggcag tttttctctt tagcccaaca cgaccgcata 300

ctgaagagct agaggcactt tggggcaaag gtaaaagcat tgcttagacc gaatgttcaa 360

tgagaccttg acgagtgctg tcacagtgtc ccctgatggc agtatggcgc atgctcttgg 420

cattacccat atgtgtatct atagggggggg ggccccctat acttagtcta gtgctacctg 480

acagggccgc ggctggtcgt cggtgtgtgg tataaccagt agtaatcccc catggattgc 540

tttaactccc cctcctccct caacaaaact ttctctaag 579

<210> 484

<211> 522

<212> DNA

<213> Rabovirus C

<400> 484

ccgggtataa cccggagttt tggggcaggt ccaagcccca cataggaaca tacgatccac 60

ggatcgtgtg ttcttttatg ctttctaacc ttacccttg taaccattac gctttacgcc 120

gcatggtgtt tggcggcacc atgacgtgga caagaggtta cgccattacg atatgtaccc 180

tccccttttg gggagagacc gaccaattat ggtacagtat ccaactgtat tgtggtcaag 240

tacttctgtt tccccggtga tgcgggatag gctgtaccca cggccaaaac ctgctgatcc 300

gttacccgac tcacatctac gaggaggcta gtaaaaggca tgaagttcaa gagtatgatc 360

caaccagatc cccactggta aactagtgat gagggttccc gttccgaaca tggcaacatg 420

tgggttccct gcgttggcac taggcccctt ccgaggggtg ctctgaagat ggattgttga 480

tgaagaccaa tttgtgcatg tgtttatcct ccggccctct ga 522

<210> 485

<211> 551

<212> DNA

<213> Rabovirus A

<400> 485

ccgaccccac tggtcgaagg ccacttggca ataagactgg tggaacaagg tcgcctgtag 60

ttgattggaa ccttctttct aatgacttat gtcagcggtg ctactcacac cgtaactctc 120

ctaccctatc cccacgcttg tggaactagg aggggatgag tgattcaagt aagtactgtc 180

agaatggtga aaataatctg attctgaaac gctatggatc catcgaaaga tggggctaca 240

cgcctgcgga acaacacatg gtaacatgtg ccccaggggc cgaaagccac ggtgatagga 300

tcacccgtgt agtttgagat catatcaatg ttcatagtct agtaagatga tttgaaatct 360

aactggtctg atggctaact gcttgtctta ttgcggccta aggatgtcct gcaggtacct 420

ttagagaacc ttaagagact attgatctga gcaggagcca aggtggtctt tcccagcctt 480

ggttaaaaag cgtctaagcc gcggcagggg gcgggaggcc ccctttcctc ccaaactata 540

atatagattg t 551

<210> 486

<211> 675

<212> DNA

<213> Parabovirus C

<400> 486

gatgtatccc catcccccag tgtgtatgcc atactgcata gctcgcctat gccctatgga 60

ttcacaaccc tttcatatac cctccctacc caaccccgta accacatgct ttactccgct 120

tggggttttg cggccccatg ttgtgacgaa atggctacgc aatcaatgcg gctaatgggg 180

cctgccgctt ttaagtggcc ccagttagaa gtttatgcac acccgcccat taggaggcca 240

ccagccaggt ggtcagaggg caagcacttc tgtttccccg gtgaagtttg ataagctgtg 300

cccacggctg aagcagacag atccgttacc cgcctcacta ctacgagacg gctagtagtg 360

tgtaatatcc gaatttcatt gatccgggtg ttccccccac ccagaaacgt gtgatgagga 420

gcggcacccc tcctatggca acatagggcc tctcctgcgc tggcacacgg gctctatgag 480

catgaaatca ggagaaagtc acacgaagac cttattgtgc tagtgttgat tcctccgccc 540

ccctgaatgc ggctaatccc aactccggag cgcccgctgg caaacccgcc agaagagcgt 600

cgtaatgcgt aagtctggag cggaaccgac tactttgggt gtggcgtgtt tcctttattt 660

cctttgtatt tgtat 675

<210> 487

<211> 675

<212> DNA

<213> Parabovirus B

<400> 487

aacccataat ccattgtcca tcaatgtttt atgggggga ccctttctcc cctcccctc 60

caaatacctt ttacccctct gtaaccaaga gtgtgcaaaa tctatttact agcccagaat 120

tgcggcttct ggggaggttt attcctcatg cctaacaaga tgttacgcaa actccgggct 180

acggccctgg gcttttgccc taaagattta gaagtttaca ctatcgtcca acaggaggac 240

aacaaaccag ttgttctaag gacaagcact tctgtttccc cggtgagact ggatagactg 300

tacccacggt tgaaactggt tgatccgtta cccgactcac tacttcgaga agattagtag 360

gaaactgtga aactgattcc attgatccgg atactttccc cgtatccaga aactactgat 420

gagggttgac ttcccgacta cggcgacgta gtgtcatccc tgcgctggca gtaggcctct 480

ttgaggatgg aagatgtgga tcggtaaccg aaggtcctat tgagctagtg tttatacctc 540

cggcctcctg aatgcggcta atcctaaccc atgatctagt gctcacaaac cagtgagtag 600

ctagtcgtaa cgcgtaagtc gtgggcggaa ccgactactt tggagtgacc gtgtttccta 660

ttttactttt gtttg 675

<210> 488

<211> 675

<212> DNA

<213> Parabovirus sp.

<400> 488

accgttacgc accactcagt tggtgtttgg tggcaccaat gatggaacaa aaggctacac 60

cacttgggct acggcccgcg ccaccttgtg gcgcaaagac attagaagaa tagcataccg 120

cccactaggg ccctgcagcc agcagggtaa cgggcaagca cttctgtctc cccggtagaa 180

cggtataggc tgtacccacg gccgaaaact gaactatcgt tacccgactc cgtacttcgc 240

aaagcttagt aggaaactgg aaagttcgag ttattgaccc ggagtgttcc ccccactcca 300

gaaacgcgtg atgagggttg ccaccccgac catggcgaca tggtgggcat ccctgcgctg 360

gcacgcggcc tctaagagga taactcgctc ctactggtaa ccgaagagcc ccgtgagcta 420

cggtttattc ctccgcctcc ctgaatgcgg ctaatcctaa cccatgagca gttgccatag 480

atccatatgg tggactgtcg taacgcgtaa gttgtgggcg gaaccgacta ctttgggatg 540

gcgtgtttcc ttgttttctc catttgttgt tgtatggtga caagttatag atctcgatct 600

atagcgtttc ttgagagatt tccaaacatt tattcaagtc gtacaattct tgtgtttaag 660

cagtacagtg taagg 675

<210> 489

<211> 316

<212> DNA

<213> Felipivirus sp.

<400> 489

gatgtcggat gacggctggc caccggggaa aaacggcaaa tgtgcaccac ctctgcaacc 60

cacgccgacc acgtttaacc atggcgttag taggagtgga ccactgcagt gggctctggt 120

gtgcgacagt cagtggtaga gtagacagtc ctgactgggc aatgggaccg cgttgcgtat 180

ccctaggtgg catcgagatt cctctgctac ccaccagcgt ggactcctat gggggggccc 240

ccataggcta ggtctatact gcctgatagg gtcgcggctg gtcgaccact gactgtataa 300

ccagttgtaa ctcact 316

<210> 490

<211> 553

<212> DNA

<213

> Boosepivirus A

<400> 490

ttgaaagacc tcggcatata tcgttgtcac aacggtatat gtcgagatct ttctccccac 60

cccctccaat tccctttcc ccctcttgca acttagaagt ttgtacacac agggcaatag 120

gatacgtgat ccagccagga cacgtgagct caagcacttc tgtttccccg tcccccttcac 180

gtactacggg aatgttagta atttgtgtgc actttagtaa ggttgatccg ggattaaccc 240

caaatcccag aaactggtga tgagcgttac caccccgcc gggcgaccgg aaggtttcgc 300

tgcgttggca ccagggcttc ggcaccagaa aaaggtaaag caaatgaagg cgctactgtg 360

ctacgagaag tttcctccag gcccctgaat gcggctaatc ctaaccagtg atccaccggt 420

gcaaaaccat gtactaggtg gtcgtaacgc gcaagtcgct ggcggaaccg actactttgg 480

gtgtcctgtg tttccatatt ttattttatt caattttatg gtgacaagag taaagagata 540

cagatttgca gcc 553

<210> 491

<211> 512

<212> DNA

<213> Boosepivirus B

<400> 491

ttttctcccc tccccctcca actacctttt ccccctcttg taacgctaga agtttgtgca 60

aaccgcctgt agggtactgc aatccagcag tgcataggct aagcttttct tgttacccca 120

ccccacatta tactgaggag gattgtgaaa ttgtgttagt atgggttagt agcggtgacc 180

cgggtaaccc caacccagaa actcacggat gagatgaaca ggaccccaca tggtaacgtg 240

tgtgttcgtc tgccccgcaa ggtgaggccg tgagagcttt gcacgcgaaa accttgaaaa 300

cccaaaagta ccttgagctc ttcgctattt tgtgtttcct ccaggaccct gaatgcggct 360

aaacctaacc cgcgatccgc acgtagcaac ccagctagag tgtggtcgta atgcgcaagt 420

tgcgggcggt accgactact ttggtgttcc tgtgtttcct ttattttatt ttgaattttt 480

atggtgacaa cagctagaaa ataagagtga ac 512

<210> 492

<211> 506

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Phacovirus Pf-CHK1 sequence"

<400> 492

gtgtgtcatt tctcccctcc ccctcccaaa ccttttcccc ctctaatcgg attgattaac 60

ccggttaaag atgattaatg gtttgtgagt tgatatgatg gcccggcatt gaatccggga 120

attcttaagt aatggaattg catccaatat gaaagtgagt gtggcaagct cacaagtagt 180

acttggctct gcccattatt tgaggaccaa ctcttcttga ctacaatgtg tttaaagtaa 240

actggaccac attgtgtatc cagacaactc catttgataa tgtacgctgg aaacgttttc 300

agtgcatagg gtcctaaagt ggtgctgaaa tattgcaagc tcaatgggat actgaacgct 360

gaaaaccgcc gctgttatca tatgggcccc tagtgggtaa atgttggctt taggcatata 420

ctgcttggga atgcagtact ggttgtagac agggtgatag cctaccggct ggcgtagttg 480

agttggtata gccagttgat tgccat 506

<210> 493

<211> 612

<212> DNA

<213> Rhinovirus sp.

<400> 493

ttaaagctgg atcatggttg ttcccaccat gattacccac gcggtgcagt ggtcttgtat 60

tacggtacat ttccatacca gttttataca ccccaccccg aaactcatag aagtttgtac 120

acaatgacca ataggtggtg gccatccagg tcgctaatgg tcaagcactt ctgtttcccc 180

ggcacccttg tatacgcttc acccgaggcg aaaaatgagg ttgtcgttat ccgcaaagtg 240

cctacgaaaa gcctagtaac actttgaaaa cccatggttg gtcgctcagc tgtttaccca 300

acagtagacc tggcagatga ggctagacat tccccaccag cgatggtggt ctagcctgcg 360

tggctgcctg cacaccctgc cgggtgtgaa gccagaaagt ggacaaggtg tgaagagcct 420

attgtgctca ctttgagtcc tccggcccct gaatgtggct aaccctaacc ccgtagctgt 480

tgcatgtaac ccaacatgta tgcagtcgta atgggcaact atgggatggg accaactact 540

ttgggtgtcc gtgtttcctg ttttactttt tcattgctta tggtgacaat tgtatctgat 600

acacttgtta cc 612

<210> 494

<211> 627

<212> DNA

<213> Rhinovirus sp.

<400> 494

ttaaaacagc ggatgggtat cccaccatcc gacccacagg gtgtagtgct ctggtatttt 60

gtacctttgc acgcctgttt ccccattgta cccctcctta aatttcctcc ccaagtaacg 120

ttagaagttt aaggaaacaa atgtacaata ggaagcatca catccagtgg tgttatgtac 180

aagcacttct gtttccccgg agcgaggtat aagtggtacc caccgccgaa agcctttaac 240

cgttatccgc caatcaacta cgtaatggct agtattacca tgtttgtgac ttggtgttcg 300

atcaggtggt tccccccact agtttggtcg atgaggctag gaactcccca cgggtgaccg 360

tgtcctagcc tgcgtggcgg ccaacccagc ttttgctggg acgccttttt acagacatgg 420

tgtgaagacc tgcatgtgct tgattgtgag tcctccggcc cctgaatgcg gctaacctta 480

accccggagc cttgcaacat aatccaatgt tgttgaggtc gtaatgagta attctgggat 540

gggaccgact actttgggtg tccgtgtttc cttttattct ttatattgtc ttatggtcac 600

agcatatata gcatatatac tgtgatc 627

<210> 495

<211> 615

<212> DNA

<213> Rhinovirus sp.

<400> 495

ttaaaactgg gtttgggttg ttcccaccca aaccacccac gcggtgttgt acactgttat 60

tccggtaacc ttgtacgcca gttttatatc ccttcccccc cttgtaactt agaagacatg 120

cgaatcgacc aatagcaggc aatcaaccag attgtcaccg gtcaagcact tctgtttccc 180

cggctctcgt tgatatgctc caacagggca aaaacaattg gagtcgttac ccgcaagatg 240

cctacgcaaa acctagtagc atcttcgaag atttttggtt ggtcgctcag ttgctacccc 300

agcaatagac ctggcagatg aggctagaaa taccccactg gtgacagtgt tctagcctgc 360

gtggctgcct gcacacccac acgggtgtga agccaaagat tggacaaggt gtgaagagtc 420

acgtgtgctc atcttgagtc ctccggcccc tgaatgcggc taaccttaac cccgtagcca 480

ttgctcgcaa tccagcgagt atatggtcgt aatgagtaat tacgggatgg gaccgactac 540

tttgggtgtc cgtgtttcac tttttactta tcaatttgct tatggtgaca atatatatag 600

atatatattg acacc 615

<210> 496

<211> 660

<212> DNA

<213> Enterovirus L

<400> 496

acatgggcca gcccaccaca cccactgggt gtagtagtct ggttctatgg aacctttcta 60

cgcctctttt gcttccctcc cccatttctc cttcgattgc tccacctgtg atctttgcaa 120

cttagaagaa ataatgaacc cgcacaatag cgggcgctga gccacagcgt caatgtgcaa 180

gcacttctgt ttccccggaa tgggcccata ggctgtaccc acggctgaaa gggaccggcc 240

cgttacccgc cttggtactg cgagaatgtt agtaactccc tcgatagctt taggcgttac 300

gctcagccct ttgagcccga agggtagttc gggtcgatga ggctcgtcat tccccactgg 360

cgacagtgtg acttgcctgc gttggcggcc cggggtgggg ggcaaccccc atccacgcct 420

actgaaggac agggtgtgaa ggcgctattg cgctactaag gagtcctccg gcccctgaat 480

gcggctaacc cgaaccccga gcccacggtg gtaaacccgc cacaagtggg tcgtaatgag 540

taatttgggg cagggaccga ctactttggg tgtccgtgtt tcctgttttt ccatacgatg 600

gctgcttatg gtgacaacca taagcaattg gattggccat ccggtgttca tattgcgaat 660

<210> 497

<211> 675

<212> DNA

<213> Enterovirus K

<400> 497

tcagcctgac gcaagtgcct ccattggagt ctctccaagc cctccggggc ttggagggcg 60

ccgacccccc gcctaggggg agcccacgac acggctggag tccattggca caccgcagcc 120

acgattcaag ccagaattga aagcgggaag cacttctgtc tccccggtgt ggatcatacg 180

ctgtacccac ggcgaaaagt gaagcatcgt tacccgactc ggtacttcga gaagcccagt 240

acagttgtgg atctctgcag ggtatacgct cagcgtgacc cctacgtagt tccttgagat 300

ggctgagaga acaccccacg ggcgaccgtg tctctcggcg cgtggctcaa ggccgggcct 360

tcagtggctc ggtgccttgc agagtgaagc ctccgaacag cctattgagc taccgtttag 420

cctccgccct cttgaatgcg gctaatccta accatggagc gcccgcccac agtccagtgg 480

gtagagcgtc gtaacgcgca agtccgtggc ggaaccgact actttagagt ggcgtgtttc 540

caatttatcc tttataaagt tgcttatggt gacaccacaa gagatccacg atttcttgtt 600

tcttatcact gagacacaag tcatattcat caatctttat tgcggaatta acttggtgcg 660

tccaaacaca tcagc 675

<210> 498

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 498

caccctgagg gcccacgtgg cgtagtactc tggtatcaag gtacctttgt acgcctattt 60

tatttccctt cccccacagt aacttagaag cttatctcat agttcaacag tagggtcact 120

aaccaagtgg ctcagcgaac aagcacttct gtttccccgg tcctagtacc tgtgaagctg 180

tacccacggc ggaaggggaa aaagatcgtt atccggcccc ctacttcgga aagcctagta 240

acaccattga agcaatcgag tgttgcgctc agcacagtaa cccctgtgta gctttggttg 300

atgagtctgg gcactcccca ctggcgacag cggcccaggc tgcgttggcg gccaaccgac 360

tcgggcaacc gggtcggacg ctcgtttgtg gacatggtgt gaagagccta ctgagctaga 420

gggtagtcct ccggcccctg aatgcggata atcctaaccc cggagcaccc acactcaatc 480

cagagtgcag gatgtcgtaa cgcgtaagtc tgggacggaa ccgactactt tgggtgtccg 540

tgtttcctgt tttacttact ttggctgctt atggtgacaa tctagtgttg ttaccatata 600

gctattggat tggccatccg gtgttttgaa ttgtgtgttt atactaattc ttttacatat 660

cacagacaac caaat 675

<210> 499

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 499

cggtaccttt gtacgcctat tttacccccct tccccttgta acttagaagc aaagcaaacc 60

agttcaatag taagcaacac aacccagtgt tgtgacgaac aagtacttct gtttccccgg 120

gagggtctga cggtaagctg tacccacggc tgaagtatga cctaccgtta accggctacc 180

tacttcgaga agtctagtaa taccattgaa gttttgttgg cgttacgctc aacacactac 240

cccgtgtgta gttttggctg atgagtcacg gcattcccca cgggcgaccg tggccgtggc 300

tgcgttgcgg ccaaccaagg ggcgcaagct ccttggacgt cacttaacag acatggtgtg 360

aagaacctat tgagctaggt agtagtcctc cggcccctga atgcggctaa tcctaactcc 420

ggagcacatc agtgcaaccc agcatttggt gtgttgtaat acgcaagtct ggagcggaac 480

cgactacttt gggtgtccgt gtttcctgtt ttaccttatt tggctgctta tggtgacaat 540

ttgatattgt taccatatag ctgttggatt ggccatccgg atttttgaaa gagacccaaa 600

actttcttct ctacttcaga ttcaagtgcg aagttttcct tttcatatat tacttactaa 660

tttgaagtac caaag 675

<210> 500

<211> 675

<212> DNA

<213> Enterovirus I

<400> 500

ttagtacttt ctcacgggga tagtggtatc cctccctagt aatttagaag acttgaaaaa 60

ccgaccaata ggcacctcgc atccagcggg gtaaaggtca agcacttctg tttccccggg 120

tcgagtagcg atagactgtg cccacggtcg aaggtgaaac aacccgttat ccgactttgt 180

acttcgggaa gcctagtacc accaaagatt atgcttgggg tttcgctcag cacgaccctg 240

gtgtagatca ggccgatgga tcaccgcatt cctcacaggc gactgtggcg gtggtcgcgt 300

ggcagcctgc cgatggggca acccatcgga cgccaagcat atgacagggt gtgaagagcc 360

tactgagcta caaagtattc ctccggcccc tgaatgcggc taatcccaac cacggagcat 420

ttgctaccaa accaggtagt ggaatgtcgt aacgggtaac tctgtggcgg aaccgactac 480

tttgggtgtc cgtgtttcct tttaatttat cattctgtat atggtgacaa ctatagtgct 540

atctcgattt gcattactat tgttgagatt aaaactttat tacattgttg cattttaccc 600

tttgagtgag ttttcacctg aacagattaa tttactcatc ctgtttatat attacaagca 660

gaaatacttg caaag 675

<210> 501

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 501

gcaatgctgc accagtgcac tggtacgcta gtacctttc acggagtaga tggtatccct 60

taccccggaa cctagaagat tgcacacaaa ccgaccaata ggcgcaccgc atccagccgt 120

gcagcggtca agcacttctg tctccccggt ctgtaaagat cgttatccgc ccgacccact 180

acgaaaagcc tagtaactgg ccaagtgaac gcgaagttgc gctccgccac aaccccagtg 240

gtagctctgg aagatggggc tcgcaccacc cccgtggtaa cacggttgcc tgcccgcgtg 300

tgcttccggg ttcggtctcg tgccgttcac ttcaacttca cgcaaccagc caagagccta 360

ttgtgctggg acggttttcc tccggggccg tgaatgctgc taatcccaac ctccgagcgt 420

gtgcgcacaa tccagtgttg ctacgtcgta acgcgtaagt tggaggcgga acagactact 480

ttcggtaccc cgtgtttcct ctcattttat ttaatatttt atggtgacaa ttgttgagat 540

ttgcgctctt gcaacgttgc cattgaatat tggcttatac tatttggttg ccttttacaa 600

aacctctgat atacccagtt cttacattga tctgcttgtt tttctcaatt tgaagtatag 660

actacaaata gcaaa 675

<210> 502

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 502

cgtggcggcc agtactctgg tatcacggta cctttgtacg cctgttttat atccccttcc 60

cccgcaactt agaagaaaac aaatcaagtt cactaggagg gggtacaaac cagtaccacc 120

acgaacaagc acttctgttt ccccggtgat gtcgtataga ctgtaaccac ggttgaaaac 180

gattgatccg ttatccgctc ttgtacttcg aaaagcccag tatcaccttg gaatcttcga 240

tgcgttgcgc tcagcactca accccagagt gtagcttagg tcgatgagtc tggacactcc 300

tcaccggcga cggtggtcca ggctgcgttg gcggcctacc tgtggtccaa agccacagga 360

cgctagttgt gaacaaggtg tgaagagcct attgagctac aagagaatcc tccggcccct 420

gaatgcggct aatcctaacc acggagcaag ggtacacaaa ccagtgtata tcttgtcgta 480

acgcgcaagt ctgtggcgga accgactact ttgggtgtcc gtgtttcctt ttgtttttat 540

catggctgct tatggtgaca atctaagatt gttatcatat agctgttgga ttggccatcc 600

ggtaatttat tgagatttga gcatttgctt gtttcttcaa caatttcacc tattcattgc 660

atttcagcag tcaaa 675

<210> 503

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

PV3 sequence"

<400> 503

taccttttgta cgcctgtttt atactccctc ccccgcaact tagaagcata caattcaagc 60

tcaataggag ggggtgcaag ccagcgcctc cgtgggcaag cactactgtt tccccggtga 120

ggccgcatag actgttccca cggttgaaag tggccgatcc gttatccgct catgtacttc 180

gagaagccta gtatcgctct ggaatcttcg acgcgttgcg ctcagcactc aaccccggag 240

tgtagcttgg gccgatgagt ctggacagtc cccactggcg acagtggtcc aggctgcgct 300

ggcggcccac ctgtggccca aagccacggg acgctagttg tgaacagggt gtgaagagcc 360

tattgagcta catgagagtc ctccggcccc tgaatgcggc taatcctaac catggagcag 420

gcagctgcaa cccagcagcc agcctgtcgt aacgcgcaag tccgtggcgg aaccgactac 480

tttgggtgtc cgtgtttcct tttattcttg aatggctgct tatggtgaca atcatagatt 540

gttatcataa agcgagttgg attggccatc cagtgtgaat cagattaatt actcccttgt 600

ttgttggatc cactcccgaa acgttttact ccttaactta ttgaaattgt ttgaagacag 660

gatttcagtg tcaca 675

<210> 504

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 504

ctttgtacgc ctgttttaca tcccctcccc cacgtaactt tagaagcaat tcaacaagtt 60

caatagaggg ggtacaaacc agtatcacca cgaacaagca cttctgtttc cccggtgatt 120

ttacataagc tgtgcccacg gctgaaagtg aatgatccgt tacccgctcg agtacttcga 180

aaagcctagt atcgctttgg gatcttcgac gcgttgcgct cagcactcta ccccgagtgt 240

agcttaggct gatgagtctg ggcattcccc atcggcgacg atggcccagg ctgcgttggc 300

ggcctaccca tggctaacgc catgggacgc tagttgtgaa caaggtgtga agagcctatt 360

gagctactcg agagtcctcc ggcccctgaa tgcggctaat cccaaccacg gatcaggtgc 420

ctccaaccca ggaggtggcc tgtcgtaacg cgcaagtctg tggcggaacc gactactttg 480

ggtgtccgtg tttcctttta tcttttaaat ggctgcttat ggtgacaatc atagattgtt 540

atcataaagc gaattggatt ggccatccgg tgaaatacaa acacattatt tacttgtttg 600

ttggatttac tccgctcaca cagcttactc ctaagataat atttattgta ttgctggtaa 660

ggagacacta ttata 675

<210> 505

<211> 636

<212> DNA

<213> Enterovirus sp.

<400> 505

aagcaaggca aacctgacca atagtaggtg tggcacacca gccgcatttt ggtcaagcac 60

ttctgtttcc ccggaccgag tatcaataag ctgctcacgc ggctgaagga gaaaccgttc 120

gttacccgac cagctacttc gagaaaccta gtaacactat gaacgttgcg gagtgtttcg 180

ttcagcactt cccccgtgta gatcaggtcg atgagtcacc gcattcctca cgggtgaccg 240

tggcggtggc tgcgttggcg gcctgcctac gggttcgccc gtaggacgct ctaataccga 300

catggtgtga agagtccatt gagctagctg gtagtcctcc ggcccctgaa tgcggctaat 360

cctaactgcg gagcaggtgc tcacagacca gtgagtagcc tgtcgtaacg ggcaactctg 420

cagcggaacc gactactttg ggtgtttttc ctttttttctt ctcttatatt ggctgcttat 480

ggtgacaatt aaagaattgt taccatatag ctattggatt ggccatccgg tgacgagcag 540

agccattgtt tacctctttg ttggatttgt acctttgaac cacaaagtct tgaataccat 600

tcatctcatt ttaaagttca actcagctaa aagaaa 636

<210> 506

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

SA5 sequence"

<400> 506

agtacttggt attccggtac ctttgtacac ctatttacaa accctacccc ttgtaacctt 60

agaagcaatt atttaaccgc tcactagggg gtgtgctatc caagcacatc aagagcaagc 120

acttctgtct ccccgggagg ggctaatggt acgctgtgcc cacggcggaa atgagcccta 180

ccgttaaccg gcagtctact tcgggaagcc cagtaactac attgaaactt tgaggcgtta 240

cactcagcac ataaccccaa tgtgtagttc tggtcgatga gccttggcat cccccacagg 300

cgactgtggc caaggctgcg ttggcggcca gcctgcggac caaaagtccg taggacgcct 360

aattgtggac atggtgtgaa gagcctactg agctagactg tagtcctccg gcccctgaat 420

gcggctaatc ctaaccctgg agcatccgcg tgcaacccag tacgtagggt gtcgtaatgc 480

gtaagtctgg gatggaaccg actactttgg gtgtccgtgt ttcttgtttt tcatactggg 540

tcgcttatgg ttacaactaa ttgttgtaat cattggcagt gcgcgctgac cacgcgatta 600

ttgatatttc catttgttgg atactccaat agtgtcaact catatacaca acttttacca 660

ctgatcaaga taaaa 675

<210> 507

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 507

tgtgcgcctg ttttgaaacc ccctccccca actcgaaacg tagaagtaat gtacactact 60

gatcagtagc aggcgtggcg caccagccat gtctcgatca agcacttctg tttccccgga 120

ctgagtatca atagactgct cacgcggttg aaggtgaaaa cgtccgttac ccggctaact 180

acttcgagaa acctagtagc accatagaaa ctgcagagtg tttcgctcag cacttccccc 240

gtgtagatca ggtcgatgag tcactgcaat ccccacgggt gaccgtggca gtggctgcgt 300

tggcggcctg cctatggggc aacccatagg acgctctaag gtggacatgg tgtgaagagt 360

ctattgagct agttagtagt cctccggccc ctgaatgcgg ctaatcctaa ctgtggagcg 420

catactccca aaccagggag cagtgcgtcg taacgggcaa ctccgcagcg gaaccgacta 480

ctttgggtgt ccgtgtttcc ttttattcct atactggctg cttatggtga caattgagag 540

attgttacca tatagctatt ggattggcca tccagtgtgt aatagagcaa tcatttacca 600

atttgttgga tttactccat taacccacac gtctctcaac acactacatt tcatcttact 660

actgaacact agaaa 675

<210> 508

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Mobovirus A sequence"

<400> 508

tattctccca caaaccttct tgtaactctg ttaagccttt tacatccatg taatttaatt 60

ttctccacct aaaaggattt cccccatggt ccttttttggc tcgaacaaat gctacatagg 120

gtcttgtttt ctccccctgg ctctcttgcc agggttccat accccaattc ctctatttcc 180

atgatttttc atcatggttt atttttactg tcttcttatt ttctgaggtg accaactcct 240

aagccgactg ggtcgcggaa gcccggactc ctcgcatcac tagggtgcgt agcgatgtag 300

gcgaaaatat tggttgctag atgcatacat atagtgaatt gatactacac caaactctgt 360

tctttttgaa actagctatt ttctaagtaa ggtaggctac gggtgaaacc ttaccattgc 420

aggtacgtga accgcaacgg acatttggcc gaagactggt gtacccacgt cagttatagg 480

acctcttcaa cgttggtgga cggcatgtca ctgattagtt aggctagtga atttaagttc 540

aggggtatc ttttagctta agcgtgtatt ctagtaggac ttgcagagcc tccccaccta 600

ggaggatctc tgtttatagc ccctttttcct tgttccgtta gttttccaca cttttacaat 660

atttgatgat ttgtt 675

<210> 509

<211> 509

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Burpengary Virus sequence"

<400> 509

ctccccccc ttccccttcc cgagtaggag attggcatgt atgctctaca tgcccgattc 60

tctcttgctc actctcttaa atcctggtgg cggtctcgga ttaaacattt atgtcgtatc 120

tgggatcgtc ttacttggtg gtaattcctc tgttgcctag ggacctccgg actgccggat 180

taaaggtctc aacagagggc aatgtacaag gaagtcatta tacgctaatt aagtatttga 240

tgaatgacta gtgtgacagg gctgaggaac tccccccggg taaccggtgc ctcagcgtcc 300

gaaagacacg tggataggat ccaccctgtt atacccagca cgatgtaata gtcaaatacc 360

tctgatttgt gtaggatgta taaattgtgc attgtaaatt ttgggcgtag agatgctccg 420

aaggtacccc gttttacggg atctgatcgg aggctaatta cccaatgcgc cctaaataac 480

ttcatataat ttctttttct ttattcaaa 509

<210> 510

<211> 675

<212> DNA

<213> Hunnivirus sp.

<400> 510

taacgtttgg caagaaccct cacctgtcaa ttgggaccac cactttcagt gaccccatgc 60

gaagtagtga gagagaataa gctttcttac ccttcatttg tgaacccttc agtcgaagcc 120

gcttggaata agataggagg aaaagttcat tctaaatgga gtgaaacatg tacttcagaa 180

tttctagcac gcgctgggct ttcttgcgtg tgacggcact gtcttgccgg agctctccac 240

actgacaccc cacgcttgtg gaccttggtg gcagatgaca acactgcagc tggaattgag 300

tgtctggtac actctgtgta acagtgaaaa caatgtgatc acttcggtga gctagtagcc 360

tgtggaccaa caactggtaa cagttgcctc aggggccaaa agccacggtg tttacagcac 420

cctactggtt tgattggagc aatccaagat gtcacagagt tagtaattgc caagcagtcc 480

gtactggtat cttgacatac cgtgcagttt tggatagtga aggatgccct gacggtaccc 540

ataggtaaca agtgacacta tggatctaag caggggctca ctctacgctg ctttacagct 600

ggctgtgagt taaaaaacgt ctagctatcc acaacctagg ggactaggtt ttccttttat 660

ttagattaca attat 675

<210> 511

<211> 675

<212> DNA

<213> Hunnivirus sp.

<400> 511

acagtttttg acaaggaccc tcacctgtca atcgggacca ccactttcag tgaccccgtg 60

cgaagtgttg agagaaagtg agctttctta cccttcattt gtgaaccctt cagtcgaagc 120

cgcttggaat aagatggaag gaaatgttca ttctaaatgg agtgaaacat atacttaatt 180

tccagtgttt agtggtcttt ccactagaca acggcactgt cttgccggaa ctctacacac 240

caacattcca cgcttgtggg actcaaatgt tggatgacac agttgtagct ggaactgagt 300

gtttagtgca ctctgtgtaa cagtgaaaac aatgtgatca cttcggtggg ctagtagcct 360

gtggactaac aactggtaac agttgcctca ggggccaaaa gccacggtgt taacagcacc 420

ctactagttt gattggagca atccatgatg ttacagagtt agtaactgcc aaacagattg 480

tactggtatc ttggcatacc gtgcaactta ggatagtgaa ggatgccctg gcggtaccca 540

taggtaacaa gtgacactat ggatctaaac aggggctcac tctacgttgc tttacaacta 600

gctgtgagtt aaaaaacgtc taactatcca caacctaggg gactaggttt tccttttat 660

ttttatacac aacta 675

<210> 512

<211> 240

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Ia Io sequence"

<400> 512

ggtgtccggg tgtgtgggat agacccagat gtgcagtgga tgcgagcatt tgagtcagag 60

taggagcaag cccaggggca aagggaccac attgtgtatc ccgaatgaag gatcgagatt 120

tctctcctca ttacccggtg tcttgtcact gttgggggggg cccaacagtc ttagtcctat 180

actgcctgat agggtcgcgg ctggccggac tcaagtgcta tagtcagttg attttcactc 240

<210> 513

<211> 416

<212> DNA

<213> Taura Syndrome Virus

<400> 513

ctttaaaagt cgtgcgtggc ttcaccacgc acgatcagta ctatcagtta accactcttg 60

aatatgctca atgaccctat tcaacactgg tgtctcttag tacattattt tagcacttaa 120

cgtgcatgag ttttgcccat ttctttcaaa aaaatgagta ttcgaggaga cgtcccgctc 180

cccgtcttat ttcaaccgta gactcgacat ctattggtgg acatttaatt ccagtcgccg 240

taagttgctt ctgccccgcg ctatattttc ttatacttat ggttctatag gtctggttta 300

aaacgtaaat agacggccca caaactatag aacgcgtacc cggaacgcca atcccggata 360

agtccctgga tatatagatg caccgcaata taagcctgca gactgtctca tatact 416

<210> 514

<211> 604

<212> DNA

<213> Acute bee paralysis virus

<400> 514

cccgtcaaaa taacaactta taacacgatg ttacccgaag aaaccatttt agtgtaacat 60

ttaagattag aagtagttca tctaatagag ataggcacta ttagaaggag gccttttcta 120

aaggagccgt tagtcagccc agacaagcgc agtactttag aagagagaag ttccccgata 180

gcgaccgaaa agacgcgttt tccgtgctaa ctaatttaaa tgtgggaacg aatattatta 240

ttgaaattat gtgagccacg tagcaatcaa gtcatgtttt tgtcactacg tttactcatc 300

taatgtagat aattttgttt aagtacctat ttaggtgtca tcccaccaga gaagaaataa 360

tacgtaccgg aacccagagt acacccctta ttttaagcct tactgggctt ctctgttagt 420

tagtaatctg gcccacgttt tgcgttgagt ggggtcccaa cagtaggaat tcgacggaca 480

agtagcaagc gagtcggtac caattggttt agccttcgaa attactctgg gcaggaagtt 540

actaaacgag aactttctgc ttaaatccca acgcacaaac aaatagagta aataaataat 600

tata 604

<210> 515

<211> 539

<212> DNA

<213> Bovine rhinitis A virus

<400> 515

tttgttttgc ggctttgccg ttgttcgggt tttacctgtt ttcacacagc aaaacaggcc 60

ttctagtttc gtgcttaaac gagatcatgc tcgaactaga actacatagc tggtcactgg 120

actcatacca caccttgtgg agctttatgg gaaaggtggc tagtgggctg tggaagtgac 180

tctgaccaca tgcctctcaa gtgtgggaaa tcacggatcg gtgtagcgac gacaacaggc 240

cttgggacac cctctccagt aatggagacc caaggggcca aaagccacgc ctcgtgccct 300

gttgttcaca accccagtgc gacccgtgtt agtacctatt tgcgagaact gtgtctggac 360

agctaaacac aaccctagtg ggagactaag gatgcccagg aggtacccgg aggtaacaag 420

tgacactctg gatctgacct ggggagagag ggcttgcttt acaggcgcct ctctttaaaa 480

agcttctatg tctcatcagg caccggaggc cgggcctttt ccttttaaaa ttacactta 539

<210> 516

<211> 608

<212> DNA

<213> Bovine rhinitis B virus

<400> 516

ccccccctac ttaaagatgt acggttttgc tgctttcaca gagtaaagca gatagaggtt 60

ctgaactggc aaactttacc tcgaaacacg cccgtttttc tgctgtgtct cacagactgt 120

cctgtcacac ttgtggcggc ttgtgacact gtgaacatag tgagaccgac caagacaaca 180

gtttcaagtg atgaacatcg aacgtctaaa ctggatccgt aactggacat gttagggcaa 240

ggacttcccc cctggtaaca ggagcctggc tggccaaaag ccccgctcat tgagcctagc 300

atgttgtcga ccctggactg ttcagtttag ttagtacatg gaattcactt gtcacggttc 360

ttctgaactc ggtctctagt atgacagcct aaggatgccc tccaggtacc ccggggtaac 420

aagtgacacc cgggatctga ggaggggact actttacgta gtttaaaaaa cgtctaagct 480

gttatggtga ccagaggctg gcacctttca cttttaaaat tacactactg actacaattg 540

aagtgataac ggttttacag gctttcaaac tagttacaca agcactgttt tcctgacaca 600

cacacttt 608

<210> 517

<211> 576

<212> DNA

<213> Equine rhinitis A virus

<400> 517

aatattggcg cgcgcatttg cgcgcccccc cccatttcag ccccctgtca ttgactggtc 60

gaaggcgttc gcaataagac tggtcgtcac ttggctgttc tatcgtttca ggctttagcg 120

cgccctcgcg cggcgggccg tcaagcccgt gcgctgtata gcgccaggta accggacagc 180

ggcttgctgg attttcccgg tgccattgct ctggatggtg tcaccaagct gacaaatgcg 240

gactgaacct cacaaagcga cacacctgtg gtagcgctgc ccaaaaggga gcggaactcc 300

cccgccgcga ggcggtcctc tctggccaaa agcccagcgt taatagcgcc ttttgggatg 360

caggagcccc acctgccagg tgtgaagtgg agtgagtgga tctccaattt ggtctgttct 420

gaactacacc atctactgct gtgaagaatg ccctggaggc aagctggtta cagccctgac 480

caggggccct gcccgtgact ctcgatcggc gcagggtcaa aaattgtcta agcagcagca 540

ggaacgcggg agcgtttctt ttccctttgt atcgac 576

<210> 518

<211> 466

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GFTV sequence"

<400> 518

atggggaagg gtatgacgtg ccccttcctt cttcggagaa ctcgctctag tggtctttcc 60

acttctggaa aagagtgagt gcacgtgatc aggaccgtcg aagacgacaa atacctggtg 120

ctctatctca tagacgtttc acagctgtag cgacccctca gtagcagcgg aagccccctc 180

ctggtgacag gagcctctgc ggccaaaagc cacgtggata agatccactg ctgagggcgg 240

tgcgacccta gcaccctgtg atgcatacta gttgtagcgt gccggactat tggtctgtca 300

taagacacct gatagagaga ccaagaatgt cctggaggta ccccgcgtgc gggatctgac 360

caggagacca ttgcccaatg ctttacaacg ggtctatggt ttaaaaactg tcgcagtctc 420

tccaaaccaa gtggtcttgg ttttcaatta ctttgaatat ttcact 466

<210> 519

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

SAFV V13C sequence"

<400> 519

ttttcgacgt ggttggaatt gccatcattt ccgacgaaag tgctatcatg cctccccgat 60

tatgtgatgt tttctgccct gctgggcgga gcattctcgg gttgagaaac cttgaatctt 120

tttctttgga accttggttc ccccggtcta agccgcttgg aatatgacag ggttattttc 180

ttgatcttat ttctactttt gcgggttcta tccgtaaaaa gggtacgtgc tgccccttcc 240

ttctctggag aattcacacg gcggtctttc cgtctctcaa caagtgtgaa tgcagcatgc 300

cggaaacggt gaagaaaaca gttttctgtg gaaatttaga gtgcacatcg aaacagctgt 360

agcgacctca cagtagcagc ggactcccct cttggcgaca agagcctctg cggccaaaag 420

ccccgtggat aagatccact gctgtgagcg gtgcaacccc agcaccctgg ttcgatgatc 480

attctctatg gaaccagaaa atggttttct caagccctcc ggtagagaag ccaagaatgt 540

cctgaaggta ccccgcgtgc gggatctgat caggagacca attggcggtg ctttacactg 600

tcactttggt ttaaaaattg tcacagcttc tccaaaccaa gtggtcttgg ttttccaatt 660

ttgttgaatg gcaat 675

<210> 520

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

SAV P-113 sequence"

<400> 520

ggagatctaa gtcaaccgac tccgacgaaa ctaccatcat gcctccccga ttatgtgatg 60

ctttctgccc tgctgggtgg agcatcctcg ggttgagaaa accttcttcc tttttccttg 120

gaccccggtc ccccggtcta agccgcttgg aataagacag ggttatcttc acctcttcct 180

tcttctactt catagtgttc tatactatga aagggtatgt gtcgcccctt ccttctttgg 240

agaacacgcg cggcggtctt tccgtctctc gaaaagcgcg tgtgcgacat gcagagaacc 300

gtgaagaaag cagtttgcgg actagcttta gtgcccacaa gaaaacagct gtagcgacca 360

cacaaaggca gcggaccccc cctcctggca acaggagcct ctgcggccaa aagccacgtg 420

gataagatcc acctttgtgt gcggcacaac cccagtgccc tggtttcttg gtgacacttc 480

agtgaaaacg caaatggcga tctgaagcgc ctctgtagga aagccaagaa tgtccaggag 540

gtaccccttc cctcgggaag ggatctgacc tggagacaca tcacatgtgc tttacacctg 600

tgcttgtgtt taaaaattgt cacagctttc ccaaaccaag tggtcttggt tttcactctt 660

taaactgatt tcact 675

<210> 521

<211> 612

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

VHEV sequence"

<400> 521

aattccttct tcctttctcc ttggacctcg gtcccccggt ctaagccgct cggaatatga 60

cagggttatt ttcacctctt ctctcttcta cttcatagtg ttctatacta tgaaagggta 120

tgtgtcgccc cttccttctt ggagaacgtg cgtggcggtc tttccgtctc tcgaaaaacg 180

tgcgtgcgac atgcagagta acgcaaagaa agcagttctt ggtctagctc tggtgcccac 240

aagaaaacag ctgtagcgac cacacaaagg cagcggaaac cccctcctgg taacaggagc 300

ctctgcggcc aaaagccacg tggataagat ccacctttgt gtgcggtgca accccagcac 360

cctggtttct tggtgacacc ttagtgaacc ctcgaatggc aatctcaagc gcctctgtag 420

gaaagccaag aatgtccagg aggtacccct tcctcatgga gggatctgac ctggagacac 480

atcacacgtg ctatacactt gtgcttgtgt ttaaaaattg tcacagcttt cccaaaccaa 540

gtggtcttgg ttttccctta acttcgaaaa gtcactatgg cctgcaaaca tggataccca 600

gacgtgtgcc ct 612

<210> 522

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

TRV NGS910 sequence"

<400> 522

atgcgacgtg gttggagatt aaaccgactc cgacgaaagt gctatcatgc ctccccgatt 60

atgtgatgtt ttctgccctg ctgggcggag cattctcggg ttgatacacc ttgaatcctt 120

catccttgga cctcaggtcc cccggtctaa gccgcttgga atacgacagg gttattttcc 180

aatcttctcc tttctacttt catgagtcct attcatgaaa agggtctgtg ctgccccttc 240

cttcttggag aatctgcgcg gcggtctttc cgtctctcga aaagcgcaga tgcagcatgc 300

tggaaccggt gaagaaaaca gttctttgtg gaaacttaga gcagacatcg aaacagctgt 360

agtgacctca cagtagcagc ggaacccccct cctggtaaca ggagcctctg cggccaaaag 420

ccccgtggat aagatccact gctgtgagcg gtgcaacccc agcaccctgg ttcgatggtt 480

gttctctgtg gaaccagaga atggtctttc tcaagccctc cagtagagaa gccaagaatg 540

tcctgaaggt accccgcatg cgggatctga tcaggagacc aatcgtcagt gctttacact 600

ggcgctttgg tttaaaaact gtcacagctt ctccaaacca agtggtcttg gttttcactt 660

ttatcaaact gtttc 675

<210> 523

<211> 551

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

EMCV2 RD1338 sequence"

<400> 523

aaatactggt cgaaaccgct tgggataaga ccggggtttg ttaatgtctc aatgttattc 60

tccacccaat tgacgtcttt tgtcaattgg agggcagtga aaccttgccc ttgcttcttg 120

cagaggattc ccagtggtct ttccgctctc gacaagggaa ttcatgatcc accaaaagtt 180

gtgaagagag caggtcccat ggaagctttc tgacgactga tgatgactgt agcgaccctt 240

tgcaggcagc ggaccccccc acctggtgac aggtgcctct gcggccaaaa gccacgtgtt 300

taacagacac ctgcaaaggc ggcacaaccc cagtgcctca tcaaaagtct gatgactgtg 360

gaaatagtca accggctttt cttaagcaaa tttggtgtcg gggctgaagg atgcccggaa 420

ggtaccacac tggttgtgat ctgatccggg gccacagtac atgtgcttta cacatgtagc 480

tgcggttaaa aaacgtctag gcccccccgaa ccacggggac gtggttttcc tttgaaaacc 540

acgattacaa t 551

<210> 524

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

EMCV1 JZ1203 sequence"

<400> 524

gtctgctcga tatcgcaggc tgggtccgtg actacccact ccccctttca acgtgaaggc 60

tacgatagtg ccagggcggg tactgccgta agtgccaccc caaaacaaca acaaaccccc 120

cctaacatta ctggccgacg ccgcttggaa taaggccggt gtgcgtttgt ctatatgtta 180

tttcccacca cattgccgtc ttttggcaat gtgtgggccc ggaaacctgg ccctgtcttc 240

ttgacgagca ttcctagggg tctttcccct ctcgccaaag gaatgcaagg tctgttgaat 300

gtcgtgaagg aagcagttcc tctggaagct tcttgaagac aaacaacgtc tgtagcggcc 360

ctttgcaggc agcggaaccc cccacctggc gacaggtgcc tctgcggcca aaagccacgt 420

gtataagata cacctgcaaa ggcggcacaa ccccagtgcc acgttgtgag ttggatagtt 480

gtggaaagag tcaaatggct ctcctcaagc gtattcaaca aggggctgaa ggatgcccag 540

aaggtacccc attgtatggg atctgatctg gggcctcggt gcacatgctt tacatgtgtt 600

tagtcgaggt taaaaaacgt ctaggccccc cgaaccacgg ggacgtggtt ttcctttgaa 660

aaacacgatg ataat 675

<210> 525

<211> 544

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

EMCV1 AnrB-3741 sequence"

<400> 525

atgtggtcga agccacttgg aataagaccg gcgtgcgctt gtctatatgt tacttccacc 60

acattgccgt cttttggcaa tgtgagggcc cggaacctgg ccctgtcttc ttgacgaaca 120

ttcctagggg actttcccct ctcgccaaag gaatgtaagg tctgttgaat gtcgtgaagg 180

aagcagttcc tctggaagct tcttgaagac aaacagcgtc tgtagcgacc ctttgcaggc 240

agcggaaccc cccacctggt aacaggtgcc tctgcggccg aaagccacgt gtataagata 300

cacctgcaaa ggcggcacaa ccccagtgcc acgttgtgcg ttggatagtt gtggaaagag 360

tcaaatggct ttccccaagc gtattcaaca aggggctgaa ggatgcccag aaggtacccc 420

actggttggg atctgatctg gggcctcggt gcaggtgctt tacacctgtt gagtcgaggt 480

taaaaaacgt ctaggccccc cgaaccacgg ggacgtggtt ttcctagaaa accactatga 540

caat 544

<210> 526

<211> 675

<212> DNA

<213> Cosavirus sp.

<400> 526

cgtgctttac acggtttttg aaccccacac cggctgtttg gcgcttgcag gacagtaggt 60

atattttctt tcatttctct tttctagccg cgtaggttct atctacgcgg gcggagtgat 120

actcccgctc cttcttggac aggcggcctc cacgctcttt gtggatctta aggctgccaa 180

gtcactggtg tttgaagtga agaatggaga gacactaggg cgtttcatgt ggctttgcca 240

gggattgtag cgatgctgtg tgtgtgtgcg gatttcccct cgtggcgaca cgagcctcac 300

aggccaaaag ccctgtccga aaggacccac acagtggggt tgccccgacc cctcccttca 360

aagctttgtg taaacaaact tttgtttaga ctttcttaag cttctctcac atcaggcccc 420

aaagatgtcc tgaaggtacc ctgtgtatct gaggatgagc accaccaact acccggactt 480

gtgggacgtg tcccacagac gcatgtggta ttccagcccc ctccttttga ggagggggct 540

tttgctcgct cagcacagga tctgatcagg agattcatct ctggtgcttt acaccagagc 600

atggatttaa aaattgccca aggcctggca aacaacctag gggactaggt tttctctatt 660

ttaaaagatg tcaat 675

<210> 527

<211> 675

<212> DNA

<213> Cosavirus sp.

<400> 527

cgtgctttac acggtttttg aaccccacac cggctgtttg gcgcttgcag gacagcaggt 60

ttattttctt ttaactctct ctttctagcc acacacgatc tatgtgtgtg ggcggagtga 120

tactcccgtt ccttcttgga caggcggcct ccacgccctt tgtggatctt aaggctacca 180

agtcactggt gttggaaagt gaagagaaag gagttccttg ggaactacat gtggcattga 240

cagaggttgt agcgatgctg tgtgtgtgtg cggattaccc ccgtggcgac acggacccca 300

caggccaaaa gccctgtccg aaaggaccca cacagtggag caaccccagc tcccctcttc 360

aatgttttgt gttagcaacc ttggtattat tttctctcaa gcttccaata caccgggccc 420

caaagatgtc ctgaaggtac cccgtgtatc tgaggatgag caccatcaac tacccggact 480

tgttctttcg agaacagacg catgtggtaa cccagccccg atcctaaggg gtcggggctt 540

ttgctcactc agcacaggat ctgatcagga gacctccccc ccctgcttta caggggggcgg 600

gggtttaaaa attgcccaag gcctggcaaa taacctaggg gactaggttt tccttttttat 660

tttaaagttg tcaat 675

<210> 528

<211> 675

<212> DNA

<213> Cosavirus A

<400> 528

ccgtgcttta cacggttttt gaaccccaca ccggctgttt ggcgcttgca ggacagcagg 60

tttattttct tatgctcttt atttctagcc aacagggttc tatcctgttg ggcggagtga 120

tactcccgtt ccttcttgga cagattgcct ccacgatctt tgtggatctc aaggtgatca 180

agtcactggt aaatagagcg aaggttgagg aaacctgagg aatttccatg tggttttgcc 240

aggagttgta gcgatgctgt gtgtgtgtgc ggatttcccc tcatggcaac atgagcctca 300

caggccaaaa gccctgtccg aaaggaccca cacagtggag caatcccagc tccctcctac 360

aaagctttgt gagaatgaac tcacgtttat tcttctttat tctctgttta catcaggccc 420

caaagatgtc ctgaaggtac cttgtgtatc tgggcatgag caccatcaac tacccggact 480

tgcatttcgg tgcagacaca tgtggttacc cagcccctct gctttggcag aggggctttt 540

gctcgctcag cacgagatct gatcaggagc ccttcccagt gtgctttaca cctggcgggg 600

ggttaaaaat tgcccaaggc ctggcaaaat aacctagggg actaggtttt ccttttatta 660

acaatgtctg tcatt 675

<210> 529

<211> 660

<212> DNA

<213> Malagasivirus B

<400> 529

ctttattttc ttatgtaact cttctttta agttttattt tgcctacttg tgagcttatg 60

cgggaccact gtcttagaca accccacatt tgtcatgagt aagtacacgc aaccattacg 120

attacttttt aaccgtctga ccttttgata acaactgaag ttaggcgtga aacatgcatt 180

tataccaaag tagccccgca tttccccact acggtggggg ggctacccta ctggctttgg 240

aactgtagcc attatgtgtt gcctggcttt caggatctca caacacaaca gttctctcac 300

aatggaatat gggtgagatt gcagtgacat gaacaagtat ctagtagtac atagactcaa 360

gcctagttgc ctgcggaaca acatgtggta acacatgccc cagggtccaa aagacaaggg 420

ttaacagccc cttctaggtg tctgtgtgtg aagaatactt tagtagtgtt gttatgatct 480

cacctgttag tacagaatga gtatggcttg gtgaaggatg tcctacaggt acccattata 540

tggatctgag taggagacca ctagtggtgg ctttaccgcc aggtgagtgg tttaaaaagc 600

gtctagccaa gccaacagca ctagggatag tgctttctat attttatatt ttcagtgtat 660

<210> 530

<211> 654

<212> DNA

<213> Mosavirus sp.

<400> 530

ccccccccctc aaattgcaac gatatagcta atggcgagat tgagatgcta tatcacctct 60

tctaagttat agacctcatc tgattgataa ggacgtaatt tggtcgaaac cgcttggaat 120

aagaccgatg cgcgtagtca tgatgatgat gtaagatcta ggaacttatc caatctgctt 180

atgtctatgt aagtagaggg gcaggcctca ttgccctaat tctttctacc gagtatctgc 240

tagggtttct agcggcttga atacttggat tgagggatac aagatactac tgatcgattg 300

tcgattggga aacttgtaga tacttcaaag ctaccagtag cgtggactca cagccagcgg 360

actacccctc atggtaacat gagcctctgg gcccacaagg cacgtcgcaa gacctgtgag 420

acggcaaccc cagcctagct ttgttgagga aacaagcgat aacatgacat gagagaccgg 480

aaggattctt gtattgtgag ccgaaggatg gcctctaggt acctcatttt atgagatctg 540

aggaggtgct cttgagttgg tgctttacac tgacaacaca gagttaaaaa gcgtctaagc 600

tcacccggaa attgggaaat ttccgttatt tccattttgt ttgcaaagtc gttc 654

<210> 531

<211> 654

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

SVV sequence"

<400> 531

ctgggccctc atgcccagtc cttcctttcc ccttccgggg ggtaaaccgg ctgtgtttgc 60

tagaggcaca gaggagcaac atccaacctg cttttgtggg gaacggtgcg gctccaattc 120

ctgcgtcgcc aaaggtgtta gcgcacccaa acggcgcatc taccaatgct attggtgtgg 180

tctgcgagtt ctagcctact cgtttctccc ctatccactc actcacgcac aaaaagtgtg 240

ctgtaattac aagatttagc cctcgcacga gatgtgcgat aaccgcaaga ttgactcaag 300

cgcggaaagc gctgtaacca catgctgtta gtcccttcat ggctgcgaga tggctatcca 360

cctcggatca ctgaactgga gctcgaccct ccttagtaag ggaaccgaga ggccttcttg 420

caacaagctc cgacacagag tccacgtgat tgctaccacc atgagtacat ggttctcccc 480

tctcgaccca ggacttcttt ttgaatatcc acggctcgat ccagagggtg gggcatgatc 540

cccctagcat agcgagctac agcgggaact gtagctaggc cttagcgtgc cttggatact 600

gcctgatagg gcgacggcct agtcgtgtcg gttctatagg tagcacatac aaat 654

<210> 532

<211> 331

<212> DNA

<213> Teschovirus A

<400> 532

actagtcctt ggacttttgt tgtgtttaaa cacagaaatt taattacctg gccatgaatt 60

cattggatta acccttctga aagacttgct ctggcgcgag ctaaagcgca attgtcacca 120

ggtattgcac cagtggtggc gacagggtac agaagagcaa gtactcctga ccgggcaatg 180

ggactgcatt gcatatccct aggcacctat tgagatttct ctggggccca ccggcgtgga 240

gttcctgtat gggaatgcag gactggactt gtgctgcctg acagggtcgc ggctggccgt 300

ctgtactttg tatagtcagt tgaaactcac c 331

<210> 533

<211> 416

<212> DNA

<213> Teschovirus B

<400> 533

cttcctttta attcgtaact gataagtgat agtccttgga agctaggtat ttgttacgct 60

agttttggat tatcttgtgc ccaacatttg ttttcgaaca tatgttgtgt ttaaacacag 120

aaatctagtt tctttggtta tgagtttaat ggaatatcct tttgaaagac ttgccttggc 180

gcgggctaga gcgcaattgt caccaggtat tgcaccaatg gtggcgacag ggtacagaag 240

agcaagtact cctgactggg taatgggact gcattgcata tccctaggca tctattgaga 300

tttctctgga gcccaccagc atggagttcc tgtatgggaa tgcaggactg gacttgtgct 360

gcctgacagg gtcgcggctg gccgtctgta ctttgtatag tcagttgaaa ctcatt 416

<210> 534

<211> 664

<212> DNA

<213> Tottorivirus sp.

<400> 534

ccccttttacg taactgcaac ttaaagagta ccctactgca ttggatgtgt ggtaaacttt 60

tacgcacaca tttgtagtag tgttagttat gttctaccta atgagtatgc atgcacccgt 120

cgaaacacgc ttgtgataag ataggtgagt ccatgtgact aatctcatta agataaataa 180

gcaccctaca acgcacggca cgctcgtgtc ttccgtgcgg ggccgggaca acagcggcct 240

aaatcttcta ggtgaccacc atgcttttgg gactatggca ccactgtgga cgtgagtacc 300

tggcagtaag tctgtgaaaa gatggaaggt gtcccaagct atggggcgta tgcatatagc 360

ctgcggaaca aacaacggcg acgttgtccc cagggcccaa aaggcacgtg gataagatcc 420

acctatatgt ttaccccata gtgtaagtca ctggaagtcc tagtaatgga tgtctggagt 480

aaggctcacg gggtagggcg aaggatgccc agaaggtacc cgtaggtaac cttaagagac 540

tatggatctg atctggggac cggatggcgc catcaccatg acgtggaggc cggtttaaaa 600

aacgtctaag cccgaccaac aacctagggg actaggtttt ccttttttat tcatgtatga 660

cgtt 664

<210> 535

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Posavirus 1 sequence"

<400> 535

acatttcctt gcgtgcgcac ccgaaaattt attgaacttg gcttgaatca taagtaatgc 60

ttcttatagc ggacactttg agaatataat gcttgtatgg attaatgtat actgttttaa 120

ataacaaacc tagacacgca gttgcgttga tggttgtatc aatcacatat aagtgttgaa 180

ctcgtgttaa tcctcgatcg ctatatgttt gccgcctact tccaataaaa tagattacat 240

gcgcgtcatg cctttgtggg ttacctattg gcctctgaca aaaacaagtc gtaagatttg 300

tagcttcccg gtgtaaaaag ctgggcgcgg tctggctctc gtagggtgga aaggtccacc 360

aatggctggt tgagtgtaag ctccggtgtc ctggttgtcg caattccagg cgtcgtaata 420

acctatattg catctgactc taactcttgt ggctctactg tatctagttc ttgttctact 480

aactctaata atactactgg ctctaatact gaaaaactta catatgttaa tatagataat 540

atccttgatc ctgatatccc tcacgtcact gaagttcgcc gaaaacgaat ttcagatcat 600

atcattgaat ctcaaggatg tacttgctct gaacctacta taactcctca tgcgttttca 660

ttttctactc ttggc 675

<210> 536

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

A105-675 sequence"

<400> 536

ccacccacag caagaatgcc atcatctgtc ctcaccccca tttctcccct ccttcccctg 60

caaccattac gcttactcgc atgtgcattg agtggtgcac gtgttgaaca aacagctaca 120

ctcacgtggg ggcgggtttt cccgcccttc ggcctctcgc gaggcccacc cttcccttc 180

ctcccataac tacagtgctt tggtaggtaa gcatcctgat cccccgcgga agctgctcgc 240

gtggcaactg tggggaccca gacaggttat caaaggcacc cggtctttcc gcctccagga 300

gtatccctgc tagtgaattc tagtggggct ctgcttggtg ccaacctccc ccaaatgcgc 360

gctgcgggag tgctcttccc caactcaccc tagtatcctc tcatgtgtgt gcttggtcag 420

catatctgag acgatgttcc gctgtcccag accagtccag caatggacgg gccagtgtgc 480

gtagtcgtct tccggcttgt ccggcgcatg tttggtgaac cggtggggta aggttggtgt 540

gcccaacgcc cgtactttgg tgacaactca agaccaccca ggaatgccag ggaggtaccc 600

cgcctcacgg cgggatctga ccctgggcta attgtctacg gtggttcttc ttgcttccat 660

ttctttcttc tgttc 675

<210> 537

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

A110-675 sequence"

<400> 537

acctttgtgc gcctgtttta taccccctcc cccaactgta acttagaagt aacacacacc 60

gatcaacagt cagcgtggca caccagccac gttttgatca agcacttctg ttaccccgga 120

ctgagtatca atagactgct cacgcggttg aaggagaaag cgttcgttat ccggccaact 180

acttcgaaaa acctagtaac accgtggaag ttgcagagtg tttcgctcag cactacccca 240

gtgtagatca ggtcgatgag tcaccgcatt ccccacgggc gaccgtggcg gtggctgcgt 300

tggcggcctg cccatgggga aacccatggg acgctctaat acagacatgg tgcgaagagt 360

ctattgagct agttggtagt cctccggccc ctgaatgcgg ctaatcctaa ctgcggagca 420

cacaccctca agccagaggg cagtgtgtcg taacgggcaa ctctgcagcg gaaccgacta 480

ctttgggtgt ccgtgtttca ttttattcct atactggctg cttatggtga caattgagag 540

atcgttacca tatagctatt ggattggcca tccggtgact aatagagcta ttatatatcc 600

ctttgttggg tttataccac ttagcttgaa agaggttaaa acattacaat tcattgttaa 660

gttgaataca gcaaa 675

<210> 538

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

18-675 sequence"

<400> 538

cccacagcaa gaatgccatc atctgtcctc acccccaatt ttccctttc ttcccctgca 60

accattacgc ttactcgcat gtgcattgag tggtgcatgt gttgaacaaa cagctacact 120

cacatggggg cgggttttcc cgccctacgg cctctcgcga ggcccacccc ttccctcccc 180

ttataactac agtgctttgg taggtaagca tcctgatccc ccgcggaagc tgctcacgtg 240

gcaactgtgg ggacccagac aggttatcaa aggcacccgg tctttccgcc ttcaggagta 300

tccctactag tgaattctag cggggctctg cttggtgcca acctccccca aatgcgcgct 360

gcgggagtgc tcttccccaa ctcaccctag tatcctctca tgtgtgtgct tggtcagcat 420

atctgagacg atgttccgct gtcccagacc agtccagtaa tggacgggcc agtgcgtgta 480

gtcgtcttcc ggcttgtccg gggcatgttt ggtgaaccgg tggggtaagg ttggtgtgcc 540

caacgcccgt actttggtga cacctcaaga ccacccagga atgccaggga ggtaccccac 600

ctcacggtgg gatctgaccc tgggctaatt gtctacggtg gttcttcttg cttccacttc 660

tttcttctgt tcacg 675

<210> 539

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

A115-675 sequence"

<400> 539

acctttgtgc gcctgtttta tacccccccc aacctcgaaa cttagaagta aagcaaaccc 60

gatcaatagc aggtgcggcg caccagtcgc atcttgatca agcacttctg taaccccgga 120

ccgagtatca atagactgct cacgcggttg aaggagaaaa cgttcgttac ccggctaact 180

acttcgagaa acccagtagc atcatgaaag ttgcagagtg tttcgctcag cactaccccc 240

gtgtagatca ggccgatgag tcaccgcact tccccacggg cgaccgtggc ggtggctgcg 300

ttggcggcct gcctatgggg caacccatag gacgctctaa tacggacatg gtgcgaagag 360

tctattgagc tagttagtag tcctccggcc cctgaatgcg gctaatccta actgcggagc 420

acataccctt aatccaaagg gcagtgtgtc gtaacgggta actctgcagc ggaaccgact 480

actttgggtg tccgtgtttc cttttaattt ttactggctg cttatggtga caattgagga 540

attgttgcca tatagctatt ggattggcca tccggtgact aacagagcta ttgtgttcca 600

atttgttgga tttaccccgc tcacactcac agtcgtaaga acccttcatt acgtgttatt 660

tctcaactca agaaa 675

<210> 540

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

A73-675 sequence"

<400> 540

ttactccatt cagcttcttc ggaacctgtt cggaggaatt aaacgggcac ccatactccc 60

cccacccccc ttttgtaact aagtatgtgt gctcgtgatc ttgactccca cggaacggac 120

cgatccgttg gtgaacaaac agctaggtcc acatcctccc ttcccctggg agggcccccg 180

ccctcccaca tcctcccccc agcctgacgt atcacaggct gtgtgaagcc cccgcgaaag 240

ctgctcacgt ggcaattgtg ggtcccccct tcatcaagac accaggtctt tcctccttaa 300

ggctagcccc ggcgtgtgaa ttcacgttgg gcaactagtg gtgtcactgt gcgctcccaa 360

tctcggccgc ggagtgctgt tccccaagcc aaaccccctgg cccttcacta tgtgcctggc 420

aagcatatct gagaaggtgt tccgctgtgg ctgccaacct ggtgacaggt gccccagtgt 480

gcgtaacctt cttccgtctc cggacggtag tgattggtta agatttggtg taaggttcat 540

gtgccaacgc cctgtgcggg atgaaacctc tactgcccta ggaatgccag gcaggtaccc 600

cacctccggg tgggatctga gcctgggcta attgtctacg ggtagtttca tttccaatcc 660

ttttatgtcg gagtc 675

<210> 541

<211> 675

<212> DNA

<213> Kobuvirus sp.

<400> 541

ttactccatt cagcttcttc ggaacctgtt cggaggaatt aaacgggcac ccatacaccc 60

ccatcccctt tctgcaactt aagtatgtgt gctcgtgatc ttgactccca cggaatggat 120

cgatccgctg gagaacaaac tgctagatcc acatcctccc ttcccctggg aggaccttgg 180

tcctcccaca tcctcccccc agcctgacgt accacaggct gtgtgaagcc cccgcgaaag 240

ctgctcacgt ggcaattgtg ggtcccccct tcatcaagac accaggtctt tcctccttaa 300

ggctagcccc gatgtgtgaa ttcacattgg gcaactagtg gtgtcactgt gcgctcccaa 360

tctcggccgc ggagtgctgt tccccaagcc aaaccccctgg cccttcacta tgtgcctggc 420

aagcatacct gagaaggtgt tccgctgtgg ctgccagcct ggtaacaggt gccccagtgt 480

gcgtaacctt cttccgtctt cggacggtag tgattggtta agatttggtg taaggtccat 540

gtgccaacgc cctgtgcggg atgaaacctc tactgcccta ggaatgccag gcaggtaccc 600

cacccccggg tgggatctga gcctgggcta attgtctacg ggtagtttca tttccaattc 660

ttttatgtcg gagtc 675

<210> 542

<211> 675

<212> DNA

<213> Aichivirus sp.

<400> 542

ttactccatt cagcttcttc ggaacctgtt cggaggaatt aaacgggcac ccatacatcc 60

ccatcccctt tctgtaactt aagtatgtgt gcttgtaatc ttgactccca cggaatggat 120

cgatccgctg gagaacaaac tgctagatcc acatcctccc ttcccctggg aggaccttgg 180

tcctcccaca tcctcccccc agcctgacgt accacaggct gtgtgaagcc cccgcgaaag 240

ctgctcacgt ggcaattgtg ggtcccccct tcatcaagac accaggtctt tcctccttaa 300

ggctagcccc gatgtgtgaa ttcacattgg gcaactagtg gtgtcactgt gcgctcccaa 360

tctcggccgc ggagtgctgt tccccaagcc aaacccctgg cccttcacta tgtgcctggc 420

aagcatatct gagaaggtgt tccgctgtgg ctgccagcct ggtaacaggt gccccagtgt 480

gcgtaacctt cttccgtctc cggacggtag tgattggtta agatttggtg taaggttcat 540

gtgccaacgc cctgtgcggg atgaaatctc tactgcccta ggaatgccag gcaggtaccc 600

caccctcggg tgggatctga gcctgggcta attgtctacg ggtagtttca tttccaatcc 660

ttttatgtcg gagtc 675

<210> 543

<211> 675

<212> DNA

<213> Aichivirus sp.

<400> 543

actccattca gcttcttcgg aacctgttcg gaggaattaa acgggcaccc atacaccccc 60

atccccttttt ttgcaactta agtatgtgtg ctcgtaatct tgactcccac ggaatggatc 120

gatccgctgg agaacaaact gctagatcca catcctccct ccccccctggg aggacctcgg 180

tcctcccaca tcctcccccc agcctgacgt atcacaggct gtgtgaagcc cccgcgaaag 240

ctgctcacgt ggcaattgtg ggtcccccct tcatcaagac accaggtctt tcctccttaa 300

ggctagtccc gatgtgtgaa ttcacatcgg gcaactagtg gtgtcactgt gcgctcccaa 360

tctcggccgc ggagtgctgt tccccaagcc aaacccctgg cccttcacta tgtgcctggc 420

aagcatatct gagaaggcgt tccgctgtgg ctgccagcct ggtaacaggt gccccagtgt 480

gcgtaacctt cttccgtccc cggacggtag tgattggtta agacttggcg taaggttcat 540

gtgccaacgc cctgtgcggg atgaaacctc tactgcccta ggaatgccag gcaggtaccc 600

caccttcggg tgggatctga gcctgggcta attgtctacg ggtagtttca tttctaattc 660

tttcatgtcg gagtc 675

<210> 544

<211> 675

<212> DNA

<213> Aichivirus sp.

<400> 544

ttactccatt cagcttcttc ggaacctgtt cggaggaatt aaacgggcac ccatacaccc 60

ccacccccttt tttgcaactt aagtatgtgt gctcgtgatc ttgactccca cggaatggat 120

cgatccgctg gagaacaaac tgctagatcc acatcctccc ttcccttggg aggacctcgg 180

tcctcccaca tcctcccccc agcctgacgt accacaggct gtgtgaagcc cccgcgaaag 240

ccgctcacgt ggcaattgtg ggtcccccct tcattaagac accaggtctt tcctccttaa 300

ggctagtccc gatgtgtgaa ttcacattgg gcaactagtg gtgtcactgt gcgctcccaa 360

tctcggccgc ggagtgctgt tccccaagcc aaacccctgg cccttcacta tgtgcctggc 420

aagcatatct gagaaggtgt tccgctgtgg ctgccagcct ggtaacaggt gccccagtgt 480

gcgtaacctt cttccgtctt cggacggtag tgattggtta agatttggcg taaggttcat 540

gtgccaacgc cctgtgcggg atgaaacctc tactacccta ggaatgccag gcaggtaccc 600

caccctcggg tgggatctga gcctgggcta attgtctacg ggtagtttca tttccaattc 660

ttctatgtcg gagtc 675

<210> 545

<211> 675

<212> DNA

<213> Aichivirus sp.

<400> 545

tactccattc agcttcttcg gaacctgttc ggaggaatta aacgggcacc catacacccc 60

cacccccttt tctgcaactt aagtatgtgt gctcgtaatc ttgactccca cggaatggat 120

cgatccgctg gagaacaaac tgctagatcc acatcctccc ttcccctggg aggaccccgg 180

tcctcccaca tcctcccccc agcctgacgt atcacaggct gtgtgaagtc cccgcgaaag 240

ctgctcacgt ggcaattgtg ggtcccccct tcatcaagac accaggtctt tcctccttaa 300

ggctagcccc gatgtgtgaa ttcacattgg gcaactagtg gtgtcactgt gcgctcccaa 360

tctcggccgc ggagtgctgt tccccaagcc aaacccctgg cccttcacta tgtgcctggc 420

aagcatatct gagaaggtgt tccgctgtgg ctgccagcct ggtaacaggt gccccagtgt 480

gcgtaacctt cttccgtctc cggacggtag tgattggtta agatttggtg taaggttcat 540

gtgccaacgc cctgtgcggg atgaaacctc tactgcccta ggaatgccag gcaggtaccc 600

caccttcggg tgggatctga gcctgggcta attgtctacg ggtagtttca tttccaattc 660

ttttatgtcg gagtc 675

<210> 546

<211> 664

<212> DNA

<213> Aichivirus sp.

<400> 546

gcttcttcgg aacctgttcg gaggaattaa acgggcaccc atacaccccc acccccttt 60

ctgcaactta agtatgtgtg ctcgtaatct tgactcccac ggaatggatc gatccgctgg 120

agaacaaact gctagatcca catcctccct tcccctggga ggaccccggt cctcccacat 180

cctcccccca gcctgacgta tcacaggctg tgtgaagtcc ccgcgaaagc tgctcacgtg 240

gcaattgtgg gtcccccctt catcaagaca ccaggtcttt cctccttaag gctagccccg 300

atgtgtgaat tcacattggg caactagtgg tgtcactgtg cgctcccaat ctcggccgcg 360

gagtgctgtt ccccaagcca aacccctggc ccttcactat gtgcctggca agcatatctg 420

agaaggtgtt ccgctgtggc tgccagcctg gtaacaggtg ccccagtgtg cgtaaccttc 480

ttccgtctcc ggacggtagt gattggttaa gatttggtgt aaggttcatg tgccaacgcc 540

ctgtgcggga tgaaacctct actgccctag gaatgccagg caggtacccc accttcgggt 600

gggatctgag cctgggctaa ttgtctacgg gtagtttcat ttccaattct tttatgtcgg 660

agtc 664

<210> 547

<211> 675

<212> DNA

<213> Aichivirus sp.

<400> 547

tactccattc agcttcttcg gaacctgttc ggaggaatta aacgggcacc cactttcctg 60

tcctctcccc ttttctgtaa ctccaagtgt gtgctcgtaa tcttgactcc cgcggattga 120

ccgctccgct ggtgaacaaa ctgctaggtc atctcctccc cacccttggg cgtccttccg 180

ggcgtccaca ccctcccccc agcctgacgt gtcacaggct gtacaaagac cccgcgaaag 240

ctgctaacgt ggcaattgtg ggtcccccct ttgtaaagga accgagtctt tctcccttaa 300

ggctagaccc ctgtgtgaat tcacaggtgg caactagtgg ttccactgca tgctcccgac 360

ctcggccgcg gagtgctgtt ccccaagtcg taacactgac cttcacttat gtgcctggca 420

agcatatctg agaagatgtt ccgctgtggc tgccaaacct ggtaacaggt gccccagtgt 480

gcgtagtctt cttccgtctt cggacggtag gtgttaggta aagatgcggc gtaaggttca 540

agtgccaacg ccctggaagg gatgaccctt ctactgccct aggaatgccg cgcaggtacc 600

ccaggttcgc ctgggatctg agcgcgggct aattgtctac gggtagtttc atttccctct 660

tcttccactg gcatc 675

<210> 548

<211> 675

<212> DNA

<213> Aichivirus sp.

<400> 548

actccattca gcttcttcgg aacctgttcg gaggaattaa acgggcaccc actttcctgt 60

cctctccccc tttctgcaac tcaagtgtgt gctcgtaatc ctgactccca cgggttgacc 120

gccccgttgg tgaacaaaca gctaggtcat tccctcccta cccctgggcg ccatttcagt 180

ggcgttcata tcctcccccc agcctgacgt gtcacaggct gtgcaaagtc cccgcgaaag 240

ctgctcacgt ggcaattgtg ggtcccccct ttgtgaagga accgagtctt tctcccttaa 300

ggctagaccc ctgtgtgaac tcacaggtgg caactagtgg ttccactgca tgctcccgac 360

ctcggccgcg gagtgctgtt ccccaagtcg tgacactgac ctccacttat gtgcctggca 420

agcatatctg agaagatgtt ccgctgtggc tgccaaacct ggtaacaggt gccccagtgc 480

gtgtagtctt cttccgtctc cggacggtaa gtgtgtggta aagatgcggc gtaaggttca 540

agtgccaacg ccctggaagg gatgaccctt ctactgccct aggaatgccg cgcaggtacc 600

ccaggttcgc ctgggatctg agcgcgggct aattgtctac gggtagtttc atttccctct 660

cttttcactg gcatc 675

<210> 549

<211> 498

<212> DNA

<213> Pestivirus sp.

<400> 549

gtataagggt tgggaacctt gtaccaagct acctctgcca ttcagtattt gggagtagaa 60

gtagatgtgt ttacaaactc acacgtgtgg gggcggggat agactgtgcc agcggtcgtg 120

taccagcacc tacgcatacg tgtggactgc gaaccaggag agcacctagg tctgacaagc 180

tgtgagaaca cagtagtcgt cagtgagtca gctggtaagg atcacccacc tggatactca 240

cgtggacgag ggagtttccc agtcagaaac ctacaccaga ggaggggtcc tctggagaca 300

tggatggtct gagtaacaga ctatctactg gggtgtgctg cctgacaggg tctcggctga 360

tagcctggct agcagtataa aaatcagttg aattggcata tgagttgtga acatctagta 420

aacaatgaaa gacaaaaaca aaaaatgagc ataataaaaa aattgtacaa tccactactc 480

aggtgtggct gcagactt 498

<210> 550

<211> 576

<212> DNA

<213> Kobuvirus sp.

<400> 550

tttgaaaagg gggtgggggg gcctcggccc cctcaccctc ttttccggtg gccattcgcc 60

cgggccaccg ttactccact ccactccttc gggactggtt tggaggaaca caacagggct 120

tcccatccct gtttaccctt tattccatca tcctttcccc aagtttaccc tatccacacc 180

ccactgactg actcctttgg attttgacct cagaatgcct atttgacctc ccactcgcct 240

ctccctttc ggattgccgg tggtgcctgg cggaaaaagc acaagtgtgt tgcaggctac 300

caaactccta cccgacaaag gtgcgtgtcc gcgtgctgag taatgggata ggagatgcct 360

acaacaggct cgcccatgag tagagcatgg actgcggtgc atgtgacttc ggtcaccacg 420

ggcatagcat tgctcacccg tgaatcaagt catcgagatt tctctgacct ctgaagtgca 480

ctgtggttgc gtggctggga atccacgctt gaccatgtac tgcttgatag agtcgcggct 540

ggccgactca tgggttaaag tcagttgaca agacac 576

<210> 551

<211> 675

<212> DNA

<213> Kobuvirus sp.

<400> 551

cctacccaag ggttacatgg gaccatattc ctcctcccct gtaactttaa gttttgtgcc 60

cgtattcttg actccaggcg gatgttgtgt cgcccgtcct gtgaacaaac agctagacac 120

tttcctcccc tccctctggg ctgctccggc agtccactcc ctccccccag cgtaacatgc 180

cccgctggag tgatgcacct ggaagtcgtg gacgtgggtt agtaacttcg gtgaaaaccc 240

actataatga caactggttg acccccacac tcaaaggact cgagtctttc tcccttaagg 300

ctagcccggc cacatgaatt tgcagctggc aactagtgag tccaccatgt cccgcaacct 360

cggctgcgga gtgctgttcc ccaagcgtat gccttccttc tgtaagagtg cgcctggcaa 420

gcacatctga gaagtcgttc cgctgcgtcg tgccaacctg gcgacaggtg acccagtgtg 480

cgtagacttc ttccggattc gtccggctct tctctaggaa acatgcgtgt aaggttcatg 540

tgccaaagcc ctgcgcgcgg tgttcttcta ctgccctagg aatgtgccgc aggtacccct 600

acttcggtag ggatctgagc ggtagctaat tgtctacggg tagtttcatt tccatcttct 660

cttcaggtcg acatc 675

<210> 552

<211> 675

<212> DNA

<213> Kobuvirus sp.

<400> 552

gaccttctgg tacttcttcg cctgggtcac aaaagcgaag aacctgcctc tctaacgcca 60

gacgagcggc attaaacttg aacttctggc actctccact ctccctttc cctgtccctt 120

tccccactgc gctctcaagg tcgcgcaatc ctgggactag cccagtttta aaagttcctg 180

gcaccctttg cccctctagg cccttaaggt aggaactgac cttgtgctgt gatctcggtg 240

cgggagtgct accacgtagt catcgtaagc ctcgtttctg gttctgccct ggcaaggcta 300

cagagtaccg tgttccgctg tggatgccat ccgggtaacc ggacccccag tgtgtgtagc 360

ggtatgttca cggtccgccg tgttcaccag attcctgacc tggctttgct agaaatggtg 420

tgtgcccaat ccctgtgacc agtatcaatt acatcaccta ggaatgctag gaaggtaccc 480

cagtcctgag ctgggatctg atcctaggct aattgtctac ggtgatgctc cttttatttt 540

cttacaactg ctattgactg tctgattgct gattctgctc ttgtgctctt ctgctctggc 600

tcattctcaa gggttctctt tgtccaagat cctttggttc tctccttgtt ccacttgcca 660

ctgccaacgc ttgtc 675

<210> 553

<211> 369

<212> DNA

<213> Pestivirus sp.

<400> 553

gtatacgcag ttagttcatc ctgtgtatac agattggaga ctctaaaaac aacgattcgg 60

aataggggcc cgcggcgaag accgaagaca ggctaaccat gccgttagta gggctagcac 120

caaaacgcgg gaactagaca cttaggagag tggtctggct actctaagag gtgagtacac 180

cttaaccgtc aagggttcta ctcctcagtt gaggactaga gatgccctgt ggacgggggc 240

atgcccaaga gttagcttag ccggggcggg ggttgttccg gtgaaagtag caatattgac 300

cacactgcct gatagggcgg agcaggcccc ctaggtagtc tagtataaaa tgtctgctgt 360

acatggcac 369

<210> 554

<211> 399

<212> DNA

<213> Pestivirus sp.

<400> 554

tacgcggggt ataacgacag tagttcaagt gtcgttatgc atcattggcc ataacaaatt 60

atctaatttg gaatagggac ctgcgacctg tacgaaggcc gagcgtcggt agccattccg 120

actagtagga ctagtacaaa taggtcaact ggttgagcag gtgagtgtgc tgcagcggct 180

aagcggtgag tacaccgtat tcgtcaacag gtgctactgg aaaggatcac ccactagcga 240

tgcctgtgtg gacgaggaca tgtccaagcc aatgttatca gtagcggggg tcgttactga 300

gaaagctgcc cagaatgggt agttgcacat acagtctgat aggatgccgg cggatgccct 360

gtattttgac cagtataaat attatccgtt gtaaagcat 399

<210> 555

<211> 123

<212> DNA

<213> Pestivirus sp.

<400> 555

gcagatatcg gtggtggacc tgggggttgg gctcaccgtg ccccttcatg gggtagacct 60

cactgcttga tagagtgccg gcggatgcct caggtaagag tataaaatcc gttgttcact 120

aac 123

<210> 556

<211> 382

<212> DNA

<213> Pestivirus sp.

<400> 556

gtatacgagt ttagctcaat cctcgtatac aatattgggc gtcaccaaat atagatttgg 60

cataggcaac accccgatgc gaaggccgaa aagggctaac catgccctta gtaggactag 120

caaaaaatcg gggactagcc caggtggtga gcttcctgga tgaccgaagc cctgagtaca 180

gggcagtcgt caacagttca acacgcagaa taggtttgcg tcttgatatg ctgtgtggac 240

gagggcatgc ccacggtaca tcttaaccta tccgggggtc ggataggcga aagtccagta 300

ttggactggg agtacagcct gatagggtgt tgcagagacc catctgatag gctagtataa 360

aaaactctgc tgtacatggc ac 382

<210> 557

<211> 373

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Classical swine fever virus sequence"

<400> 557

gtatacgagg ttagttcatt ctcgtatgca tgattggaca aattaaaatt tcaatttgga 60

tcagggcctc cctccagcga cggccgaact gggctagcca tgcccacagt aggactagca 120

aacggaggga ctagccgtag tggcgagctc cctgggtggt ctaagtcctg agtacaggac 180

agtcgtcagt agttcgacgt gagcagaagc ccacctcgat atgctatgtg gacgagggca 240

tgcccaagac acaccttaac cctagcgggg gtcgctaggg tgaaatcaca ccacgtgatg 300

ggagtacgac ctgatagggt gctgcagagg cccactatta ggctagtata aaaatctctg 360

ctgtacatgg cac 373

<210> 558

<211> 222

<212> DNA

<213> Pegivirus sp.

<400> 558

tcagggttgg taggtcgtaa atcccggtca ccttggtagc cactataggt gggtcttaag 60

agaaggttaa gattcctctt gtgcctgcgg cgagaccgcg cacggtccac aggtgttggc 120

cctaccggtg ggaataaggg cccgacgtca ggctcgtcgt taaaccgagc ccgtcaccca 180

cctgggcaaa cgacgcccac gtacggtcca cgtcgccctt ca 222

<210> 559

<211> 527

<212> DNA

<213> Pegivirus sp.

<400> 559

cccggcactg ggtgcaagcc ccagaaaccg acgcctattt aaacagacgt tatgaaccgg 60

cgccgacccg gcgaccggcc aaaaggtggt ggatgggtga tgccagggtt ggtaggtcgt 120

aaatcccggt catcttggta gccactatag gtgggtctta agggttggtc aaggtccctc 180

tggcgcttgt ggcgagaaag cgcacggtcc acaggtgttg gccctaccgg tgtgaataag 240

ggcccgacgt caggctcgtc gttaaaccga gcccattacc cacctgggca aacaacgccc 300

acgtacggtc cacgtcgccc tacaatgtct ctcttgacca ataggctttg ccggcgagtt 360

gacaaggacc agtgggggct gggcgacggg ggtcgtatag gaagaaaaat gccacccgcc 420

ctcacccgaa ggttcttggg ctaccccggc tgcaggccgc cgcggagctg gggtagccca 480

agaaccttcg ggtgagggcg ggtggcattt ttcttcctat accgatc 527

<210> 560

<211> 402

<212> DNA

<213> Pegivirus sp.

<400> 560

tggtcacctt ggtagccact ataggtgggt cttaagagaa ggttaagatt cctcttgtgc 60

ctgcggcgag accgcgcacg gtccacaggt gttggcccta ccggtgtgaa taagggcccg 120

acgtcaggct cgtcgttaaa ccgagcccat ttcccgcctg ggcaaacgac gcccacgtac 180

ggtccacgtc gccctttaa tgtctctctt gaccaatagg ttcatccggc gagttgacaa 240

ggaccagtgg gggccggggg tcacagggat ggaccctggg ccctgccctt cccggcgggg 300

tggggaaagc atggggccac ccagctccgc ggcggcctgc agccggggta gcccaagaac 360

cttcgggtga gggcgggtgg cattttтctt cctataccga tc 402

<210> 561

<211> 531

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hepatitis GB virus A sequence"

<400> 561

gccgggtgga aggcccggaa ccgccccacc acctcaacta ggtggtaagg gtacgtctat 60

cggtccggct ggcccgaaag gcggtggatc ctgtgtgtta gggttcgtag gtggtaaatc 120

ccagcacagg tggtaatcgc tatagggcag gcttatcccg gtgaccgctt ccctggatcc 180

tggagcgggt cgtggcggca cggtccacag gagtggggcc tccggtgtga ataagccctc 240

gtctggagca tcagacgtta aactgagacg tcccgaagag atcggaacga cgccccacgt 300

atggcaacgc cgcttaaaac ccttcgggga cagctatgcg ggttgacaat gccagtgggg 360

ggccgggccc actattgttg tgggctccga gttcctctag ggatggccga aaggcagcca 420

tggggccacc caggcggcgc cgtgctacag gcggcaaggg gaaaaatcct tcgggtgacc 480

ccgggtggca ttccctccct tagcagcatg agtgtggtgg tagctgcaac c 531

<210> 562

<211> 520

<212> DNA

<213> Pegivirus sp.

<400> 562

ggggaatctc acccccgtc cggttccgga agaatcggaa accgacaccc tgaccaatca 60

ttcttgatca tagagtggat gttagtgaaa gccagacgaa agccggcgga tgggtggtga 120

cagggttggt aggtcgtaaa tcccggccac cctggtaccc ggtataagtt gggcggaagc 180

tgactgaagc tccgtgctct tttctgtgcg ttcttggtgc acggtccaca ggtgacgcct 240

ataccggtgt gaataatagg ccgactcgag cggagtcgtt aaactgagaa cctccatacg 300

gatggcaact tggcttgcgt acggggacgc cgctaaagtc acagtgggtt aagtccggcg 360

ggttgacaac cccagcaagg cgagggggtc ctattgttgg actctgccag ttcccggtgg 420

aggtaggcat ggggtggccc agctccgcgg cgcgctacag ccggggtagc ccaaaatccg 480

aaaggtgagg gcgggccaca tgtccgaaat ttagtcaagc 520

<210> 563

<211> 524

<212> DNA

<213> Pegivirus I

<400> 563

agaatggtct aagtggttgc caccgtggtc cgaaggggag gaggacctac gctgccaggg 60

ttggcaggtc gtaaatcccg ggtgtaggag atccctcctt gttaggactg ctggtagctg 120

gggggtcggt gacccccctgg gcaaccgcca aacccggacg accgggtggc ggctccatgt 180

tggcacggtc cacaggtgtg aaccctaccg gtgtgaataa gggttggtgg ttgcggtcca 240

ccttaaacgt agtatgcatt gggcttggta aaacaccgct cgtagtacgg aacgccgcct 300

ttaaagacac agtaggcgta gccggcgggt tgacaatcca tacgggggggt ggggtgtggt 360

catggatctg tccacaccac cttcatgcgg ccctctaagc aagccatacc gggggggaggc 420

gcgcggcacc gcactgccgg gcaaggggaa gaaccttcgg gtgacccccc cccaaccacc 480

gtccgatcaa tgctaatgtt gcgtttaggc gtgacaccgg caca 524

<210> 564

<211> 537

<212> DNA

<213> Pegivirus K

<400> 564

agaatggtgt gatccgtcgc cgctccagcg gaaagcgggc gggatctagt ggttagggtt 60

gttcgcgtaa atcccacact agtggtacgc tcgtataacg tgggagcagc cggtggggtc 120

gacccccccac ctggcggctg ctgagcaccg gacgaagcgc ggggggtgaa cgctaacccg 180

cggcccgggc tgccaacgtt aggcacgtct tggctggaag acgttaaaca cagggccccc 240

cctcaaccct gatccgaggc cagagaccaa ggtacgccgc ccttttaaag gcgttactcg 300

tccaatagga tctctccggc gggttgtcaa accttgctgg ccctggtgat ggttacggga 360

gggggtgggg cggggagtag aagccccgcc cggcatgggg gtaccaagct cggcacgccc 420

agcacgcgtg gcgtagggga aaaatccttc gggtgacccc tggtaccata aagtaattaa 480

catgagcatg ccgctagggt gtgctttttc ttccttcctt gggaaggcgg tggcacc 537

<210> 565

<211> 617

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Theiler's disease-associated virus sequence"

<400> 565

tgataccgtg tcccggtacg acctcgcgcg tccccaagct cgccctgagg ggggagcgta 60

agggcgcgta gtggggtagc cccccaaacc gagccaccct agtgagtgac tttagaatgg 120

ttagggagac taccgccttc gctgtttggg gacctaatga tccgcgtgcc agggttcttc 180

gggtaaatcc cggcgcggtg ttttgggttc agggcagtag gggcagacgg gccagcagtc 240

gctggttcct ggtaccacca ccctatccgg acgacctccc tcacgaaagg tcgccacggt 300

ctgtggctcg acgacgccta taattcagtc cgaggggcgc agccctcgtt aaacttaggc 360

aaggttcctc gccattgatt tggccagggg tttaagtgaa cgccgccctt ttaatgttta 420

atagggttct ttcccggcgg gttgacaaac acttccctgg gctcttcgtt ggcctcggtt 480

ccttgatgct tcggcaccca tgagcgcaca gggggggggac cctgcgacag tccgccaaga 540

ggaaaatcct tcgggtgacc tcgtgcgcaa cccaatccct tcttcttcca catggcgtgt 600

ctgtggtgca tgctgtg 617

<210> 566

<211> 349

<212> DNA

<213> Pegivirus sp.

<400> 566

ggacttcggt ccccctgtta ctctgcgagc caccgcagag ccagggttgg tacgcccgag 60

gtgttagacc ccggccgaaa gctcctaacc atggggttag taggacgtgg taaatgccac 120

tgaggggttg gagagctggt agagcgagta agtcggcgta aggcccgagt acgggcctcc 180

cagcccgggt cagcctaaac ctggctgtga tacccggtgc atggagggcg tgtcccaacg 240

ctcgatcgct gtagggtggg tccctgcagt tgggtgtggc taccctgctc gtactgcttg 300

atagagtccc ggcggacgga ccagctctcg tcagtccgtg gagttgcac 349

<210> 567

<211> 327

<212> DNA

<213> Pegivirus sp.

<400> 567

aactgttgtt gtagcaatgc gcatattgct acttcggtac gcctaattgg taggcgcccg 60

gccgaccggc cccgcaaggg cctagtagga cgtgtgacaa tgccatgagg gatcatgaca 120

ctggggtgag cggaggcagc accgaagtcg ggtgaactcg actcccagtg cgaccacctg 180

gcttggtcgt tcatggaggg catgcccacg ggaacgctga tcgtgcaaag ggatgggtcc 240

ctgcactggt gccatgcgcg gcaccactcc gtacagcctg atagggtggc ggcgggcccc 300

cccagtgtga cgtccgtgga gcgcaac 327

<210> 568

<211> 436

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GB virus C/Hepatitis G virus sequence"

<400> 568

cccccggcac tgggtgcaag ccccagaaac cgacgcctat ctaagtagac gcaatgactc 60

ggcgccgact cggcgaccgg ccaaaaggtg gtggatgggt gatgacaggg ttggtaggtc 120

gtaaatcccg gtcaccttgg tagccactat aggtgggtct taagagaagg ttaagattcc 180

tcttgtgcct gcggcgagac cgcgcacggt ccacaggtgt tggccctacc ggtgggaata 240

agggcccgac gtcaggctcg tcgttaaacc gagcccgtta cccacctggg caaacgacgc 300

ccacgtacgg tccacgtcgc ccttcaatgt ctctcttgac caataggcgt agccggcgag 360

ttgacaagga ccagtggggg ccgggggctt ggagagggac tccaagtccc gcccttcccg 420

gtgggccggg aaatgc 436

<210> 569

<211> 393

<212> DNA

<213> Pegivirus sp.

<400> 569

agaatgggga gttaactcct ggcactggcc cgaagcatga actgatcgcg gtggcagggt 60

tcttcgggta aatcccggcc gcgtgttgtg attgtgttag ggcaggtgac agtcggcagg 120

gtcgaccccc tgcttcagga ccactgtctt cctggacgac cgttgctgaa aaagggccgc 180

cacggtctgt agctcgccga cgcttctaat tcaggccgga ggaccacgct ccgtaatcga 240

gcccaagtac tcaaacccca gcacccctgg gtcacgccct acgccgccct tttaacgctt 300

cggctaatag ggtctatccg gcgggttgac aaagggcgca gggttacctg gtactacgag 360

cttgggtgtc cctgggagta atcccagggt gcc 393

<210> 570

<211> 343

<212> DNA

<213> Flavivirus sp.

<400> 570

atataaatcc cagtttggtt aaacctattt caaggcttaa gttgtttatt attttatcgc 60

cgctcgtgac tataaagttg cctagcggag agagataaag aagaaggagt tcaaggctca 120

gggcagggcg caagttccct ggtccctagg ccgctcgcag gaaggaggag tgaagaagaa 180

gaaagagaag gagaggacca ccgccgaaag aaggcaggtg cctcacaaga gggccaacca 240

gcgtgttgga ccagtggcca acgccggacg gcgtggtggc ctgctgggac gcctggggat 300

tggatggagt gccttcctac aggaagacat cgttcaagcc atc 343

<210> 571

<211> 104

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bussuquara virus sequence"

<400> 571

agtatttctt ctgcgtgaga ccattgcgac agttcgtacc ggtgagtttt gacttaacgc 60

agtgagaaaa gttttcgagg aaagacgaga agcgaattct ctga 104

<210> 572

<211> 106

<212> DNA

<213> Zika virus

<400> 572

agttgttgat ctgtgtgagt cagactgcga cagttcgagt ctgaagcgag agctaacaac 60

agtatcaaca ggtttaattt ggatttggaa acgagagttt ctggtc 106

<210> 573

<211> 150

<212> DNA

<213> Yokose virus

<400> 573

agtaaatttt gcgtgctagt cgctgagcgt cagaccgcaa agtgagtttt tagtgatcta 60

aagtgaggag ttattcttac tgtcatcaaa cactacaaat aaacacgttg aaattatttc 120

cggaagaaca actgtccgga atcaaagacg 150

<210> 574

<211> 118

<212> DNA

<213> Wesselsbron virus

<400> 574

agtatattct gcgtgctaat cgttcgacgt tagtccgtgg agtgagcttc tattagagtc 60

gttaacacgt ttgaataatt tctactgaaa ggagtagaag aaaggagatt cattccca 118

<210> 575

<211> 389

<212> DNA

<213> Hepacivirus sp.

<400> 575

acctccgtgc tatgcacggt gcgttgtcag cgttttgcgc ttgcatgcgc tacacgcgtc 60

gtccaacgcg gagggattct tccacattac catgtgtcac tccccctatg gagggttcca 120

ccccgcccac acggaaatag gttaaccata cctatagtac gggtgagcgg gtcctcctag 180

ggcccccccg gcaggtcgag ggagctgaaa ttcgtgaatc cgtgagtaca cggaaatcgc 240

ggcttgaacg tcatacgtga ccttcggagc cgaaatttgg gcgtgcccca cgaaggaagg 300

cgggggcggt gttgggccgc cgcccccttt atcccacggt ctgataggat gcttgcgagg 360

gcacctgccg gtctcgtaga ccataggac 389

<210> 576

<211> 445

<212> DNA

<213> Hepacivirus B

<400> 576

accacaaaca ctccagtttg ttacactccg ctaggaatgc tcctggagca ccccccctag 60

cagggcgtgg gggatttccc ctgcccgtct gcagaagggt ggagccaacc accttagtat 120

gtaggcggcg ggactcatga cgctcgcgtg atgacaagcg ccaagcttga cttggatggc 180

cctgatgggc gttcatgggt tcggtggtgg tggcgcttta ggcagcctcc acgcccacca 240

cctcccagat agagcggcgg cactgtaggg aagaccgggg accggtcact accaaggacg 300

cagacctctt tttgagtatc acgcctccgg aagtagttgg gcaagcccac ctatatgtgt 360

tgggatggtt ggggttagcc atccataccg tactgcctga tagggtcctt gcgaggggat 420

ctgggagtct cgtagaccgt agcac 445

<210> 577

<211> 379

<212> DNA

<213> Hepacivirus I

<400> 577

cagggtttcg accctggccc ggatacctat cgccttacgc cgaaaggtaa cgagtaggag 60

tcgggtcccc aggcccttac cgccaccaag ccaggtgggg aggtatggga gccgggggggt 120

gcagctggta gctccatggg ggacgccccg tgagcggatg ctgcatcgat accgggttag 180

ctctctggga gagcggcact tgacaccacg aatccgggaa ccggacaatc gccggcgtgg 240

gacgcgttgc ctccgtggcc gagcaatttg gcatgcccgt ggtgaagagt gatggtgggg 300

gggggccccc cttccagtac cgtactgcct gatagggtct tgcctcaagc ccagagagtc 360

gaggctgaaa accgccatc 379

<210> 578

<211> 557

<212> DNA

<213> Hepacivirus J

<400> 578

gccgctcccg aaagggagtc cggcgcgtca tcccactccg aggagtgggg tggcgtcccc 60

gtgtgccggg gaaccatgaa gcctaagggc atccacattt tagaatgaac ttgaagcttc 120

gtttcgctgg ccggaaagtc ctgggttccc atggccaggg ttccgcaggt gggtaaatcc 180

cggtggggtt ccatccagga tatacggcag gcgggcgtag tccggcggtt cggacgacgt 240

gtgggtcgcc tacggtggat tgttcacagg atgggcactc cggtgtgaat aggccccgtc 300

agggtgcgct gacgttaaac tcaggccttg cctggtgttc ggggaggatt gcagggccac 360

gccgcctcta agggccgtat ggcacagtac ttcttcgggc gggttgtcaa ggccctccaa 420

cgcgacacca gtgcctcggc aggcatgggg ccacccagct cggcgtcccg cacacagacg 480

gcgtagggga aaatcagcaa tgtgaccccg ggtggcattt tccttctctc tacttccatg 540

catgatcaac cgcaatc 557

<210> 579

<211> 280

<212> DNA

<213> Hepacivirus K

<400> 579

gggaacaatg gtccgtccgc ggaacgactc tagccatgag tctagtacga gtgcgtgcca 60

cccattagca caaaaaccac tgactgagcc acacccctcc cggaatcctg agtacaggac 120

attcgctcgg acgacgcatg agcctccatg ccgagaaaat tgggtatacc cacgggtaag 180

gggtggccac ccagcgggaa tctgggggct ggtcactgac tatggtacag cctgataggg 240

tgctgccgca gcgtcagtgg tatgcggctg ttcatggaac 280

<210> 580

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Icavirus sequence"

<400> 580

cgaagtttaa gctaagcacc ctcgggcgtt cccggattat gtgatcacat caatttgatg 60

gctggtcacc acgcaacgcc tggagagata ctcttacttt tctcttaaga tccccggtca 120

tttgacgctt gtaggatgat agggttattt tccactataa atactttcat actcttggat 180

gttctatatc caagacggga ggacctaccc cgtacccctt agaggtgaga tgccaagaac 240

aggccctttc tgttctctcg acaatggcat cataggcaac aagcatcaca ccaagattgc 300

taagttttgt taagagttct tcaagctata gggtggctgt agcgaccttc tgatgcctgc 360

ggatttcccc acggagcgat ccgtgccaca ggggccaaaa gccacggcta acgcccatca 420

ggagcggcac ttaccccgtg ccccacccctt gaaacttgaa tgttcacact ggcttctctc 480

ggctttctga actgtctgct tgttggggcc ccgaaggatg ccctggaggt accccatttt 540

atgggatctg accaggggac acctcagctc tctaagttgc tggtgtttaa aaaacgtcta 600

agggccccca ccccttaggt ggagggatcc acctttcctt tatttttaa aactctttta 660

tggtcacaat tgttt 675

<210> 581

<211> 113

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Antarctic penguin virus A sequence"

<400> 581

ctaggagact acgcagtggg ataagatgac tatgatgtcg tacgggcaga agccagtaca 60

gtcgaagtcg agaccgacgt cgaggatttg actctgcctg acctagtgcc atc 113

<210> 582

<211> 214

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Forest pouched giant rat arterivirus sequence"

<400> 582

tattggatcc gcctccgggc aaaggttact ttcttgtacc tcggcttagc cacagggtga 60

ccccttgtac gtaggggccc cgacgtagca ctggtctgac aacaccttct cggcatttca 120

ccttctgccc gctcttccgg gcggtggtgt caagaagcag cagtgctctt ctcttttctt 180

cctgcagttc accgagccct acgggggga ggtg 214

<210> 583

<211> 561

<212> DNA

<213> Avisivirus sp.

<400> 583

cattcccctt tccccagcca tgggttaaat ggcccctcac caggttcggt gctgtctagg 60

cttccagtaa agaagtcaac cgagcattga acaaaacctc agtgggtatg gtagttaacc 120

ccgtccactg gacaactttg tcctcttaaa agtggatcaa tccaccccaa ctcccccccct 180

agccacctga gccatggtgg atagcagtga cgaaactagg gaccccaata cctctagtgc 240

caagagaatt ccccccctcgc gagaggtgct cttgggcccg aaaggctagt tggcagggtg 300

aagtgaagga agctgctagc gtggcaacct taagcgtagc ccgaagctga ccttagaggt 360

taaccctagt ggaccactgg atgaagctgt ggaggtggtg gataggaaag ttggccactt 420

gtgagtagat gcccagaagg cataaggctg atctggggcc agtgactata ccgttccggt 480

aaacctggta taaaaaccat gaaagcaagt gggtttaaaa tttcttctaa ttccttcatt 540

tcagtagtga taactggcag a 561

<210> 584

<211> 134

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Avian paramyxovirus penguin sequence"

<400> 584

accaaacaag gactagataa cccacgtgac cgttaactgg aaaataagat gttgtagggg 60

cgacctagtt ggaattcgac cccggctccg aaacctctaa ttgtggttat tggcagtcta 120

gtctacttct aacg 134

<210> 585

<211> 121

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Newcastle disease virus sequence"

<400> 585

accaaacaga gaatctgtga ggtacgataa aaggcgaaga agcaatcgag atcgtacggg 60

tagaaggtgt gaaccccgag cgcgaggccg aagctcgaac ctgagggaac cttctaccga 120

t 121

<210> 586

<211> 302

<212> DNA

<213> Bat Hp-betacoronavirus

<400> 586

ttaagcttcg gcttgttgca taggaccgga aaggtactat ctaccctaac tcttgtagtt 60

agactctcta aacgaacttt aaaactggtt gtgtccttca gtagtctgta tggccattgg 120

aggcacaccg gtaattatca aatactaaga agattcatag tacatccttg tctagctttt 180

ggttggcagt gagcctacgg tttcgtccgt gtcgctcaca attatccaca cagtaggttt 240

cgtccgctgt ggttgagttg ctagtccgtt gctgtttcgt cagccatcta caactcgaca 300

cc 302

<210> 587

<211> 458

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Basella alba endornavirus sequence"

<400> 587

ggaatatggc taatcggctt attctaatca aacgcaaaag acttatgaca cagaccggac 60

ctgaacgagg tgataaaaca cctcgttcag gttcaaaacg tagaagattc attcctccga 120

ttagaaatac aactacgtct aagcacgata gagatggtat caagatcggt tttagaccac 180

gagaaaatcg gcaaatgaaa gtacaagttg ggtggtttaa attaccaaga acagtaacat 240

tcaagaacaa cggcaacccg tttgttacct catttcgtaa attgtttaga agtaacaaag 300

ataaattatt taatggtggg aagaacctaa gtacagtacc agccagaagt agtgaaatga 360

cagaaatgtt tatgttcatg tccacgctag agggccaatt gtcaatccaa gatcgagatc 420

caaaaataat caataagtct atatacatga tagaggta 458

<210> 588

<211> 675

<212> DNA

<213> Ball python nidovirus

<400> 588

ccccttcacc cataggcact aggagaacag gataaccct aacggggcat cctgcctgtg 60

acctttcaga ttcgctagtt agatatcttc acagactctg ctaggcttct gacccagtcc 120

gttcccaaag tccgttaccg cccgagtagc gcttaggcgc gaaagggacg gaagtacctc 180

cagtaagcga aagctgaagt aagggaaata cggcaagact aacttgttag tcttacagtg 240

tggataacct ggtagttatc cccgacaaga gacctgactc gatgtgaaaa catccaacta 300

ggttggcttc aaactagcta caggcaggat atcccccaac ggggcctcca aggaatcgag 360

aaccaaccct cattccacgt ctgtagtaag caaaaacagg ggcgatcttc accgacacct 420

ctcaccacag agcacaccaa cctctgtgaa gccaatttcc tcgtccaagg acaggttatt 480

gagggtcaac tttcttccga ccagaagaag ggatttccta ccaaaagaaa aaccaaatcc 540

accaacacca caaggtaaaa caacaacttg tgaagccaat acttagtcaa agactaacta 600

ttgagggtca actttctctt caatagagaa gggatttcct ggtaaaacaa ataacaacaa 660

ctaacatcag caact 675

<210> 589

<211> 150

<212> DNA

<213> Sapelovirus sp.

<400> 589

gacaggtgtt ttggagggcg gatgacgata tctggctggc caccagggaa taacggcaaa 60

tgtctgatca tacggttcac aagtctaccg gcgatagtgg ttcaacacca tgtgtagcag 120

ggattcttgc gtatgtgaag gcgacagtgc 150

<210> 590

<211> 496

<212> DNA

<213> Unknown

<220

><221> source

<223> /note="Description of Unknown:
Bat Picornavirus 3 sequence"

<400> 590

taagcggaaa gcattcttgt cccccggtca gtaacctata ggctgttccc acggctgaaa 60

gggtgaacat ccgttacccg cctcagtact tcgagaaacc tagtacgcct gatgattcca 120

aattggtatg atccggtcaa ccccagacca gaaactgtgg atgggggtca ccattcctag 180

tatggcaaca tacaggtgtc cccgcgtgtg tcacaggccc ttacgggtgc catttcggat 240

gagtctggcc gaagagtcta ttgagctact gttgatacct ccggcccccct gaatgcggct 300

aatctcaacc ccggagccac tgggtggtga accaaccact tggtggtcgt aatgagcaat 360

tctgggacgg aaccgactac tttggggtgt ccgtgtttct tttgttcata ttaaactgtt 420

ttatggtcac aacacaactt ggtacgattt gtgattattc actgctcact tgtcacagta 480

aatatacaca atcatc 496

<210> 591

<211> 419

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Bat Picornavirus 2 sequence"

<400> 591

gggttttacg aaacccgtat acaccagacc ttttctcccc tccccctcca cctacctttt 60

ccccctcttt ggaccgaaac aaggacacgt aagtggaaac gcgattttat atgtggttgg 120

ccaccacgga ataacggcaa ttgtctacat gtgggaagtg caacctccct agccgataac 180

ccctgaccgg gtgtgtagga taggaaaggt gcccactgtg ggcgacaggt tatggtagag 240

tggataccta gccaggggca atgggactgc tttgcatatc cctaatgaag tattgagatt 300

tctctgctca ttacccggtg atggttgtgt gggggggccc ccatacacta gatccatact 360

gcctgatagg gtcgcggctg gccgaccata acctgtatag tcagttgaat tcagccaag 419

<210> 592

<211> 410

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bat Picornavirus 1 sequence"

<400> 592

gaaacccgta tacaccggac cttttctccc ctccctctcc acttaccttt ttcccctctt 60

cggcatgaaa caaggattat tcaagtggaa acgcgattta atatgcggct ggccaccgcg 120

gaataacggc aattgtgtat ctgctggaag ccaagcctgc ctagccgata gcccttgacc 180

gggtgtgtag gatagcccag gaaccagcaa tacgcgacag gttatggtag agtagatacc 240

tagccagggg caatgggact gcattgcata tccctaatga accattgaga tttctctggt 300

cattacccgg tgatggttac tagaggggggg cctctagtac tagatctata ctgcctgata 360

gggtcgcggc tggccgacca tgacctgtat agtcagttga tttgagcaat 410

<210> 593

<211> 675

<212> DNA

<213> Iflavirus sp.

<400> 593

acgaatcggt atacgcttcg gtacctattg ttgcaagttc gttccctatt ttcgatttgc 60

ctgcccgaat ttgactcaaa caattgtgac atactatgtc tctgtttgaa agcactacac 120

gagtttgcgc ccagcatgta tgttttcaag tcttttgtat aagtctgcct ctatagtggt 180

tttctttgac cttaaagcct tgtcaaccat cctatatgct gcatcgagac ttgatgtcaa 240

tctgcctcta ctacgcaaat gtctagtaat tagttataag gttttactat tttccctcat 300

ttccaatttt agtttgtagt gtatgtgagt atcattctta ctccgactgt taagagaaac 360

caatttatag tcgttaaata tgataaatgg aatgaatgat ggtgtcattt taaaaacact 420

cttctctata ggcgtaagca ttctcgctct tagagtcgta aagaagaaat gccgtgtcta 480

tcagtatgtt atgcgattta ttttctgcca cgcgcttcta gtgcaatcta gttgacatac 540

agacattgcc taccactcgc gagggtcgac cggtagtgta aggagtaagt gatgatttcc 600

gcttattctg taccctttgc ctggtgagga cagatcctga ctaattttaa atataaatga 660

acactagctt ccaag 675

<210> 594

<211> 210

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bat dicibavirus sequence"

<400> 594

gtatagcacc ggaatggtat tttactactc caagtatacg tactaggagt taaaccctgt 60

aatttacagg ggatttagtg acttttatcc gtaaaagtcg attggacgtt aatcggtaac 120

gaggccaagt accgtgaacc aatttaaaaa cgtattttct catgtggtag aaccaacttg 180

gaaatagcat ggcatatagg ttgtttaggg 210

<210> 595

<211> 212

<212> DNA

<213> Betacoronavirus sp.

<400> 595

gataaagtgt gaatcgcttc cgtagcatcg caccctcgat ctcttgttag atctaatcta 60

atctaaactt tataaaaaca ctaggtccct gctagcctat gcctgagggt ttaggcgttg 120

catactagtg tcttaggaat ttgactgata acacttccct gctaacggcg tgttgcactc 180

tcagtctaag cctcccaccc ataggaggta tc 212

<210> 596

<211> 277

<212> DNA

<213> Betacoronavirus sp.

<400> 596

atttaagtga atagcttggc tatctcactt cccctcgttc tcttgcagaa ctttgatttt 60

aacgaactta aataaaagcc ctgttgttta gcgtattgtt gcacttgtct ggtgggattg 120

tggcattaat ttgcctgctc atctaggcag tggacatatg ctcaacactg ggtataattc 180

taattgaata ctatttttca gttagagcgt cgtgtctctt gtacgtctcg gtcacaatac 240

acggtttcgt ccggtgcgtg gcaattcggg gcacatc 277

<210> 597

<211> 675

<212> DNA

<213> Cardiovirus sp.

<400> 597

tacgatcgct gtacattcca ctactgccaa ttagctcccc cttcccgttg ctcccctcta 60

taaggagagc cttctcttgc aaaggtgaag ccttcacccc cggtcgaagc cgcttggaat 120

aagacagggt tattttctcc tctcctcggc gcttgcctct tctaagctga ataggttcta 180

tctattcagg cggatggtct ggtccgttcc ttcttggaca gagtgtgtat ctgggttttc 240

cggatctcga ccacacactc accagagctc aggagtgatt aagtcaaggc ccgatctgcg 300

gcgaaaagga aatgaagtat tttgcagctg tagcgacctc tcaaggccag cggatttccc 360

cacctggtga caggtgcctc tggggccaaa agccacgtgt taatagcacc cttgagagcg 420

gtggtacccc accaccctgc aaattatgga tttgacttag taactaaaag attgacttgg 480

catacctcaa cctgagcggc ggctaaggat gccctgaagg tacccgtgtt gaaatcgctt 540

cggcgaccat ggatctgatc aggggccctg cctggagtgg ttctatccca cacagcgtag 600

ggttaaaaaa cgtctaaccg ccccacaaag accccggcag ggatgccggt ttccttttta 660

ccaattcttg acact 675

<210> 598

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Breda virus sequence"

<400> 598

atcacctagt acttacaagc gggtcaaacc gccctccgga acggtcataa ccccctcccg 60

aacgtgcgct tgacgtgact ggtctttcag tctagctttc tgagaaatac tccggggttg 120

taacccacca ttttgacctt tggtcagttt ggtaacactc caaccaaaca gcatctgacc 180

cacctccagc ttgctgcagg ccatttggac caaacgggtt cagatatctt gtggctaaac 240

ctctgaccca cctccagttt actgcaggcc ttttggacta aacgggtaca gactctttgt 300

ggttagtttt taactaccca ctgtttagcc gccaacctga tttttattgt tacaaaattt 360

tgtgtttaca cattattttc ttacggttgg cagtttgttt ggttgtttgc acagtttttg 420

ctgataccaa tttttactgt gcttttggtg ttttcggcta aggctgtttt tcacatactt 480

agtttgcttg aagtaacctt cacaacatct gttttgtttt tggtttctaa ggtaaagagt 540

ttcaggaaaa aacataggcg cccatcttgt ggtgtctagt tttaattaat ctggcaaaca 600

agtatcaagt catcgactcc ctttggagtg agacttacga gtaccaattc gcctattttg 660

gccatccata taaaa 675

<210> 599

<211> 383

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bovine viral diarrhea virus 3 sequence"

<400> 599

gtatacgccc agttagttca ggtggacgtg tacgattggg tatcccaaat taataatttg 60

gtttagggac taaatcccct ggcgaaggcc gaaacaggtt aaccatacct ttagtaggac 120

gagcataatg ggggactagt ggtggcagtg agctccctgg atcaccgaag ccccgagtac 180

ggggtagtcg tcaatggttc gacgcatcaa ggaatgcctc gagatgccat gtggacgagg 240

gcgtgcccac ggtgtatctt aacccaggcg ggggccgctt gggtgaaata gggttgttat 300

acaagccttt gggagtacag cctgataggg tgttgcagag acctgctaca ccactagtat 360

aaaaactctg ctgtacatgg cac 383

<210> 600

<211> 537

<212> DNA

<213> Bovine rhinitis A virus

<400> 600

ttttgcggct ctgccgccgt tcgggtttta cctgttttca cagagcaaaa caggacctct 60

agtttcgtgc ttaaacgaga tcatgctcga actagaacta taacgctggt cactggaccc 120

gtgccgcgcc ttgcggatct ttgcgggaat ggtggctagt gggctgtgga agtgactcta 180

accacacgcc cctcaagtgt gggaaaacac gaactggtgt agcgacgacg ataggccttg 240

ggacaccctc tccagtgatg gagacccaag gggccaaaag ccacgccttg tgccctgtcg 300

ttcacaaccc cagtgcagtt cgtgccagta cctgcttttg ggaagtgtgc tttggacagc 360

tgaaaacagt cctagtggga gactaaggat gcccaggagg tacccggagg taacaagtga 420

cactctggat ctgacttggg gagagcgggt ctgctttaca gacgccactc tttaaaaaac 480

ttctatgtct cgtcaggcac cggaggccgg gccttttcct ttaaaacaat acacttt 537

<210> 601

<211> 248

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bovine picornavirus isolate TCH6 sequence"

<400> 601

ttggatctga gcaggggccc cctttgggtt gctttacaac tcaactgggg gttaaaaaac 60

gtctaacccg acacgccaga gggatctggt ttccttttat ttcttttact caccactgga 120

tgcagattga cgataaacgt tgttgtttgt gactattgac ttgatctgct tctacgggtt 180

tactttcact gttatacttc ttgctttgtt tggtgttcac tgtactttgt ctccttctac 240

atttcaca 248

<210> 602

<211> 486

<212> DNA

<213> Bovine nidovirus

<400> 602

caccaataga ttagtcaagc tgtctatagg cataaactaa cccccaaccc cattaccccg 60

gggccaggtg ggccgccgcc ttcgggcaaa cccgtgcgct ggtataatca aggttcacag 120

ccagattcac tgccggttag ctagtggggc ggtagcctgg caaaacccga agaggttgga 180

aagggaactt cagggtagtt tatcctaggc tagcgtagct acagttcggt caagataacc 240

gtcctggtgc tagggctagt agagacagtg gtaacttgga caagggtcca gggccacttt 300

agggaatacc ctacggaagg ctaggtccgt aaggaagacc cccgcagttg tccgcggttg 360

agcagagctc ctgcgtagac aaaaggcaaa aagtggatta cattcgcctg caggaaaagg 420

caaacgtcgt tggagtcgga gctaaagtac tggacgattg ataccacgcc tgctgcggta 480

gataaa 486

<210> 603

<211> 293

<212> DNA

<213> Hepacivirus sp.

<400> 603

ccatcgacac tccaggctca cggattaagt taggttccgc cgaagcgggc taaccaggcc 60

cctagtagga ggcgcctatc ccgtgagccc tttccccacg gattgagtgg agctggagct 120

gggaaggacc gagtacggtc caatcgagaa gaaccctgat gaacattcca ggcctcttcg 180

gtagatttgg atatatccac cagtgaaggc ggggtcgtgg gtacaggccc cctagtccac 240

acagcctgat agggtcctgc cgcaggatcc gtgggtgcgg ctgtacatgt acc 293

<210> 604

<211> 472

<212> DNA

<213> Mitovirus sp.

<400> 604

gccccgccga ccttcctatt tatttctaaa taggaaggtc ccgactagtc ggataattcg 60

gtttaacgaa ttatggttag atctattaaa gttaaaatag attgaatttc tttctcatcc 120

tccttattcc tatactttgg gagtaaatga caaatgtcta tcctcaaacc gaaatggctt 180

aagtgatgaa tttgaaagaa aggtaggttt taagaatata aggcatcaat atattatacc 240

cttgaatgtt aagtgaccac ggcgtgacga ttagggctat cttaggatag acagccatct 300

aacgcgacag cagtggaaat cagcttagca tctcaagatc atgtataata tatacataac 360

cttacaatta taaaaccaaa ccaaaacaca ctataatttt atataaatta tagaagtatc 420

ggacctggac ggtacctact attaactgat agtagccaaa tgcaggaagc tc 472

<210> 605

<211> 315

<212> DNA

<213> Mitovirus sp.

<400> 605

ggaacttttc agttccagaa gtggctttat taagccttca aagtttacac tttgaacctg 60

cgattaattc cccatagtga ctcttgttac tgtgaattag gaatagttgt agttcaactt 120

ctaatgaggt gaacaatata ataactcatc ttattaaccc tatacgtaga caattgtcca 180

aagagacagt tggaattctg ccaatctgga atgtttggta cgcgtagaag ataataagag 240

accctctatt cccctgcctc atgactaagt catggcccgg ggtgtaatag agatactttt 300

atatattata caatc 315

<210> 606

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Canine picodicistrovirus strain 209 sequence"

<400> 606

cgctctttat acaaatctgt caacccttg tataactcta agccgaacaa ttatagctag 60

gctttttatt atataacatt aattaggcat tagcgttgtc gccaatctct tggtaatcct 120

aaggatacct ttcctgttga ctaagatgaa gcgccttcgg ttaccgatgc ccggtgtcca 180

cgaagccatc gtggtcggcc gcgtccccca cctctcccaa cttggactcc atgttttcag 240

taggtgtaat gattagtatt attgattctg ctcgttcaat gtgtttatct tcacgatctg 300

ggacccaaca catgcttcac tcatgtttaa atgttggttc cctcattttg aagacaccca 360

aaccatagag tgcgagaatg aggatttcta cttccattct ggtaacagaa atgaattcct 420

gcgtgtgtct cgtaaatgga atctttaaga acttcagata aatcgaacaa tacactaata 480

caagttgttt tctaccaaca tgttcaatgc ggctaatctg accgtggagc tgtgaagcgc 540

tcaaacccga gtgttgtata cagtcgtaat gcgtaagtcc atgaggaacc gactactgtt 600

acctctgttg gtgtgtttct cctttcctct cttttattat tatttgttat tgcaaatact 660

acaactttga tcaac 675

<210> 607

<211> 107

<212> DNA

<213> Canine morbillivirus

<400> 607

accagacaaa gttggctaag gatagttaaa ttattgaata ttttattaaa aacttagggt 60

caatgatcct accttaaaga acaaggctag ggttcagacc taccaat 107

<210> 608

<211> 601

<212> DNA

<213> Kobuvirus sp.

<400> 608

tttaagtgtt gtgcccaatc tcttgactcc tgctggaacc accgaccagt agtgtccaaa 60

atgccaggtg gaaaatcctc ccttcccctc tgggcttcat gcccggcatc ctccccccag 120

cctgacgtgc cacaggctgt gcaaagaccc cgcgaaagct gccaaaagtg gcaattgtgg 180

gtcccccctt tgtcaaggcg tcgagtcttt ctcccttaag gctagtcctg tcagtgaact 240

ctgtcgggca actagtgacg ccactgcatg cctccgacct cggccgcgga gtgctgcccc 300

ccaagtcatg cccctgacca caagttgtgc tgtctggcaa acattgtctg tgagaatgtt 360

ccgctgtggc tgccaagcct ggtaacaggc tgccccagtg tgcgtaattc tcatccagac 420

ttcggtctgg caacttgctg ttaagacatg gcgtaagggg cgtgtgccaa cgccctggaa 480

cgagtgtcca ctctaatacc ccgaggaatg ctacgcaggt acccctggct cgccagggat 540

ctgagcgtag gctaattgtc taagggtatt ttcatttccc accctcttct tcttgttcat 600

a 601

<210> 609

<211> 291

<212> DNA

<213> Alphacoronavirus sp.

<400> 609

cttaagtgtc ttatctatct atagatagaa aagtcgcttt ttagactttg tgtctactct 60

tctcaactaa acgaaatttt tgctacggcc ggcatctctg atgctggagt cgtggcgtaa 120

ttgaaatttc atttgggttg caacagtttg gaaataagtg ctgtgcgtcc tagtctaagg 180

gttctgtgtt ctgtcacggg attccattct acaaacgcct tactcgaggt tctgtctcgt 240

gtttgtgtgg aagcaaagtt ctgtctttgt ggaaaccagt aactgttcct a 291

<210> 610

<211> 623

<212> DNA

<213> Cripavirus sp.

<400> 610

gcaaaatcgg tagtacgtta aacgtacgac caccgatgag actgaaatga cactagttga 60

gattatttca atatcctagt gttataaagt caatatttgt tggttgatcg tttcgtcaat 120

cgatggcgct gacagccgga aagacggcaa taataaaaac caagatttag tttttaagtt 180

ttgattgaat tgcaaaagct atcttgaata gacaatcaaa atattaagta aagcaaaagc 240

ttcttaaaga agacaatatt taattagtta gtaaccaaac ctcatcgtgc ccctaagggt 300

taaccggtta cgtaaaagcg tagaggtatt aaggtcactg cggagaccta aaatccgcaa 360

ttttatgttt tgtaatgttt tagttataga cttagatgta actataagag tttataaata 420

cttgtttcaa gatttataga caagatctga tcctatggat tttagataac cttcatgtta 480

gtggatagtg tgtgtaccta tctaaacgca taaggctctt atttcatatt taaagtagga 540

ctatgtatta cggcgcatct aacggtaacg ttagtcaaga ccggagaatc tcggaatgaa 600

ttttagtaat tcccaaattt ata 623

<210> 611

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human coxsackievirus A2 sequence"

<400> 611

acctttgtgc gcctgtttta tatccccacc ccgagtaaac gttagaagtt acgcaacccc 60

gatcaatagt aggtgtagca ctccagctgc atcgagatca agcacttctg tctccccgga 120

ccgagtatca atagactgct aacgcggttg aaggagaaaa cgttcgttac ccggccaatt 180

acttcgagaa gcccagtagt gccgtgaaag ttgcggagtg tttcgctcag cacttccccc 240

gtgtagatca ggctgatgag tcaccgcgat ccccacaggt gactgtggcg gtggctgcgt 300

tggcggcctg cctatggggc aacccatagg acgctctaat acagacatgg tgcgaagagc 360

ctattgagct aattggtagt cctccggccc ctgaatgcgg ctaatcctaa ctgcggagca 420

catgccctca aaccaggggg tggtgtgtcg taacgggtaa ctctgcagcg gaaccgacta 480

ctttgggtgt ccgtgtttct ttttattctt ataatggctg cttatggtga caattaaaga 540

attgttacca tatagctatt ggattggcca tccggtgact aacaaatcgc tcatatacca 600

gtttgttggt tttgttccct tatcacatac agctcataac accctcttat atttactaca 660

attgaatagc aagaa 675

<210> 612

<211> 436

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Coronavirus AcCoV-JC34 sequence"

<400> 612

agaaacaagt agtgttttaa aaaccttcaa attagtgcct gtaacatctt tgcaatgaaa 60

gtagcgctca ctagcctcta tgcaaagaat gttaaaagaa atacgaagca tttaaagaat 120

acaatctatc taggataggt acaattctcc tccccctctt gacttcggtc aactcaactc 180

aactaaacga aatccccctt gcatggttcc gacccgtgta aggttgtgta tttcgtgcag 240

tcgttgccct tactagtgta agcgtaacgg catctaggtt tgcacgtctt ggaggaaacg 300

gtgtgtacgt ttctagtgtt tacgccgtat cggttccggc ccgataggta ttgcattaga 360

cgtcctgggt ggttctgcct gcccttgtgt gattcggctg ttccgtcagt ttggtcacct 420

cacacgtcct taagac 436

<210> 613

<211> 457

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Chicken picornavirus 3 sequence"

<400> 613

gggtatggtg gttaaccccg tccactgggc atctttgccc tctcagaagt ggatcaatcc 60

accccaactc ccccccctggt tacctgagcc acagtggact ccggtgacga agctagggac 120

cccaatacct caagtgccaa gagagtcccc ccctcgcgag aggtgctctt gggcccaaaa 180

ggctagttgg cagagtgaag tgaaggaagc tgctaacgtg gtgaccttaa gcgtaattcg 240

aagctgacct ttgaggttaa ccctagtgga ccactggagg aatctgtgga ggtggtggtt 300

aggaaagttg gccacttgtg agtagatgcc cagaaggcat aaggctgatc tggggccagt 360

gactataccg ttccggtaaa cctggtataa aaaccatgaa agcaagtggg tgaaattttc 420

tctttttatc cttcattcag cagttgatat tggcaaa 457

<210> 614

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Chicken picornavirus 1 sequence"

<400> 614

gtggccgact tgcagaactt cctaccgaac caccacctca cccccataac tccttccctc 60

tacaccttcc gctatggtgg ttaccactgc tttattgcct tgactgagaa tggccacccc 120

ctcgacacct gccccctact gccccaccgc gcaacctttt gcttgtccac tcggttggag 180

aggcatgggg gccccgttca tatccccagt ccagttggtg accttccccc tccccgtccg 240

gtagatggtc cagagggctt tgccgacgcc ctctatgatg cttgcttgtc tttcctccgt 300

cagcgcgagc atgcacttgt cgagcccacg gaaacacagc ctagcttttg cttctctcac 360

cctgcgtacc ctgggcgcct tccgctcgag attcgctttg ttcgacaccc tggcgtcccc 420

ccaccgctac gtgattttac tcgtggcata caccgccctg gcgttcagta cattccactg 480

cctaatttgg gtggcctccc tcaatctccc gcacccccca ttgcgcacgt catcaccgcc 540

gccgctaacg cgatccggcg cggttctcac tggcactgtc ccctcgtccg ccgggtttcc 600

actcatggtt ggcttttcac ttattctggt tactggtgtt ccatcctact tcatgatggt 660

cgccatgacc atgac 675

<210> 615

<211> 675

<212> DNA

<213> Orivirus sp.

<400> 615

aaaccctcac gagtgcttgt ggtaggtccc aggccaatat tcttcgtaag gcttggttcc 60

aattttccac cactcgtgtt tgggttctgg cctatggtac ccagaggggc ggtttggggg 120

aattaactcc ccctcccctg tggtcctata ccaccccaca cctctgtggg ctttctttac 180

tatcttcttg ttttccgact tttaaacact aggcaggcgc gcctagtcat acaccgcccg 240

gctggtcttt ccagctcttg tgggcggtgc gcgctggtcc atcgtgccca gcgacatagc 300

accttgtgga cacctccgaa cgccctcccc tgtatggggt ggtgcccagg ggtttcagtg 360

tggtgacaca ctccctgggg cccgaaaggc tagtgtgcaa caggtgaggt acagccagct 420

gcccccgtgg ctggagggac caagcttgtg aagcacacct caccttcttg ggggtgggct 480

agtaagtggt gaaagcatag tgtccgtgtc gctggccaac actttgggtc aagtccagcc 540

actcagtgag tagatgccca ggaggtaccc ctagtggatc tgacttgggg cctgttactt 600

aatgcaggtt aaaaactatg aaagctgagt agtgtagccc ggctggtggc ttctcttcct 660

tattcattct atttt 675

<210> 616

<211> 651

<212> DNA

<213> Gallivirus sp.

<400> 616

ggttaacttg tttaaccaag gcttccgtgc agggcttgca cgttaggagt tcatgttgat 60

tcatgctcca atgcccaaaa ctttgtgtgt ttgtttatgt ctttcccaaa gtttcccccа 120

aggtttcggt actcaaacct taattcctag tcccatcttt tgggccttag tatctaggaa 180

atgtacccgt gccttgacga acgtaagaaa gctgtctttt attgaacggt tctaatgaac 240

taagtataac tggctcgcgc cacctggtgt gtgccgttgg aattccccca tggtaacatg 300

gtccaacggg cccgaaaggc tagtgggcaa tcggtcctcc aaggaagggg ttcccacccc 360

gacctgaaca ggatttgatg aagctcacct cccaggctcc taaccccaag gaagtttact 420

tatagtaatt agaatttagt atgtaattgc tggcaatctt gctagtagtc aggaacgtta 480

tgaccaaatg agtagacccc cagaaggtac cccattatat gggatctgat ctgggcctca 540

tactgtgtgt ctccccacat atgaggttaa aaccatgaaa gtttggtcca aaatattctt 600

ttccttttat cttttcttta gtggtgacgc cattatatca gcagtttgct g 651

<210> 617

<211> 303

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Chicken calicivirus sequence"

<400> 617

ggtgcatcat cactgaacac cctcgggcag agatgcaagg gtggaagtca ctcctgcccc 60

ctggcaacat gcaggtgccc gatcccaagc ttagttctga cttcctctcc tgggtggtgc 120

aacactccaa ggttgatgaa caaacctgga gggacctctg ggcaacctgg tctctcgagg 180

atctccggcg catctccaca gactacctca tgctcccgga acctgagaag aacactgatg 240

cctatgatgg ctggctgatg gtcggcgaga tgttgacctt tgccggtctc attggctggg 300

agg 303

<210> 618

<211> 524

<212> DNA

<213

> Unknown

<220><221> source

<223> /note="Description of Unknown:
Carp picornavirus 1 sequence"

<400> 618

agctacagga aagagagaga taatcacagc acataaatac aactacagaa gagagccttt 60

ccctgagcac tatttacagc aaaccacgct gggaaaagtg gtagcatgac ccacttacgg 120

gttacttagt ataggatttt aatatcttcg attctttatt ttcttttctt aactaagttc 180

gcccggactt tatccgtgtt ttattttagt ttaagcaaaa ttgagtaact aagtttttaa 240

cctgccaaat ggtgagaagt aactctgtga aaataccatt tgtgcatgaa attgtcttga 300

aaactcttag gcttttgggg ggtcccactg ctgtttggag gactattgac agactctatt 360

gtagttgagt agtgactaat gataacgatt tgcgtattac gcaatgggct gtacccgtta 420

gatttagtat gccgggggga ggggtcccac tggattgcac tatgtaacct gacagggcgt 480

ctgccgacgc actacaatga ggataagatc ggctgttttt attt 524

<210> 619

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Falcon picornavirus sequence"

<400> 619

cgcttggaat aagttgagag gaattatgca tgctagttgt gtttgtttac aactaattgt 60

tctaatccaa gtgaagctct tcgcttgggg cggcacgaca cttgccgtaa ttcttctacc 120

gtccctccac accttgtgga tgaagggccg gatgtgtggc ctctggctaa cccctctctc 180

tggggtgatg ctactggatg ttttactcct agaccaaatc acatgaactc ctcttgatcc 240

acttcggtgg ggctatgagc ctgcggatta atagctggcg acagctaccc caggggccaa 300

aagccacggt gttagcagca ccctcatagt ctgatgccca agggctgatg ttgggagcta 360

gtagtgtgtg tctggcctat gtttaggact tctggccaag cgcagaggag tggggctgaa 420

ggatgcccag aaggtacccg taggtaacct taagagacta tggatctgat ctggggcccc 480

ctcacgtggc tttaccacgt gttgggggtt aaaaaacgtc taggccccac cagcccacgg 540

gagtgggctt tcccttaaaa agcccaacaa tatttatggt gacaattcac tgtttcttct 600

ttgcaatttt tgtattcttg gactccttat tttttgtttg cttgatttta gtggacattc 660

agattcaaat acaaa 675

<210> 620

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Equine rhinitis B virus 1 sequence"

<400> 620

cgacaggcac aggtcgctcc gagttctagt agtgtgggaa cttgttacta ctgatgaaac 60

gaggtagtga cactcagtac ctgcgaacga ggtcggggcc ctcccttctt ccttcaccca 120

actttcactt ttcgttccac tttagcaggg gtcttctttc tatccccctg gcggcattgg 180

aactagccgt cgcgtcttaa cgcgcagccc tgaaggcccc acaccttgtg gatcttgccg 240

tgggtatgtt tctggcatgt gtttctcaag cctgcaaccg aagccgaaca gccacatgaa 300

cagtttgagc gtggtagcgc tgtgtgagtt ggcggtggat ccccctcgtg gtaacacgag 360

cccccgtggc caaaagccca gtgtttacag cacctctcac atccaggacg accccatcct 420

ggcgctcact cttagtagta tggcttagta cgcattaggt ggtaagccga gctctccctc 480

ggccttgttc tgaatgcaca catgtctagg ggctaaggat gtcctacagg tacccgcacg 540

taaccttcag agagtgcgga tctgagtagg agaccgtggt gcactgcttt acagatgcag 600

cccggtttaa aaagcgtcta tgcccctaca gggtagcggt gggccgcgcc ctttcctttt 660

aaaactactt gttct 675

<210> 621

<211> 675

<212> DNA

<213> Equine rhinitis A virus

<400> 621

aagggttact gctcgtaatg agagcacatg acattttgcc aagatttcct ggcaattgtc 60

acgggagaga ggagcccgtt ctcgggcact tttctctcaa acaatgttgg cgcgcctcgg 120

cgcgcccccc ctttttcagc cccctgtcat tgactggtcg aaggcgctcg caataagact 180

ggtcgttgct tggcttttct attgtttcag gctttagcgc gcccttgcgc ggcgggccgt 240

caagcccgtg tgctgtacag caccaggtaa ccggacagcg gcttgctgga ttttcccggt 300

gccattgctc tggatggtgt caccaagctg gcagatgcgg agtgaacctt acgaagcgac 360

acacctgtgg tagcgctgcc cagaagggag cggagctccc ccgccgcgag gcggtcctct 420

ctggccaaaa gcccagcgtt aatagcgcct tctgggatgc aggaacccca cctgccaggt 480

gtgaagtgga ctaagtggat ctccaatttg gcctgttctg aactacacca tctactgctg 540

tgaagaatgt cctgaaggca agctggttac agccctgatc aggagccccg ctcgtgactc 600

tcgatcgacg cggggtcaaa aactgtctaa gcagcagcag aaacgcggga gcgtttcttt 660

ttcctcattt gtttc 675

<210> 622

<211> 224

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Equine arteritis virus sequence"

<400> 622

gctcgaagtg tgtatggtgc catatacggc tcaccaccat atacactgca agaattacta 60

ttcttgtggg cccctctcgg taaatcctag agggctttcc tctcgttatt gcgagattcg 120

tcgttagata acggcaagtt ccctttctta ctatcctatt ttcatcttgt ggcttgacgg 180

gtcactgcca tcgtcgtcga tctctatcaa ctacccttgc gact 224

<210> 623

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 623

actctggtat cacggtacct ttgcacgcct attttatacc ccttccccat cgtaacttag 60

aagcaacaaa caaactgccc aatagcagca caacacccag ttgtgttagg ggcaagcact 120

tctgtttccc cggaagggtc tgacggtatg ctgtacccac ggcagaagta tgacctaccg 180

ttaaccggcc atgtacttcg agaagcctag taccattatg aaggttgatt gatgttacgc 240

tccccagcaa ccccagctgg tagttctggt cgatgagtct cggcattccc cacgggcgac 300

cgtggccgag gctgcgttgg cggccagcct acaccatacg gtgtaggacg tcaagatact 360

gacatggtgt gaagagccta ttgagctacg tggtagtcct ccggcccctg aatgcggcta 420

atcctaactc cggagcatcc gccagtaagc ccactggaag ggtgtcgtaa tgcgaaagtc 480

tggagcggaa ccgactactt tgggtgtccg tgtttcctgt tttacttatt gtttggctgc 540

ttatggtgac aacttatagt tatcatcata agctacttgg tcttgccaac cggagaatta 600

tttggttatt tgttggtttc ataaacctac agtcgtattt tcctgtctta ttaattgttc 660

tcaaaattaa caaca 675

<210> 624

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 624

taccgctgca ccagtgagct ggtacgctag taccttcgca cggagtagat ggcatccccc 60

accccgtaac ttagaagcaa agtacacatc tggccaatag tggcgctgca tccagccgcg 120

caacggtcaa gcacttctgt ttccccggtc cgcaagggtc gttatccgcc cagtccacta 180

cggaaagcct actaaccatt gaagctatcg agaggttgcg ctcggccacg accccggtgg 240

tagctctgag tgatggggct cgcaaacacc cccgtggtaa cacggatgct tgcccgcgcg 300

tgcactcggg ttcagcctat tggttgttca cctcaacata gtgtaaatgg ccaagagcct 360

actgtgctgg attggttttc ctccggagcc gtgaatgctg ctaatcccaa cctccgagcg 420

tgtgcgcaca atccagtgtt gctacgtcgt aacgcgtaag ttggaggcgg aacagactac 480

tttcggtact ccgtgtttct tttgttttat tttgaatttt atggtgacaa ttgctgagat 540

ttgcgaatta gcgactctac cgctgaacat tgccctgtac tacctaatcg catttcacaa 600

aacctcagag ataccaagct cttacattga tctgcttgtt ttcctgaatc tcaaatataa 660

attggaacaa gcaaa 675

<210> 625

<211> 107

<212> DNA

<213> Morbillivirus sp.

<400> 625

accagacaaa gctggctagg ggtagaataa cagataatga taaattatca tacttaggat 60

taatgatcct atcaattggc acaggatttg gataaaggtt cacagtc 107

<210> 626

<211> 348

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Dianke virus sequence"

<400> 626

tgttttcaac cataatacta ctactacaag tataaaaccc cgtccgtctg tcggagacgc 60

taaactctga ccaccaatct agccacatca gttgcttaaa gaacctcttg agacactctc 120

ccacttaaca tcttttagga atcttcgatg ctacaacaac ttggctagtg aacaataaat 180

ccgtacaatt cacagttgta agaggccata ggtccagact ttgaaaggtt tgtttctatt 240

gttacaaata cttagattaa cagaggctat ttaatagtgc tcatcacgtt aacagagtaa 300

ccttgtgcaa tagtatgagc ttgttgtaaa acgtcttgat acgacacc 348

<210> 627

<211> 270

<212> DNA

<213> Hepacivirus sp.

<400> 627

cactcaatac tacactccgc atttggggag aagcgctggc gttcgcggaa ccgcgttaac 60

catacgcgta gtacgagtgc gacagacccc ggtgctactg gtggtagcga gacacgagcc 120

gaagtctgtg gggggaactc cacttagagg gcatgcccgg gcgtaggctt ctgagttggg 180

atgggcccca acttggcccc tgagtggggg ggtgttacga cctgataggg tgcgggctgg 240

cgcctaccac tttccagtcg tacatgagtc 270

<210> 628

<211> 415

<212> DNA

<213> Narnavirus sp.

<400> 628

gcccgggggg tgcagtcctg tgaaagggtc tgcaccatac tatatatgta tatgattaca 60

tcccaaaagg cgacttcgtt caggttttaa atctgacgta ggtccagtaa ataagcatgt 120

caaaacatgt aagtttatcc tgtaatctac tctcataaga tgagataaga tgatattgca 180

gttcccatgt aaataaatcc attatgaatt cattcatata aggtagaagt ggtaactatg 240

gttgaaacat taatataaaa cggtcatttt gcatgaacgt cattaaggaa ctggcatacc 300

aatgtctatt tagtgactat gatatttaga gtatccctta tattaattaa caattattcc 360

ttttagcata tcatccgaca acaaatttta aaagaagaaa tattactcat taaaa 415

<210> 629

<211> 675

<212> DNA

<213> Torovirus sp.

<400> 629

gtacttacaa gcgggttaaa ccgccctccg gaacggttac aacccccctcc cgaacgtgcg 60

cttgacgtga ctggttctca gtctggcttt ctgagaaata ctccagggtt gtatcccacc 120

atcttgacct ctggtcattt tggtaacacc ataaccaaac aaactctaca cacctaaccc 180

acctccagct tgctgcaggc cttttggact aaacgggttt aggtgttttg tgaccaactc 240

gtctacccac ctccagtttt actgcaggcc tttttggact aaacggcttt agacttttgt 300

ggttagtttt taactaccca ctgtttagcc gccaacctga ttttcattgt tgtaaaattt 360

tgtgtttata cactattttc atacggttgg cagtttgttt ggttgtttgc acagtttttg 420

ttgataccaa tttttactgt gcttttagtg tattctgcta aggctgtatt ttacatactt 480

agtttggttg aagcaatttt tacaacattt atattgtttt tgatttctaa ggtaaagagt 540

cttaggaaac accatagacg ccattcttgt ggtgtctagt tccaactaat ctggcaaaca 600

agtaccaagt cattgactca ctttggagtg agacttacga gtaccaattt gcctatttcg 660

gacatccata taaag 675

<210> 630

<211> 675

<212> DNA

<213> Foot-and-mouth disease virus

<400> 630

acaagcttga caccgcctgt cccggcgtta aagggaagta accacaagct tacaaccgcc 60

taccccggtg ttaatgggat gtaaccacaa gatacacctt cacccggaag taaaacggca 120

aattcacaca gttttgcccg tttttcatga gaaacgggac gtctgcgcac gaaacgcgcc 180

gtcgcttgag gaggacttgt acaaacacga tctaaacagg tttccccaac tgacatacac 240

cgtgcaattt gaaactccgc ctggtctttc caggtctaga ggggtaacac tttgtactgt 300

gcttgactcc acgctcggtc cactggcgag tgttagtaac agcactggtg cttcgtagcg 360

gagcatggtg gccgtgggaa ctcctccttg gtaacaagga cccacggggc cgaaagccac 420

gtcctgacgg acccaccatg tgtgcaaccc cagcacggca acttttctgt gaaactcact 480

ctaaggtgac actgatactg gtattcaagt actggtgaca ggctaaggat gcccttcagg 540

taccccgagg taacacgcga cactcgggat ctgagaaggg gactggggct tctgtaaaag 600

cgcccagttt aaaaagcttc tatgcctgga taggtgaccg gaggccggcg cctttccatt 660

ataactactg acttt 675

<210> 631

<211> 310

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Feline infectious peritonitis virus sequence"

<400> 631

actttaaag taaagtgagt gtagcgtggc tataactctt cttttacttt aactagcctt 60

gtgctagatt tgtcttcgga caccaactcg aactaaacga aatatttgtc tctctatgaa 120

accatagaag acaagcgttg attatttcac cagtttggca atcactccta ggaacggggt 180

tgagagaacg gcgcaccagg gttccgtccc tgtttggtaa gtcgtctagt attagctgcg 240

gcggttccgc ccgtcgtagt tgggtagacc gggttccgtc ctgtgatctc cctcgccggc 300

cgccaggaga 310

<210> 632

<211> 143

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Farmington virus sequence"

<400> 632

acgacgcata agcagagaaa cataagagac tatgttcata gtcaccctgt attcattatt 60

gacttttatg acctattatt agacccttca cgggtaaatc cttctccttg cagttctcgc 120

caagtacctc caaagtcaga acg 143

<210> 633

<211> 528

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Avian infectious bronchitis virus sequence"

<400> 633

acttaagata gatattaata tatatctatt acactagcct tgcgctagat ttttaactta 60

acaaaacgga cttaaatacc tacagctggt cctcataggt gttccattgc agtgcacttt 120

agtgccctgg atggcacctg gccacctgtc aggtttttgt tattaaaatc ttattgttgc 180

tggtatcact gcttgttttg ccgtgtctca ctttatacat ctgttgcttg ggctacctag 240

tgtccagcgt cctacgggcg tcgtggctgg ttcgagtgcg aggaacctct ggttcatcta 300

gcggtaggcg ggtgtgtgga agtagcactt cagacgtacc ggttctgttg tgtgaaatac 360

ggggtcacct ccccccacat acctctaagg gcttttgagc ctagcgttgg gctacgttct 420

cgcataaggt cggctatacg acgtttgtag ggggtagtgc caaacaaccc ctgaggtgac 480

aggttctggt ggtgtttagt gagcagacat acaatagaca gtgacaac 528

<210> 634

<211> 626

<212> DNA

<213> Rhinovirus sp.

<400> 634

ttaaaactgg gtgtgggttg ttcccaccca caccacccaa tgggtgttgt actctgttat 60

tccggtaact ttgtacgcca gtttttccct cccctcccca tccttttacg taacttagaa 120

gttttaaata caagaccaat agtaggcaac tctccaggtt gtctaaggtc aagcacttct 180

gtttccccgg ttgatgttga tatgctccaa cagggcaaaa acaacagata ccgttatccg 240

caaagtgcct acacagagct tagtaggatt ctgaaagatc tttggttggt cgttcagctg 300

catacccagc agtagacctt gcagatgagg ctggacattc cccactggta acagtggtcc 360

agcctgcgtg gctgcctgcg cacctctcat gaggtgtgaa gccaaagatc ggacagggtg 420

tgaagagccg cgtgtgctca ctttgagtcc tccggcccct gaatgcggct aaccttaaac 480

ctgcagccat ggctcataag ccaatgagtt tatggtcgta acgagtaatt gcgggatggg 540

accgactact ttgggtgtcc gtgtttcact ttttccttta ttaattgctt atggtgacaa 600

tatatatatt gatatatatt ggcatc 626

<210> 635

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
EV22 sequence"

<400> 635

ccttataacc cgacttgctg agcttctata ggaaaaaacc ctttcccagc cttggggtgg 60

ctggtcaata aaaaccccca tagtaaccaa cacctaagac aatttgatca accctatgcc 120

tggtccccac tattcgaagg caacttgcaa taagaagagt ggaacaagga tgcttaaagc 180

atagtgtaaa tgatcttttc taacctgtat tatgtacagg gtggcagatg gcgtgccata 240

aatctattag tgggatacca cgcttgtgga ccttatgccc acacagccat cctctagtaa 300

gtttgtaaaa tgtctggtga gatgtgggaa cttattggaa acaacaattt gcttaatagc 360

atcctagtgc cagcggaaca acatctggta acagatgcct ctggggccaa aagccaaggt 420

ttgacagacc cattaggatt ggtttcaaaa cctgaattgt tgtggaagat attcagtacc 480

tatcaatctg gtagtggtgc aaacactagt tgtaaggccc acgaaggatg cccagaaggt 540

acccgcaggt aacaagagac actgtggatc tgatctgggg ccaactacct ctatcaggtg 600

agttagttaa aaaacgtcta gtgggccaaa cccaggggg atccctggtt tccttttatt 660

gttaatattg acatt 675

<210> 636

<211> 675

<212> DNA

<213> Cardiovirus sp.

<400> 636

tccgacgtgg ttggaattaa catcttttcc gacgaaagtg ctattatgcc tccccgattg 60

tgtgatgctt tctgccctgc tgggcggagc gtcctcgggt tgagaaacct tgaatctttt 120

cctttggagc cttggctccc ccggtctaag ccgcttggaa tatgacaggg ttattttcca 180

aactctttat ttctactttc atgggttcta tccatgaaaa gggtatgtgt tgccccttcc 240

ttctttggag aatctgcgcg gcggtctttc cgtctctcaa caggcgtgga tgcaacatgc 300

cggaaacggt gaagaaaaca gttttctgtg gaaatttaga gtggacatcg aaacagctgt 360

agcgacctca cagtagcagc ggattcccct cttggcgaca agagcctctg cggccaaaag 420

ccccgtggat aagatccact gctgtgagcg gtgcaacccc agcaccctgg ttcgatggcc 480

attctctatg gaaccagaaa atggttttct caagccctcc ggtagagaag ccaagaatgt 540

cctgaaggta ccccgcgcgc gggatctgat caggagacca attggcagtg ctttacgctg 600

ccactttggt ttaaaaactg tcacagcttc tccaaaccaa gtggtcttgg ttttccaatt 660

ttgttgactg acaat 675

<210> 637

<211> 292

<212> DNA

<213> Human coronavirus 229E

<400> 637

acttaagtac cttatctatc tacagataga aaagttgctt tttagacttt gtgtctactt 60

ttctcaacta aacgaaattt ttgctatggc cggcatcttt gatgctggag tcgtagtgta 120

attgaaattt catttgggtt gcaacagttt ggaagcaagt gctgtgtgtc ctagtctaag 180

ggtttcgtgt tccgtcacga gattccattc tacaaacgcc ttactcgagg ttccgtctcg 240

tgtttgtgtg gaagcaaagt tctgtctttg tggaaaccag taactgttcc ta 292

<210> 638

<211> 282

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hubei zhaovirus-like virus 1 sequence"

<400> 638

gtgcaggatg gcctttccca tcttaagtgg tttgtaggat ttcgtgggtc catacccccc 60

gatttcttgg tacgtattcc atgcacggag aatacgacca aaactcttat ttcaaaaaat 120

attatttttt actcttgtgg gctgagtgcg acccaccagt tccagcttag caacctggaa 180

gttgttggag atttatggaa ccaaattaca catgcgtgga gtgccgccac tccgtatctg 240

ttcactcatt acgcgattaa gttctgcgac gagacgagcg aa 282

<210> 639

<211> 600

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hubei tombus-like virus 9 sequence"

<400> 639

ggaccatcca ggcaggtgta ggctagtacc ctcacctgac ctgtcgcgat gtttggcttt 60

gtgaggcttg tgggaggatc ccttggccca tgcattgctg ctgtcatcgt gaaaaatgtt 120

gtatgctcgc acctggcgtg gaggaaacgg catttgacgg atgctaagga ggatttgaag 180

gtcctcaact cccatgcctt attacaagtc ccctttcctt caagcttcga gcccacgtct 240

gtgacgagca gaggtgaggt tggagtcaag aggagtccta ttcaagctgt tcgcagcaag 300

aaccataaaa ctttcattgc tcaatgggca agagcggcta aggctcggtt ctcgtttgca 360

cgacagtgtg agcggagccc tgtaaatgtc gcggcccttc atcggtggtt tgcatcggag 420

tttaagaaaa ttggcatgaa cttgctccag gtttcgtacg tgatcgatga ggttgttgaa 480

cttagtttgg aaccaacgta tgagcgtatg gtgtccgaaa ctaagaagca attccgtcat 540

cgggcacgta tggaatacta caacgagaaa aaatgccttg aaaagatcca ctaggaacgc 600

<210> 640

<211> 251

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hubei tombus-like virus 32 sequence"

<400> 640

ttcgggttta cccgcgtaag cggccacact gactggttgt cggtgttgaa tttgtatacc 60

agatgaggag acgttaccac cgtctcggca gtgctacgtc tgggaaagga ctgtgatagt 120

ggacagtcct accgtcttta gatacattgc gttgtgtata gcccgggagg gattaactaa 180

tagcaacgca atgcacacgg cggttcggat ttgcttgact gatggaaaga catctaagtt 240

ctaattgaac c 251

<210> 641

<211> 341

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hubei sobemo-like virus 3 sequence"

<400> 641

gagatgtttg cgtgggccgt tgcgctgcgg gcggcccacc tccctacggg gaccgtgaga 60

caccgctggg gaaggccccc acccccggcc aaggggatcc tgccgagagg caggagaaag 120

aggcccagcc ctctggggcg cctttagggg tgcctgggag ggaagtaccc gagccgggcg 180

gccggtcggg tgcggctgtg cagttcgagg ctaaccgtaa ggaaggcctg agctgcctcg 240

gcttgtcgga aaggaagacg aaggcactta tcaaggcttt caggaagcag gagcgcaatt 300

acagcgcgcg ccgtgcggcc tggattaacc gcatttggcc g 341

<210> 642

<211> 675

<212> DNA

<213

> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hubei picorna-like virus 2 sequence"

<400> 642

agcaacttct ttctgaaaac tagctagagt tcgacgatct ctctggctaa tgacaaataa 60

ccaatcaaaa agtcaaatgt tcatgtatat atatatttag tagtgacctt tatttagaaa 120

aactttagat gatttatcgt caagttgccc tttgtgaagc gatcagcttt ttatatcgtt 180

tcatttttgt atacgtctta attgacgttg taagtacgtt ttgcatacct cacattgagg 240

atagtatcgt ttcctgacta agaagttaaa ctagtctacc aatagcaacc atataggata 300

tagctttgtt tttaacaagg tttaatctga tcttatgctc ttgcttacgg tgttttatgt 360

atagaaaagt ttttataaaa actacataat tgtcaaaaga aaaagcgtta cgtactgacg 420

catttatgtt cacagtgtga cacaaacctt ctatattttg attgtaaaat aggctttgcc 480

tgaccattta tcaaatacaa actgtttcaa acgcctctcc gagccatatt ggccgacgcg 540

aatcgacata acagggtgag tttacagctg cagttgcagc cgaggatcca cttatttgag 600

agagaattac tcttaacgaa attttgaagt cacttctaca cagttaagtc tgagtaggcg 660

ttatccaaac gtaaa 675

<210> 643

<211> 174

<212> DNA

<213> Hepacivirus sp.

<400> 643

acatgggggg gggctgacag tgagtacact gtgccaagca ggtgctacgc tatgcctagg 60

tgctgctgta ggccaaggac atgtcccagt catcccaggt gagggggggg gttcccctca 120

ccgctgccac tgcctgatag ggtcctgccg gagggtctcg gtgtccggct gtac 174

<210> 644

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
Harrier picornavirus 1 sequence"
<400> 644
gatgtgtgac ggtgtaatta ctttccggat cccactttcc tattataact ctttcatccc 60

aaggttaggg aaagaacctg gctcggtacc accagaccct ccgccacgct agtggactct 120
ccggagataa cggtaccccc tttgtagtca cctgtgctgg tgaagaacca cctagtattg 180
cttgggtgcg tgccgcctag cttccatttc ttctggagca ctgtgcaatg aggtttcccc 240
acttggtaac aagtgcctca ggtcccgcaa ggatactgtg gggtggtgtg accgcaggga 300
gctgtctcca cggctcctct aatgttacgc cgctatccac aggccagtgc gtgtcatcga 360
tcccggatga cagagctagt attgcgaacc cccaagtaag aaaagtggct agtaacctga 420
tagctggtga agagggtggg tcagttgagt agatgcccta gaggtacccg aaaggatctg 480

actagggacc cgtgactata cattaggtaa accgggtata aaaaccatga aaaactgacc 540
acttttcttt taacctcttc tttcttttta tgtgtgttaa gtgttttgag tttggactgt 600
accttgcccg cctttcttgg attttctcta tcgctcttct tacacctact gttatcaagg 660
cactctttag agata 675
<210> 645
<211> 500
<212> DNA
<213> Kunsagivirus sp.
<400> 645
tttcaaatcg gactccggta gttataccgg agcccggttt ggacgcttgg gccgcgttaa 60
cagccccca ccccttccc actgactgat ttctcggatt ggactcattt gcattgctaa 120

ctctgattct ggatttcccc gtttatgtcg tcgcggtcgg aagtgcacgt acttcgacgt 180
tgatctgatg gcgtttgttt ccaggggga ggtggcggca gaaacgcccc cgccgtaaac 240
ttcggcgggc cacgcctgtc aagccactcc ctggggccga gcgcctgagg tgatacagag 300
agataagcac actgggcgct gacaacgccc gggacctcag tgagaagagc agtagggccg 360
tgtttatggg actccattgg atatcccccg cttgtcggaa ctcacggcta ctccgggttg 420
ggaagcccgc gactggtact gtactgggtg atagcctggt gccttccctc tcactgttgt 480
atgaaggctg aaaaccccct 500

<210> 646
<211> 675

<212> DNA

<213> Kobuvirus sp.

<400> 646

tatagtttgc ctgtttctcg caccgttacc gctcgttcgg gaatgtgaaa ctggcacccc 60

tcctctcccc taccaccctt tctccttcgc ccccattcat aatttacaac gccgcacaca 120

gcggcggccg ccaagggcta gcctggcggt tataaaagga acctgggtct ttccctcttc 180

aagccaaaag gtaggttccc tgtgtccctg aatgctcggt gaggaatgct gcaccgtaac 240

gctttgtgaa gtgtttgcaa gttctggccc ggcaagccta cagagtgctg tgatccgctg 300

cggacgccat cctggtaaca ggacccccag tgtgcgcaac agtatgttca gacttcggtt 360

tgttcacttg ctttcatgga ccattgcgcg aaagtgcgtg cgccatatcc ctgtacttca 420

ggtgtgcttc tctggaccct aggaatgctg cgaaggtacc ccgtttcggc gggatctgat 480

cgcaggctaa ttgtctatgg gttcagtttc ctttttcttt actccacaat tgactgctta 540

actgactctg gatcttgtgc ttccactgct ctttctgctc tcaaaacggc ttcacttacc 600

aactctcacc tttcgaccaa caccatttac acactaactt ttttcgactc ttctgactcc 660

tggcttggtg aagac 675

<210> 647

<211> 608

<212> DNA

<213> Kashmir bee virus

<400> 647

tacgtacaat tttgacgctt cgttcatgca acaatgaact cacatgtggc gctcggtagt 60

aaccagaggg gcgtcattcc cccgtatgga gtggagaata taagctaccg actcgagctg 120

tagaataatt cagcaactta taacgaacac gaattttagt cgacgaaacc attttagcct 180

ttaattctat gatttgatta taatgataaa cagctcatgt aactgtctaa actacataaa 240

tacaactgga ttacgaacct ttaagtaact atcttagatg aagtctagta gtctcccaat 300

atttccgtga aaagaatgag ggacgagata gctctattta aagacgtgag gctttaaatt 360

ctgataaata cattacctga gaattcctcc tttaggagtt agaatttgaa aagattagta 420

cctatactta atagaattta aatatgttaa taatgctgaa gacaagtatt tcgattatta 480

acctctatat ttctatataa agtatctgtg agtctcagtg gacatcacag taaggtcgca 540

gttaacttgt aatcttttca ttcctgtgtc ggagcagtgg taatggagcc ggacgatttc 600

gccaaaac 608

<210> 648

<211> 332

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Jingmen picorna-like virus sequence"

<400> 648

tgtgtttttg tcaagataat tgttctgtga ttaacagtga ttgtggttcg tgtaatgcga 60

cgcagtcaaa tgctagtttt gatgaagtgt atgagagagt ggaaaactta tctcataaga 120

agattgaaga gtgtgtagat caggctattg atcgagcttc taagcttcgt gattacaagc 180

ttaatgttca caatggctcc cgacgggaat catctgatcc tctctttatt tcgccccatt 240

cgttgttatc gcttggggta tctaagtttg ttgcgtttga gcagcatcac agtttcgctt 300

cagttgagtc tctgaagttg cttgctctgt ct 332

<210> 649

<211> 145

<212> DNA

<213> Mumps orthorubulavirus

<400> 649

accaagggga aaatgaagat gggatgttgg tagaacaaat agtgtaagaa acagtaagcc 60

cggaagtggt gttttgcgat ttcgaggccg ggctcgatcc tcacctttca ttgtcgatag 120

gggacatttt gacactacct ggaaa 145

<210> 650

<211> 138

<212> DNA

<213> Mosavirus sp.

<400> 650

gaagttgatc atgaacttgg ttattggtgg aacgcacatg aactcccaac aatgatcttg 60

aagacacagc gtggtaacaa ttaccatgcc cttgtggctg cccaagacat tgatggctat 120

tgggtgattt atgatgac 138

<210> 651

<211> 675

<212> DNA

<213> Unknown

<220><

221> source

<223> /note="Description of Unknown:

Miniopterus schreibersii picornavirus 1 sequence"

<400> 651

gttgtcgacc cgttgcttgg tttaagcatg aggtggattc ccccgattat gtctacccgt 60

tactatggcg ggcggtcgtt tcagggtgtt tctactgagg actgcaccaa gtctttatct 120

tttattctca gatctccggc tgtttgacgc ttgtaggaca gcaggactat tttctcttaa 180

tctttttcta cccactaggg tcctatccta gtggagggag ggtgccaccc cttctctctt 240

tagagagtgc gcctggcggt ctttccgtct ctggaaaaag gagcacatgg catgctacaa 300

ttggcacaag aaaacaagct ttgcggattc tttctttgta ctagaggaag ctgtagcgac 360

cctgtatggc gagcggactc ccctctcggc gacgagagcc tctcgggcca aaagccaagt 420

gttaatagca cccatacagg cggcagtacc ccactgccct ttctcaacat acaatgactg 480

atgaaccact gttggttttt ctgacacctt tgtagtagga ttccaaggaa tgtcctgaag 540

gtaccctgtt agcttacgcg caggatctga tcaggagtct ctttacagtg ctgtacactg 600

tgcaagggat taaaaattgt ttgaggaatc cccgagatag tggtctatct cttcctattt 660

tgttttacag acacg 675

<210> 652

<211> 381

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Linda virus sequence"

<400> 652

gtatagcagc agtagctcaa ggctgctata cgattggaca taccaaattc caattggtgt 60

tagggaccac ctaggtgaag gccgacgaca ggtagccatt cctgttagta ggacgaaccg 120

ttatggtgga ctggttgctc aggtgagcag gctgcaatgc gtaagtggtg agtacaccac 180

agccgtcaaa ggtgccactg gtaaggatca cccactggcg atgccttgtg gacgggggcg 240

tgcccaacgc aatgttagcg gtggcggggg ctgccatcgt gaaagctagg tcttgatgga 300

ccttgttgcc tgtacagtct gataggatgc cggcggatgc cctgtgacag ccagtataaa 360

gaatatccgt tgtgattgca c 381

<210> 653

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Lesavirus 2 sequence"

<400> 653

tctttcttta ttttcttatg taactcttct ttttaagttt tattttgcct acttgtgagc 60

ttatgcggga ccactgtctt agacaacccc acatttgtca tgagtaagta cacgcaacca 120

ttacgattac tttttaaccg tctgaccttt tgataacaac tgaagttagg cgtgaaacat 180

gcatttatac caaagtagcc ccgcatttcc ccactacggt gggggggcta ccctactggc 240

tttggaactg tagccattat gtgttgcctg gctttcagga tctcacaaca caacagttct 300

ctcacaatgg aatatgggtg agattgcagt gacatgaaca agtatctagt agtacataga 360

ctcaagccta gttgcctgcg gaacaacatg tggtaacaca tgccccaggg tccaaaagac 420

aagggttaac agccccttct aggtgtctgt gtgtgaagaa tactttagta gtgttgttat 480

gatctcacct gttagtacag aatgagtatg gcttggtgaa ggatgtccta caggtaccca 540

ttatatggat ctgagtagga gaccactagt ggtggcttta ccgccaggtg agtggtttaa 600

aaagcgtcta gccaagccaa cagcactagg gatagtgctt tctatatttt atattttcag 660

tgtatatggt gacaa 675

<210> 654

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Lesavirus 1 sequence"

<400> 654

gtaactaata agcaagtttt actgcctgca aactgcttca atgggaccac cgcttcggcg 60

acccctttgt tgagtttgta tgtttttaag taatctttgc aaccatacga ttattttagc 120

cgcctctcta taatgatctt gttatagtgg gacgtgaaac attggttttt ctcacacacg 180

tccggtcacc cgggcgtgtg ttcttccgta agtcctatcc acataccatc gtgggtaggc 240

cagcatgttt gcacaggctg tgacagtgtg ggtgggcttt ccacctctca acaacacact 300

gaattgcaat gcactcacgg aggaaatgac aatttggtta tagttttgaa ctgtgctagt 360

aatttctttc acattaagcc atgttgcctg cggaatcaca tgtggtaaca catgcctcag 420

ggcccaaaag gcacgggtta gcagcccctt catggtgtgt tagaagtgaa aacacatagt 480

atgagctata atatctttgt tgtcttctct gtagtgtacc ccgccaaatg taaggatgcc 540

cagcaggtac ccattttatg gatctgagct ggggattgat agtgtatcta taaatgcact 600

gatcaattta aaaagcgtct aagtttggca caaacactgg ggacagtgtt tttcctttat 660

tcttttattt gatta 675

<210> 655

<211> 638

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Phopivirus strain NewEngland sequence"

<400> 655

gggagtaaac ctcaccaccg tttgccgtgg tttacggcta cctatttttg gatgtaaata 60

ttaattcctg caggttcagg tctcttgaat tatgtccacg ctagtggcac tctcttaccc 120

ataagtgacg ccttagcgga acctttctac acttgatgtg gttaggggtt acattatttc 180

cctgggcctt ctttggccct ttttcccctg cactatcatt ctttcttccg ggctctcagc 240

atgccaatgt tccgaccggt gcgcccgccg gggttaactc catggttagc atggagctgt 300

aggccctaaa agtgctgaca ctggaactgg actattgaag catacactgt taactgaaac 360

atgtaactcc aatcgatctt ctacaagggg taggctacgg gtgaaacccc ttaggttaat 420

actcatattg agagatactt ctgataggtt aaggttgctg gataatggtg agtttaacga 480

caaaaaccat tcaacagctg tgggccaacc tcatcaggta gatgcttttg gagccaagtg 540

cgtaggggtg tgtgtggaaa tgcttcagtg gaaggtgccc tcccgaaagg tcgtaggggt 600

aatcaggggc agttaggttt ccacaattac aatttgaa 638

<210> 656

<211> 377

<212> DNA

<213> Pestivirus sp.

<400> 656

gtatacgaga ttagctcata ctcgtgtaca aattggacgt agcaaattta aaaattcgga 60

tagggtcccc atccagcgac ggccgaacgg ggttaaccat acctctagta ggactagcag 120

acggatggac tagccacagt ggtgagctcc ctgggaggtc taagttctga gtacagaaca 180

gtcgtcagta gttcaacgct ggtaaacccc agccttgaga tgctacgtgg acgagggcat 240

gcccaagaca caccttaacc tggacggggg tcgtccaggt gaaagtaccc atctttgggt 300

gctgggagta cagcctgata gggcgctgca gaggcccact gacaggctag tataaaaatc 360

tctgctgtac atggaac 377

<210> 657

<211> 494

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Quail picornavirus QPV1 sequence"

<400> 657

tttgcatcag ttcgcccctc ccctcaccat acctttttcc cctctttagg actgatactt 60

ggttatgatg agcagaggat ttcgcaagtt atgcttcttg ataaaaagta attcacgaat 120

catgggatta tagcctggaa gtgaacactc atgtggcaag tgggttagta gctctccatg 180

ttcccatgtg cagtggactg acaacagtga gttcggggtt gtgtagtaaa gggaaagtat 240

tacttacccg cacctgctat acgtggtgta cgtaggatac gagttagtag tgcttagcaa 300

ctttaaactg gtgctgaaat attgcaaggt cactgaagtt gtgaacgcga acgctccgcc 360

actgccatgt atagcgtgca atgcataaat ggtgcactac atgatacgag ggaatgggaa 420

accctccatg gccgaatgca gggtgacagc ctgccggcgg atgcctgttg ttagtataat 480

ccgttgtttg ccac 494

<210> 658

<211> 443

<212> DNA

<213> Sapelovirus sp.

<400> 658

acactcattt ccccccctcca cccttaaggt ggttgtatcc cctacaccct accctccctt 60

ccacatagga cgaataaacg gacttgagat taaggcaagt acataaggta tggtttttgg 120

atacacttaa atggcagtag cgtggcgagc tatggaaaaa tcgcaattgt cgatagccat 180

gttagcgacg cgcttcggcg tgctcctttg gtgattcggc gactggttac aggagagtag 240

acagtgagct atgggcaaac ccctacagta ttacttaggg gaatgtgcaa ttgagacttg 300

acgagcgtct ctttgagatg tggcgcatgc tcttggcatt accatagtga gcttccaggt 360

tgggaaacct ggactgggtc tatactgcct gatagggtcg cggctggccg cctgtaacta 420

gtatagtcag ttgaaaaccc ccc 443

<210> 659

<211> 189

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Porcine reproductive and respiratory syndrome virus 2 sequence"

<400> 659

atgacgtata ggtgttggct ctatgccttg gcatttgtat tgtcaggagc tgtgaccatt 60

ggcacagccc aaaacttgct gcacagaaac acccttctgt gatagcctcc ttcaggggag 120

cttagggttt gtccctagca ccttgcttcc ggagttgcac tgctttacgg tctctccacc 180

cctttaacc 189

<210> 660

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 660

gaaccttaga agtttacaca aacaaagacc aataggagtc caacacccag ttggattgcg 60

gtcaagcact tctgtttccc cggacctagt agtgataggc tgtacccacg gccgaagatg 120

aacccgtccg ttatccggct acctacttcg ggaagcctag taacattctg aagtctctga 180

ggcgtttcgc tcagcacgac cccggtgtag atcgggctga tgggtctccg catacccCac 240

gggcgaccgt ggcggaggcc gcgttggcgg cccgcctatg gcgaaagcca taggacgcct 300

cttagatgac agggtgtgaa gagcctactg agctgggtag tagtcctccg gcccctgaat 360

gcggctaatc ctaaccacgg agcgtccacc agcaatccag ctggcagggc gtcgtaacgg 420

gcaactctgt ggcggaaccg actactttgg gtgtccgtgt ttccttttga tcctattttg 480

gctgcttatg gtgacaacga taagttgtta tcataaagct cttgggttgg ccacctggaa 540

aaagttatca gtgtttgata ttgttcggct ctcacgccta ccaataagac aagccctata 600

tttacttgtt gcattttact cgtcagaaga aatcacagag tatcatttgg atttgttact 660

cacattaagg acaag 675

<210> 661

<211> 591

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Pigeon picornavirus B sequence"

<400> 661

tttatttagc tgttaagttt tatttgtgcc gagaccccat agtaggatct tggtgttcca 60

cattaagctc tcccgaccac acatccaaac gataggcggt gtaagggctc cctggctaag 120

tgtttactca ttgctaggga agtgttgcga cccgttacca gtaggaatac aggaggtctt 180

agttgcctaa ccagataaag tggtgctgaa atattgcaag ctcaatgtct ggcgaacgac 240

ggactaccgt tgaactattg ttaacgcccg cgtgtgtagg caacacacgg gttagtaggt 300

cacttacatt gacatccgtg ccgggaaagc ggatctgagc tatcgattgc ctgatagggt 360

gccggcgggc gcggtacgtg tggtatagtc cgctgtcttg gggtatggcg tctactcggt 420

tttgttgtcg ttttggaatg tcccatgtcg gagatgtcgt taccggtgtt tgctttactt 480

gtgcacgaat aagaaaacag agtttggagt tttggaagtt aatggataac atcaagtttc 540

cattgcgttc ctcgtgcatt ttaggccaga aaatcgacca caaaatcgag a 591

<210> 662

<211> 620

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Picornavirus HK21 sequence"

<400> 662

aaacgggagg atcggctttg gcttatcctc ttaatagtct acaaaactgg gctgactggt 60

gggggagcta ctagatccgg gagaggacta gttaccccgc gtaacttccc tttttgtcac 120

tccctccacc taccccttcc ctgtccctta gactcttaag taggtgtgca ggcctggccc 180

cccggaaatg ggcaaatcgg acgtttgcgt gtagggtcgt tctgaaaggt ggggcccact 240

ccaccgtagt aggatcttcc tgtttcacgt ttaaacctct ccgggcaggt atcgctagac 300

actaggctgt ataggatggc acgcacttag gtttcgcacc ttcctagtgc gccaagatcc 360

cgcttttgag ctcaagtaca tggttctttg taagatgtct aaccagagga ggtggtgctg 420

aaatattgca agccactctg gcgaacgtcc actgcattgc ctggaacagg ctctctagcg 480

ccctccactg gtagcgcggt ggtgggttag taggatacct atatggacag gggatgcggg 540

aataccctc actagctagt gactggttga tcgactggcg gcggatccag tgatacttgc 600

ataatccgca gacttgggag 620

<210> 663

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Picornavirales Tottori-HG1 sequence"

<400> 663

agccatgaac tagtgtgcga ttcccacagt gttggtcgaa agccccgcac tgtctactcg 60

catatttgac tccagtcctt ccgcagccag cggttaggtt ctggttaaag tgcattatgt 120

gcacggcgcc acgcagaact gccagaaatg gtaagctgcg cccaacgcca acggtttgtg 180

tatcccgtta gtcacacgtt tacagctgtt cgttaacagt agggttttgt cgacccggac 240

cccttaatgc gcgcgaattc aaccgcgcct agtcggcaat ggtatcattt aatcccatgc 300

actacgggag aaatttgaga ccaaagaatt cctgagggcc actgcttgct ctaagtgcaa 360

tgcctcggga gacttctgtc aggagcctag cggctttcaa ccgcgacagc taactcctgc 420

gggatgtttg gtgtccatac ttactggcgt cctcacaacg ctaagtggat gttgtccaca 480

ggtaggcaaa caccgagccc cacattcagg agacctgtat gaacgatcct atcagcatta 540

gagttggaat tgggtgtgct aacgtccgca taagtgcacc ccgtggtaac gctgggaaac 600

tatccagcgc aacgtactgt cctcaatgtc tagggaagga ccgccctaag cgtacaaccg 660

ggccatgtgt cgagc 675

<210> 664

<211> 549

<212> DNA

<213> Hepatovirus sp.

<400> 664

ttcaaagagg ggtccgggat tttcctggtc cccctctttg ggcacccttg gctcgggggt 60

gtgaataccg tgctcgcgtt tgccgtgcgt taacggcttc atttatgttt gtttgtctgt 120

tttattatgt tggtttgtct gtttgttatg ttggtattgt tcgtgtttaa tgttatgacc 180

acattacact ccagccaatg aagaacagat ggtgcggtta ttgctggcgg aattcctaac 240

gtcctggatc cgttggtacg catcacaaaa caatttgcag agagagtggt gaaacggctt 300

gggaatccct gagtacaggg aaatcacact gatagctcat cttggctgtt ttcagtcatg 360

gaccttatgc agtgtaattt gggtgtaccc cccatagctt aggaggaatg ttctgtcttg 420

gcactagagt gggacgctga tgcctccgtg tctaggatgg tctaagggac agaatggggt 480

gcctctgatg ccatactacc tgatagggtg ctctcacggc ctctgcatct tagtgagaag 540

ttcaatttt 549

<210> 665

<211> 107

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Rinderpest virus sequence"

<400> 665

accaaacaaa gttgggtaag gatcggtcta tcaatgatta tgatttagca cacttaggat 60

tcaagatcct atcgactgga gcaggcttaa ggtaaaggtt ctttaaa 107

<210> 666

<211> 675

<212> DNA

<213> Rabovirus A

<400> 666

ctacggatat ttgcatgacc cgctttctat cgccccaaca atccccttg taaccacaag 60

ctttactcag gctagcagcc cgactagctg tttggaagaa aaggctaggg cacacaccaa 120

caacaccgac cccactggtc gaaggccgct tggcaataag actggtggaa cagggtcgcc 180

tgtagttgtt tggaacattc tttctaatga ctttgtcagc ggtgctactc acaccgtaac 240

tcttctaccc tatccccacg cttgtggaac taggagggga tgagtgattc aagtaagtac 300

tgtcagaatg gtgaaaatga tctgattctg aaacgctatg gatccatcga aagatggggc 360

tacacgcctg cggaacaaca catggtaaca tgtgccccag gggccgaaag ccacggtgat 420

aggatcaccc gtgtagtttg agatcatatc aatgttcata gtctagtaag atgatttgaa 480

atctaactga gctgatggct aactgcttgt cttattgcgg cctaaggatg tcctgcaggt 540

acctttagat aaccttaaga gactattgat ctgagcagga gccaaagtgg tctttcccag 600

ctttggttaa aaaacgtcta agccgcggca gggggcggga ggcccccttt cctcccaaaa 660

cttaatattg attgt 675

<210> 667

<211> 410

<212> DNA

<213> Shingleback nidovirus 1

<400> 667

ctgtgagtac cgacaggctc gaagtctatt atgaggcgtc gaaacagaaa acctgtaaca 60

actccggttt catctatcac tgccgtcaag aggcagaaga ggacgaccac gtgtcaccag 120

atcacttgta tctgtttcag tcaggaagtc aacttttcga cgaagttcga ccattcatcg 180

acccgctgaa aagcgtagaa gtcgatgaag atgagctcca aagagccatc gccgatttcg 240

acaaccaaag tgactgtttc cactccttcg agctcgtgaa tttcgagctg aaacaacaaa 300

tcaacgagaa cgagtggtac ggttattata attacgacaa ccaaaactgc aaagttcagt 360

tgccagtcac atgtcgaatc gaggacgtaa cctgggatca ggtttacgtg 410

<210> 668

<211> 666

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Seneca valley virus sequence"

<400> 668

tttgaaatgg ggggctgggc cctgatgccc agtccttcct ttccccttcc ggggggttaa 60

ccggctgtgt ttgctagagg cacagagggg caacatccaa cctgcttttg cggggaacgg 120

tgcggctccg attcctgcgt cgccaaaggt gttagcgcac ccaaacggcg cacctaccaa 180

tgttattggt gtggtctgcg agttctagcc tactcgtttc tcccccgacc attcactcac 240

ccacgaaaag tgtgttgtaa ccataagatt taacccccgc acgggatgtg cgataaccgt 300

aagactggct caagcgcgga aagcgctgta accacatgct gttagtccct ttatggctgc 360

aagatggcta cccacctcgg atcactgaac tggagctcga ccctccttag taagggaacc 420

gagaggcctt cgtgcaacaa gctccgacac agagtccacg tgactgctac caccatgagt 480

acatggttct cccctctcga cccaggactt ctttttgaat atccacggct cgatccagag 540

ggtggggcat gacccctagc atagcgagct acagcgggaa ctgtagctag gccttagcgt 600

gccttggata ctgcctgata gggcgacggc ctagtcgtgt cggttctata ggtagcacat 660

acaaat 666

<210> 669

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Sclerotinia sclerotiorum dsRNA mycovirus-L sequence"

<400> 669

ttgaattaat cttttacgtt tacgcgcata aaatcaggac acatctcttg tatactttag 60

tatatcaatg atgtttttgt tttatgcgat taatcgtaag agaacttctt tccatccgcc 120

tgtatgggcg ggataataag ttcaccgcct tggtcgaggc gcaaacttgt atgtgcaaag 180

gtgagctata tgctcgaaat agtcgtaact aacacacagc cactacctgt agagctctat 240

tgatccggaa tcctttagtg ggaatgcaga gctcacaccg gacctgcggg tatcttcggc 300

gttagggact tctgtttcag ccttgaatca tttaccttta taccttctct gaggcgcctg 360

ggccgggcgc gatattaagt acaagtcaag gacatcgcgg gtagtggtct aatcagccgc 420

tagtcctgct ggagagttcc aacttagttg ggtgtggtgc atactagctg gatagagtag 480

gtatgtattg ctaacgtatg ccggaggcta tccgtcctcg gtagaacgtg ccgaggagta 540

gtctctgcag acccccgaac gcgtggggtc tttacttaaa tgtaggcgga gggagcgctc 600

gtaggtggaa cgactgcctc ccagtcgaat gcaagatttt gcacgcggac cagtctgccc 660

ggcaattccc gggtg 675

<210> 670

<211> 675

<212> DNA

<213> Enterovirus sp.

<400> 670

ctccggcaca gccgcaccag tgcactggta cgctagtacc ttttcacggg gtagtcggta 60

tccccccccg taacttagaa gcatgtaaca aaccgaccaa taggtgcgcg gcagccagct 120

gcgttgcggt caagcacttc tgtctccccg gtccgcaagg atcgttaccc gcccactcca 180

ctacgaggag cctagtaact ggccaagtga ttgcggagtt gcgttcagcc acaaccccag 240

tggtagctct ggaagatggg gctcgcacat cccccgtggt aacacggttg cttgcccgcg 300

tgtgcttccg ggttcagtct ccgactgttc acttcaacat cacgcaacca gccaagagcc 360

gattgtgctg gagtggtctt cctccggggc cgtgaatgct gctaatccta acctccgagc 420

gtgtgcgcac aatccagtgt tgctacgtcg taacgcgtaa gttggaggcg gaacagacta 480

ctttcggcac cccgtgtttc ctttatttta ttcttatttt atggtgacaa ttgcagagat 540

ttgtgatatt gcgactttac cgttaaacat agcactgcat tacctggttg cattccacaa 600

aacttcagag attcctagtt cctacattga cctacttgtt tatttgaatc ttaaatacaa 660

acttgagcaa gtgaa 675

<210> 671

<211> 245

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Wobbly possum disease virus sequence"

<400> 671

cggctgtgag tgcttagcat atgctagagt actacagccg ggtgttggag tcatatgcac 60

tggttgcctg tataatagtc gggatctgtc tgacctacat tatctttggg agttgcttat 120

cacgacaatt ctcgaagtgt ctgtcgacag cttacccccg attcgacaag gccccttgtc 180

caccgcagac ctatcgattt tcaacgagaa cactatcaga ggtttaaatt taaaactcac 240

caaca 245

<210> 672

<211> 631

<212> DNA

<213> Avian orthoreovirus

<400> 672

gctttttcaa tcccttgtgt cgatgttccc gtatgtcatc tggttcatgt aacggtgcaa 60

cttctatttt tggtaacgtt cactgtcagg cagcgcaaaa ctcggcgggt ggtgatcttc 120

aagcgacttc ctctcttgtt gcttattggc cctacttggc tgcgggcggt ggtctgctcg 180

ttgttattat tataattgtt ggcgctgttt gttgttgcaa ggctaaggtt aaagcggacg 240

cagcccgaaa cgttttctac cgagagctgt tcgcacttaa ttcgggtaaa agtgatgcag 300

gacctccgat ttaccaggtt tagtgtacga cgatttgagt tttcaccttt cgtcttagag 360

gagtgcacta ctccatcttt cacgactata actaataccg atccggctct ctactttaac 420

attgagtttc cgtcaagtca tcgtctctcc cccttcattc cagaactgtt gtctcagcct 480

tgtaccgttc acgtttcatt gattcggaga ttcgctctct gtgcaacctt atctagtatt 540

tgtgaatacg actgtgcgct actgccatcc atcaacgcta ttacgacgat ccctacacca 600

ggtgcgtcat catctctgat tgttcattgg g 631

<210> 673

<211> 665

<212> DNA

<213> Kobuvirus sp.

<400> 673

ggggcctcgg ccccctcacc ctcttttccg gtggccacgc ccgggccacc gatacttccc 60

ttcactcctt cgggactgtt ggggaggaac acaacagggc tcccctgttt tcccattcct 120

tccccctttt cccaacccca accgccgtat ctggtggcgg caagacacac gggtctttcc 180

ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg ggagtgctcc 240

cacccaactg ttgtaagcct gtccaacgcg tcgtcctggc aagactatga cgtcgcatgt 300

tccgctgcgg atgccgaccg ggtaaccggt tccccagtgt gtgtagtgcg atcttccagg 360

tcctcctggt tggcgttgtc cagaaactgc ttcaggtaag tggggtgtgc ccaatcccta 420

caaaggttga ttctttcacc accttaggaa tgctccggag gtaccccagc aacagctggg 480

atctgaccgg aggctaattg tctacgggtg gtgtttcctt tttcttttca cacaactcta 540

ctgctgacaa ctcactgact atccacttgc tctcttgtgc ctttctgctc tggttcaagt 600

tccttgattg tttttgactg cttttcactg cttttcttct cacaatcctt gctcagttca 660

aagtc 665

<210> 674

<211> 655

<212> DNA

<213> Kobuvirus sp.

<400> 674

ccccctcacc ctcttttccg gtggccacgc ccgggccacc gatacttccc ttcactcctt 60

cgggactgtt ggggaggaac acaacagggc tcccctgttt tcccattcct tcccccttt 120

cccaacccca accgccgtat ctggtggcgg caagacacac gggtctttcc ctctaaagca 180

caattgtgtg tgtgtcccag gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg 240

ttgtaagcct gtccaacgcg tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg 300

atgccgaccg ggtaaccggt tccccagtgt gtgtagtgcg atcttccagg tcctcctggt 360

tggcgttgtc cagaaactgc ttcaggtaag tggggtgtgc ccaatcccta caaaggttga 420

ttctttcacc accttaggaa tgctccggag gtaccccagc aacagctggg atctgaccgg 480

aggctaattg tctacgggtg gtgtttcctt tttcttttca cacaactcta ctgctgacaa 540

ctcactgact atccacttgc tctcttgtgc ctttctgctc tggttcaagt tccttgattg 600

tttttgactg cttttcactg cttttcttct cacaatcctt gctcagttca aagtc 655

<210> 675

<211> 645

<212> DNA

<213> Kobuvirus sp.

<400> 675

ctcttttccg gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt 60

ggggaggaac acaacagggc tcccctgttt tcccattcct tcccccttt cccaacccca 120

accgccgtat ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg 180

tgtgtcccag gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct 240

gtccaacgcg tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg 300

ggtaaccggt tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc 360

cagaaactgc ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc 420

accttaggaa tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg 480

tctacgggtg gtgtttcctt tttcttttca cacaactcta ctgctgacaa ctcactgact 540

atccacttgc tctcttgtgc ctttctgctc tggttcaagt tccttgattg tttttgactg 600

cttttcactg cttttcttct cacaatcctt gctcagttca aagtc 645

<210> 676

<211> 635

<212> DNA

<213> Kobuvirus sp.

<400> 676

gtggccacgc ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac 60

acaacagggc tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat 120

ctggtggcgg caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag 180

gtcctcctgc gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg 240

tcgtcctggc aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt 300

tccccagtgt gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc 360

ttcaggtaag tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa 420

tgctccggag gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg 480

gtgtttcctt tttcttttca cacaactcta ctgctgacaa ctcactgact atccacttgc 540

tctcttgtgc ctttctgctc tggttcaagt tccttgattg tttttgactg cttttcactg 600

cttttcttct cacaatcctt gctcagttca aagtc 635

<210> 677

<211> 625

<212> DNA

<213> Kobuvirus sp.

<400> 677

ccgggccacc gatacttccc ttcactcctt cgggactgtt ggggaggaac acaacagggc 60

tcccctgttt tcccattcct tccccctttt cccaacccca accgccgtat ctggtggcgg 120

caagacacac gggtctttcc ctctaaagca caattgtgtg tgtgtcccag gtcctcctgc 180

gtacggtgcg ggagtgctcc cacccaactg ttgtaagcct gtccaacgcg tcgtcctggc 240

aagactatga cgtcgcatgt tccgctgcgg atgccgaccg ggtaaccggt tccccagtgt 300

gtgtagtgcg atcttccagg tcctcctggt tggcgttgtc cagaaactgc ttcaggtaag 360

tggggtgtgc ccaatcccta caaaggttga ttctttcacc accttaggaa tgctccggag 420

gtaccccagc aacagctggg atctgaccgg aggctaattg tctacgggtg gtgtttcctt 480

tttcttttca cacaactcta ctgctgacaa ctcactgact atccacttgc tctcttgtgc 540

ctttctgctc tggttcaagt tccttgattg tttttgactg cttttcactg cttttcttct 600

cacaatcctt gctcagttca aagtc 625

<210> 678

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Picornavirales sp. isolate RtMruf-PicoV sequence"

<400> 678

tttgctcagc gtaacttctc cgggttacgt ggagaccaaa aggctacgga gactcgggct 60

acggccctgg agcacctagg tgctcctaaa gacgttagaa gttgtacaaa ctcgcccaat 120

agggcccccc aaccaggggg gtagcgggca agcacttctg tttccccggt atgatctcat 180

aggctgtacc cacggctgaa agagagatta tcgttacccg cctcactact tcgagaagcc 240

cagtaatggt tcatgaagtt gatctcgttg acccggtgtt tcccccacac cagaaacctg 300

tgatgggggt ggtcatcccg gtcatggcga catgacggac ctccccgcgc cggcacaggg 360

cctcttcgga ggacgagtga catggattca accgtgaaga gcctattgag ctagtgttga 420

ttcctccgcc cccgtgaatg cggctaatcc caactccgga gcaggcgggc ccaaaccagg 480

gtctggcctg tcgtaacgcg aaagtctgga gcggaaccga ctactttcgg gaaggcgtgt 540

ttccttttgt tccttttatc aagttttatg gtgacaactc ctggtagacg ttttattgcg 600

tttattgaga gatttccaac aattgaacag actagaacca cttgttttat caaaccctca 660

cagaataaga taaca 675

<210> 679

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Apodemus agrarius picornavirus strain Longquan-Aa118 sequence"

<400> 679

ttactcagcg taactactcc gggttacgtg atgaagaaga ggctacggag attctcgggc 60

tacggccctg gagccactcc ggctcctaaa gatttagaag tttgagcaca cccgcccact 120

agggcccccc atccaggggg gcaacgggca agcacttctg tttccccggt atgatctgat 180

aggctgtaac cacggctgaa acagagatta tcgttatccg cttcactact tcgagaagcc 240

tagtaatgat gggtgaaatt gaatccgttg atccggtgtc tcccccacac cagaaactca 300

tgatgagggt tgccatcccg gctacggcga cgtagcgggc atccctgcgc tggcatgagg 360

cctcttagga ggacggatga tatggatctt gtcgtgaaga gcctattgag ctagtgtcga 420

ctcctccgcc cccgtgaatg cggctaatcc taaccccgga gcaggtgggt ccaatccagg 480

gcctggcctg tcgtaatgcg taagtctggg acggaaccga ctactttcgg gaaggcgtgt 540

ttccatttgt tcattatttg tgtgtttatg gtgacaactc tgggtaaacg ttctattgcg 600

tttattgaga gattcccaac aattgaacaa acgagaacta cctgttttat taaatttaca 660

cagagaagaa ttaca 675

<210> 680

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Niviventer confucianus picornavirus sequence"

<400> 680

ccctttcata acccccccct tttaacccaa cccttcgtaa ccgtacgctt cactcgcctt 60

tgggtatagc ggcccaatgt gctgaagaaa aggatacgct ataaggggcc aacgggtggt 120

ggcccttaag accacccaac ctagaagctt gtacactcgg gcaatagtga ggcccacatc 180

cagtgggtca agcccaaagc attcttgttc cccggtatga tctcataagc tgtacccacg 240

gctgaaagag tgattatcgt tatcccactc agtacttcgg agagcctagt acaccacttg 300

gaaatggaag tctgtgatcc ggggttgacc ctgaacccca gaaactcatg atgaggctaa 360

ccttcccgaa cacggcgacg tgtggttagc ctgcgctggc atgaggcctc tttgtaggca 420

gactgaaatg gaagggtgac gaagagccga ctgagctact gttttattcc tccggccccc 480

tgaatgcggc taatcctaac tcctggtcca gtacttgtaa cccaacaggt ggctggtcgt 540

aatgcgtaag ccgggagcgg aaccgactac tttggggcgt ccgtgtttct caatattatt 600

catttctagc ttatggtgac aatttatgat tgcagagatt gtgctgtatt tgtgtctgag 660

agaagaagta acaat 675

<210> 681

<211> 619

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bat picornavirus isolate BtRs-PicoV sequence"

<400> 681

tttcaaaagg ccctgggcat acggcgttat tcgtaacgtc gtatgtccag ggcggtagca 60

tcaggccaag gcctgatgct accacgtgtg gactaaacca cacactcttc ttgtgacacg 120

ttgtgtcacc tatccctttc ttggtaactt agaagcttgt acacttacgc acgtaggtgc 180

cccacatcca gtggggtttg tgcaaagcaa tcttgttccc cggtaaaccc tgataggctg 240

taaccacggc cgaaacaagg tttgtcgtta cccgactcac tactacacaa agcctagtaa 300

agttcaatga aagtgcgcag cgtgatccgg tcaaaacccc cttgaccaga aacacatgat 360

gagggtcacc aacccccact ggcgacagtg tggtgtccct gcgttggcat gtggcctcgt 420

agaggcgttg caatctggat ttgctccgaa gagccccgtg tgctagtgtt tatacctccg 480

gccccttgaa tgcggctaat cctaaccccc gagcatgtac acacaagcca gtgtgtagca 540

tgtcgtaatg agcaatttgg ggatggaacc gactacttta gggtgtccgt gtttctcatt 600

attctttgtt tgatgtttt 619

<210> 682

<211> 593

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Rhinolophus picornavirus strain Guizhou-Rr100 sequence"

<400> 682

tttttttttct caggcggtag catccagcca aggcctgatg ctaccaacgt gtgactaaac 60

cacactctct ttttgtgata cattgtgtca cctatccctt tcttggtaac ttagaagctt 120

gtacacccac gcacgtaggt accccacatc cagtggggtt tgtgcaaagc attcttgttc 180

cccggtaaac cctgataggc tgtaaccacg gctgaaacaa ggtttgtcgt tacccgactc 240

actactacgc aaagcctagt aaagttcaat gaaagtgcgc agcgtgatcc ggtcaaaacc 300

cccttgacca gaaacacatg atgagggtca ccaaccccca ctggcgacag tgtggtgtcc 360

ctgcgttggc atgtggcctc atagaggcgt tgcaatctgg atttgctccg aagagccccg 420

tgtgctagtg tttatacctc cggccccttg aatgcggcta atcctaaccc ccgagcatgt 480

acacacacgc cagtgtgcag catgtcgtaa tgagcaattt ggggatggaa ccgactactt 540

tagggtgtcc gtgtttctca ttattctttg tttgatgttt tatggtgaca aca 593

<210> 683

<211> 597

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Rhinolophus picornavirus strain Henan-Rf265 sequence"

<400> 683

cggaacgttg tatgctcagg gcgtaggcac cacccacggg tggtgcctac acgtgtggac 60

taaaccacac actcttttca gcacttagtg ctgctatctc tttttgtaac ttagaagttt 120

gtacacaatg cgttagggcc acacatccag tgtggtatcg caaagcactt ctgtttcccc 180

ggtgctagta ggagggtggc tgctccacgg ccacttgccg aacccatcgt tacccgactc 240

attacttcgc aaagcctagt aacccagttg aagcaagccc ggcgtgttcc ggtcaggaaa 300

aacccccccc ggccagaaac atgtgatgag ggtgggctat ccccactggt gacagtgagc 360

cctccctgcg ttggcacatg gcccgatctg ggcgtggttc ttgtggatgc tgccgaagag 420

ccccgtgagc tagtgtttat accgccggcc tcgtgaatgc ggctaaccct aaccccggag 480

cagaggctac tgaagccaca gtagtcgctg tcgtaacgag taattctggg atgggaccga 540

ctactttcga gtgtccgtgt ttcctttatt cttttattgt tgtttatggt gacaaac 597

<210> 684

<211> 636

<212> DNA

<213> Enterovirus sp.

<400> 684

cccctaggat ccactggatg tcagtacact ggtatcgtgg tacctttgta cgcctgtttt 60

atacccccct ccccgcaact ttagaagcat caaaagcacc gctcaatagt caccacaccc 120

ccagtgtggt ttcgagcaag cacttctgtt ttcccggttg cgtcccatat gctgtgcaaa 180

cggcaaaaag ggacaatatc gttacccgct tgtatactac gggaaaccta gtaccaccat 240

tgattgtgtt gagagttgcg ctcatcacct ttccccggtg tagctcaggc cgatgaggct 300

cagaatcccc cacaggtgac tgtgtctgag cctgcgttgg cggcctgccc tcgccttatg 360

gcgtgggacg cttgatacat gacatggtgc gaagagtcta ctgtgctatg caagagtcct 420

ccggcccctg aatgtggcta atcctaacca ctgatcccac gcacgcaaac cagtgtgtag 480

tgggtcgtaa cgcgcaagtc ggtggcggaa ccaactactt tgggtgaccg tgtttccttt 540

attacttatt gaatgtttat ggtgacaatt gtttgattca gttgttgcca ttctctacat 600

tcatttaccc agcatcaaac caattgaact gttaca 636

<210> 685

<211> 421

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human poliovirus 1 strain NIE1116623 sequence"

<400> 685

agtctggaca tccctcaccg gcgacggtgg tccaggctgc gctggcggcc tacctgtggt 60

ccaaagccac gggacgctac atgtgaacaa ggtgtgaaga gcctattgag ctacaaaaga 120

gtcctccggc ccctgaatgc ggctaatccc aaccacggat caagggtgca caaaccagtg 180

tacaccttgt cgtaacgcgc aagtctgtgg cggaaccgac tactttgggt gtccgtgttt 240

cctttttaat tttgatggct gcttatggtg acaatcatag attgttatca taaagctaat 300

tggattggcc atccggtgag agtgaaatat attgtttacc tccctgttgg gtttactcta 360

actaacttct ccatttataa acttgtcatc acagttttaa taattagaag tgcagtttac 420

a 421

<210> 686

<211> 663

<212> DNA

<213> Enterovirus sp.

<400> 686

tttaaaacag cctgggggtt gttcccaccc ccagggccca ctgggcgtta gtactctggt 60

atcgcggtac cttagtatgc ctgttttatg tctcctttcc cccgcaactt tagaagtaat 120

caagttatgg ctcaacagtc gccacacccc cagtgtggtt ccgagcaagc acttctgttc 180

cccggttgcg tcttatatgc tgtgtgaacg gcagaaaggg acaatatcgt tatccgctca 240

actactacgg gaagcctagt accaccatgg attgacctga aagttgcgtt cagcgcaccc 300

ccagcgcagc tcaggccgat gaggctccga ataccccacg ggcgaccgtg tcggagcctg 360

cgttggcggc ctgcccacgt tgcaaaacgt gggacgctca tttcatgaca tggtgcgaag 420

agcctactgt gctagttgag agtcctccgg cccctgaatg tggataatcc taaccactga 480

acctacgggc gcaaaccagc gtctggtagg ccgtaacgcg caagtcggtg gcggaaccaa 540

ctactttggg tgtccgtgtt tccttttatc tttttgaatg tttatggtga caattgttgt 600

gtacagttgt taccatagtt tgcattcaga aataaaccta acactttcca ttatttgtt 660

aca 663

<210> 687

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human poliovirus 2 strain NIE0811460 sequence"

<400> 687

ttgtgcgcct gttttatatt ccccccccgc aacttagaag cacgaaacca agttcaatag 60

aaggggggtac aaaccagtac caccacgaac aagcacttct gtttccccgg tgacattgca 120

tagactgctc acgcggttga aagtgatcga tccgttaccc gcttgtgtac ttcgaaaagc 180

ctagtattgc cttggaatct tcgacgcgtt gcgctcagca cccgaccccg gggtgtagct 240

taggctgatg agtctggaca ttcctcaccg gtgacggtgg tccaggctgc gttggcggcc 300

tacctatggc taatgccata ggacgctaga tgtgaacaag gtgtgaagag cctattgagc 360

tacataagag tcctccggcc cctgaatgcg gctaatccta accacggagc aggcggtcgc 420

aaaccagtga ctagcttgtc gtaacgcgca agtctgtggc ggaaccgact actttgggtg 480

tccgtgtttc ctgttatttt tattatggct gcttatggtg acaatcagag attgttatca 540

taaagcgaat tggattggcc atccggtgag tgttgtgtca ggtatacaac tgtttgttgg 600

aaccactgtg ttagtttaac ctctctttca accaattagt caaaaacaat acgaagatag 660

aacaacaata ctaca 675

<210> 688

<211> 512

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Bovine picornavirus sequence"

<400> 688

ttttctcccc tccccctcca actacctttt ccccctcttg taacgctaga agtttgtgca 60

aaccgcctgt agggtactgc aatccagcag tgcataggct aagcttttct tgttacccca 120

ccccacatta tactgaggag gattgtgaaa ttgtgttagt atgggttagt agcggtgacc 180

cgggtaaccc caacccagaa actcacggat gagatgaaca ggaccccaca tggtaacgtg 240

tgtgttcgtc tgccccgcaa ggtgaggccg tgagagcttt gcacgcgaaa accttgaaaa 300

cccaaaagta ccttgagctc ttcgctattt tgtgtttcct ccaggaccct gaatgcggct 360

aaacctaacc cgcgatccgc acgtagcaac ccagctagag tgtggtcgta atgcgcaagt 420

tgcgggcggt accgactact ttggtgttcc tgtgtttcct ttattttatt ttgaattttt 480

atggtgacaa cagctagaaa ataagagtga ac 512

<210> 689

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human poliovirus 1 strain EQG1419328 sequence"

<400> 689

acccttgtac gcctgtttta tactcccctc cccgtaactt agaagaaaca aaataagttc 60

aataggaggg ggtacaaacc agtaccacca cgaacaagca cttctgtctc cccggtgaca 120

ttgcatagac tgtccccacg gttgaaagca attgatccgt tacccgctct tgtacttcga 180

gaagcctagt accatcttgg aatcatcgat gcgttgcgct ccacactcag tcccagagtg 240

tagcttaggc tgatgagtct ggacattcct caccggcgac ggtggtccag gctgcgttgg 300

cggcctacct gtggcccaaa gccacaggac gctagatgtg aacaaggtgt gaagagccta 360

ttgagctata agagagtcct ccggcccctg aatgcggcta atcccaacca cggatcaagg 420

gtgcacgaac cagtgtatac cttgtcgtaa cgcgcaagtc cgtggcggaa ccgactactt 480

tgggtgaccg tgtttccttt tattatttca atggctgctt atggtgacaa tcattgattg 540

ttatcataaa gcgaattgga ctggccatcc ggtgaaagtg aaacatattg tttgcctcct 600

cgttgggtct acttcaacca atctttactt acaatcttac cactacagtt ttgctggtta 660

gaagtgtgtt tcacg 675

<210> 690

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Human poliovirus 2 isolate IS_061 sequence"

<400> 690

ttgtgcgcct gttttatact cccctcccgc aacttagaag cacgaaacca agttcaatag 60

aaggggtac aaaccagtac cactacgaac aagcacttct gtttccccgg tgacattgca 120

tagactgctc acgcggttga aagtgatcga tccgttaccc gcttgtgtac ttcgaaaagc 180

ctagtatcgc cttggaatct tcgacgcgtt gcgctcagca cccgaccccg gggtgtagct 240

taggccgatg agtctggaca ttcctcaccg gtgacggtgg tccaggctgc gttggcggcc 300

tacctatggc taacgccata ggacgttaga tgtgaacaag gtgtgaagag cctattgagc 360

tacataagag tcctccggcc cctgaatgcg gctaatccta accacggagc aggcggtcgc 420

gaaccagtga ctggcttgtc gtaacgcgca agtctgtggc ggaaccgact actttgggtg 480

tccgtgtttc ctgttatttt tatcatggct gcttatggtg acaatcagag attgttatca 540

taaagcgaat tggattggcc atccggtgag tgttgtgtca ggtatacaac tgtttgttgg 600

aaccactgtg ttagctttgc ttctcattta accaattaat caaaaacaat acgaggataa 660

aacaacaata ctaca 675

<210> 691

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Coxsackievirus B5 sequence"

<400> 691

cctttgtgcg cctgttttat gccccccttcc cccaattgaa acttagaagt tacacacacc 60

gatcaacagc gggcgtggca taccagccgc gtcttgatca agcactcctg tttccccgga 120

ccgagtatca atagactgct cacgcggttg aaggagaaaa cgttcgttac ccggctaact 180

acttcgagaa acctagtagc atcatgaaag ttgcgaagcg tttcgctcag cacatcccca 240

gtgtagatca ggtcgatgag tcaccgcatt ccccacgggc gaccgtggcg gtggctgcgt 300

tggcggcctg cctacggggc aacccgtagg acgcttcaat acagacatgg tgcgaagagt 360

cgattgagct agttagtagt cctccggccc ctgaatccgg ctaatcctaa ctgcggagca 420

cataccctca acccaggggg cattgtgtcg taacgggtaa ctctgcagcg gaaccgacta 480

ctttgggtgt ccgtgtttcc ttttattctt ataatggctg cttatggtga caattgaaag 540

attgttacca tatagctatt ggattggcca tccggtgtct aacagagcta ttatatacct 600

ctttgttgga tttgtaccac ttgatctaaa ggaagtcaag acactacaat tcatcataca 660

attgaacaca gcaaa 675

<210> 692

<211> 675

<212> DNA

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Coxsackievirus A10 sequence"

<400> 692

tttgtgcgcc tgttttacaa cccttcccca acttgtaacg tagaagtaat acacactact 60

gatcaatagc aggcatggcg cgccagtcat gtctcgatca agcacttctg ttcccccgga 120

ctgagtatca atagactgct cacgcggttg aaggagaaaa cgttcgttac ccggctaact 180

acttcgagaa acctagtagc accatagaag ctgcagagtg tttcgctcag cacttccccc 240

gtgtagatca ggctgatgag tcactgcaat ccccacgggt gaccgtggca gtggctgcgt 300

tggcggcctg cctatggggc aacccatagg acgctctaat gtggacatgg tgcgaagagt 360

ctattgagct agttagtagt cctccggccc ctgaatgcgg ctaatcctaa ctgcggagca 420

catgccttca acccagaagg tagtgtgtcg taacgggcaa ctctgcagcg gaaccgacta 480

ctttgggtgt ccgtgtttct ttttattcct atattggctg cttatggtga caatcacgga 540

attgttgcca tatagctatt ggattggcca tccggtgtct aatagagcta ttgtgtacct 600

atttgttgga tttactccgc tatcacataa atctctgaac actttgtgct ttatattgaa 660

cttaaacacc cgaaa 675

<210> 693

<400> 693

000

<210> 694

<211> 30

<212> DNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 694

tataattcta ccctattgag gcattgacta 30

<210> 695

<211> 493

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 695

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

1 5 10 15

His Ala Ala Arg Pro Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu

20 25 30

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

35 40 45

Ser Val Ser Val Ile Gly Ala His Leu Ile His Trp Tyr Gln Gln Lys

50 55 60

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

65 70 75 80

Thr Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe

85 90 95

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

100 105 110

Cys Leu Gln Ser Arg Ile Phe Pro Arg Thr Phe Gly Gln Gly Thr Lys

115 120 125

Leu Glu Ile Lys Gly Ser Thr Ser Gly Ser Gly Lys Pro Gly Ser Gly

130 135 140

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

145 150 155 160

Leu Lys Lys Pro Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly

165 170 175

```
Tyr Thr Phe Thr Asp Tyr Ser Ile Asn Trp Val Arg Gln Ala Pro Gly
            180                 185                 190
Gln Gly Leu Glu Trp Met Gly Trp Ile Asn Thr Glu Thr Arg Glu Pro

        195                 200                 205
Ala Tyr Ala Tyr Asp Phe Arg Gly Arg Phe Val Phe Ser Leu Asp Thr
    210                 215                 220
Ser Val Ser Thr Ala Tyr Leu Gln Ile Ser Ser Leu Lys Ala Glu Asp
225                 230                 235                 240
Thr Ala Val Tyr Tyr Cys Ala Arg Asp Tyr Ser Tyr Ala Met Asp Tyr
                245                 250                 255
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ala Ala Thr Thr

            260                 265                 270
Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln
        275                 280                 285
Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala
    290                 295                 300
Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp Ala
305                 310                 315                 320
Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr

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

385                 390                 395                 400
Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu
                405                 410                 415
Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg Arg
```

420 425 430

Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met

435 440 445

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

450 455 460

Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp

465 470 475 480

Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg

485 490

<210> 696

<211> 207

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 696

Lys Asn Gln Val Glu Gln Ser Pro Gln Ser Leu Ile Ile Leu Glu Gly

1 5 10 15

Lys Asn Cys Thr Leu Gln Cys Asn Tyr Thr Val Ser Pro Phe Ser Asn

20 25 30

Leu Arg Trp Tyr Lys Gln Asp Thr Gly Arg Gly Pro Val Ser Leu Thr

35 40 45

Ile Met Thr Phe Ser Glu Asn Thr Lys Ser Asn Gly Arg Tyr Thr Ala

50 55 60

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

65 70 75 80

Gln Leu Ser Asp Ser Ala Ser Tyr Ile Cys Val Val Asn His Ser Gly

85 90 95

Gly Ser Tyr Ile Pro Thr Phe Gly Arg Gly Thr Ser Leu Ile Val His

100 105 110

Pro Tyr Ile Gln Lys Pro Asp Pro Ala Val Tyr Gln Leu Arg Asp Ser

115 120 125

Lys Ser Ser Asp Lys Ser Val Cys Leu Phe Thr Asp Phe Asp Ser Gln

130 135 140

Thr Asn Val Ser Gln Ser Lys Asp Ser Asp Val Tyr Ile Thr Asp Lys

145 150 155 160

Thr Val Leu Asp Met Arg Ser Met Asp Phe Lys Ser Asn Ser Ala Val

165 170 175

Ala Trp Ser Asn Lys Ser Asp Phe Ala Cys Ala Asn Ala Phe Asn Asn

180 185 190

Ser Ile Ile Pro Glu Asp Thr Phe Phe Pro Ser Pro Glu Ser Ser

195 200 205

<210> 697

<211> 246

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 697

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

1 5 10 15

Glu Lys Val Phe Leu Glu Cys Val Gln Asp Met Asp His Glu Asn Met

20 25 30

Phe Trp Tyr Arg Gln Asp Pro Gly Leu Gly Leu Arg Leu Ile Tyr Phe

35 40 45

Ser Tyr Asp Val Lys Met Lys Glu Lys Gly Asp Ile Pro Glu Gly Tyr

50 55 60

Ser Val Ser Arg Glu Lys Lys Glu Arg Phe Ser Leu Ile Leu Glu Ser

65 70 75 80

Ala Ser Thr Asn Gln Thr Ser Met Tyr Leu Cys Ala Ser Ser Phe Leu

85 90 95

Met Thr Ser Gly Asp Pro Tyr Glu Gln Tyr Phe Gly Pro Gly Thr Arg

100 105 110

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

115 120 125

Val Phe Glu Pro Ser Glu Ala Glu Ile Ser His Thr Gln Lys Ala Thr

130 135 140

Leu Val Cys Leu Ala Thr Gly Phe Tyr Pro Asp His Val Glu Leu Ser

145 150 155 160

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

165 170 175

Gln Pro Leu Lys Glu Gln Pro Ala Leu Asn Asp Ser Arg Tyr Cys Leu

180 185 190

Ser Ser Arg Leu Arg Val Ser Ala Thr Phe Trp Gln Asn Pro Arg Asn

195 200 205

His Phe Arg Cys Gln Val Gln Phe Tyr Gly Leu Ser Glu Asn Asp Glu

210 215 220

Trp Thr Gln Asp Arg Ala Lys Pro Val Thr Gln Ile Val Ser Ala Glu

225 230 235 240

Ala Trp Gly Arg Ala Asp

245

<210> 698

<211> 115

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 698

Met Gln Glu Val Thr Gln Ile Pro Ala Ala Leu Ser Val Pro Glu Gly

1 5 10 15

Glu Asn Leu Val Leu Asn Cys Ser Phe Thr Asp Ser Ala Ile Tyr Asn

20 25 30

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

Gly Gly Ser Tyr Ile Pro Thr Phe Gly Arg Gly Thr Ser Leu Ile Val
100 105 110
His Pro Tyr
115

<210> 699
<211> 112
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 699
Met Gly Val Thr Gln Thr Pro Lys Phe Gln Val Leu Lys Thr Gly Gln
1 5 10 15
Ser Met Thr Leu Gln Cys Ala Gln Asp Met Asn His Glu Tyr Met Ser

20 25 30
Trp Tyr Arg Gln Asp Pro Gly Met Gly Leu Arg Leu Ile His Tyr Ser
35 40 45
Val Gly Ala Gly Ile Thr Asp Gln Gly Glu Val Pro Asn Gly Tyr Asn
50 55 60
Val Ser Arg Ser Thr Thr Glu Asp Phe Pro Leu Arg Leu Leu Ser Ala
65 70 75 80

Ala Pro Ser Gln Thr Ser Val Tyr Phe Cys Ala Ser Ser Tyr Val Gly

85 90 95

Asn Thr Gly Glu Leu Phe Phe Gly Glu Gly Ser Arg Leu Thr Val Leu

100 105 110

<210> 700

<211> 166

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 700

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu

1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His

20 25 30

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

35 40 45

Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp

50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu

65 70 75 80

Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp

85 90 95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu

100 105 110

Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala

115 120 125

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

130 135 140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala

145 150 155 160

Cys Arg Thr Gly Asp Arg

165

<210> 701

<211> 452

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 701

Met Ser Thr Ala Val Leu Glu Asn Pro Gly Leu Gly Arg Lys Leu Ser

1 5 10 15

Asp Phe Gly Gln Glu Thr Ser Tyr Ile Glu Asp Asn Cys Asn Gln Asn

20 25 30

Gly Ala Ile Ser Leu Ile Phe Ser Leu Lys Glu Glu Val Gly Ala Leu

35 40 45

Ala Lys Val Leu Arg Leu Phe Glu Glu Asn Asp Val Asn Leu Thr His

50 55 60

Ile Glu Ser Arg Pro Ser Arg Leu Lys Lys Asp Glu Tyr Glu Phe Phe

65 70 75 80

Thr His Leu Asp Lys Arg Ser Leu Pro Ala Leu Thr Asn Ile Ile Lys

85 90 95

Ile Leu Arg His Asp Ile Gly Ala Thr Val His Glu Leu Ser Arg Asp

100 105 110

Lys Lys Lys Asp Thr Val Pro Trp Phe Pro Arg Thr Ile Gln Glu Leu

115 120 125

Asp Arg Phe Ala Asn Gln Ile Leu Ser Tyr Gly Ala Glu Leu Asp Ala

130 135 140

Asp His Pro Gly Phe Lys Asp Pro Val Tyr Arg Ala Arg Arg Lys Gln

145 150 155 160

Phe Ala Asp Ile Ala Tyr Asn Tyr Arg His Gly Gln Pro Ile Pro Arg

165 170 175

Val Glu Tyr Met Glu Glu Glu Lys Lys Thr Trp Gly Thr Val Phe Lys

180 185 190

Thr Leu Lys Ser Leu Tyr Lys Thr His Ala Cys Tyr Glu Tyr Asn His

195 200 205

Ile Phe Pro Leu Leu Glu Lys Tyr Cys Gly Phe His Glu Asp Asn Ile

210 215 220

Pro Gln Leu Glu Asp Val Ser Gln Phe Leu Gln Thr Cys Thr Gly Phe

225 230 235 240

Arg Leu Arg Pro Val Ala Gly Leu Leu Ser Ser Arg Asp Phe Leu Gly

245 250 255

Gly Leu Ala Phe Arg Val Phe His Cys Thr Gln Tyr Ile Arg His Gly

260 265 270

Ser Lys Pro Met Tyr Thr Pro Glu Pro Asp Ile Cys His Glu Leu Leu

275 280 285

Gly His Val Pro Leu Phe Ser Asp Arg Ser Phe Ala Gln Phe Ser Gln

290 295 300

Glu Ile Gly Leu Ala Ser Leu Gly Ala Pro Asp Glu Tyr Ile Glu Lys

305 310 315 320

Leu Ala Thr Ile Tyr Trp Phe Thr Val Glu Phe Gly Leu Cys Lys Gln

325 330 335

Gly Asp Ser Ile Lys Ala Tyr Gly Ala Gly Leu Leu Ser Ser Phe Gly

340 345 350

Glu Leu Gln Tyr Cys Leu Ser Glu Lys Pro Lys Leu Leu Pro Leu Glu

355 360 365

Leu Glu Lys Thr Ala Ile Gln Asn Tyr Thr Val Thr Glu Phe Gln Pro

370 375 380

Leu Tyr Tyr Val Ala Glu Ser Phe Asn Asp Ala Lys Glu Lys Val Arg

385 390 395 400

Asn Phe Ala Ala Thr Ile Pro Arg Pro Phe Ser Val Arg Tyr Asp Pro

405 410 415

Tyr Thr Gln Arg Ile Glu Val Leu Asp Asn Thr Gln Gln Leu Lys Ile
420 425 430
Leu Ala Asp Ser Ile Asn Ser Glu Ile Gly Ile Leu Cys Ser Ala Leu
435 440 445
Gln Lys Ile Lys
450

<210> 702
<211> 1462
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 702

Leu Ser Val Lys Ala Gln Thr Ala His Ile Val Leu Glu Asp Gly Thr
1 5 10 15
Lys Met Lys Gly Tyr Ser Phe Gly His Pro Ser Ser Val Ala Gly Glu
20 25 30
Val Val Phe Asn Thr Gly Leu Gly Gly Tyr Pro Glu Ala Ile Thr Asp
35 40 45
Pro Ala Tyr Lys Gly Gln Ile Leu Thr Met Ala Asn Pro Ile Ile Gly
50 55 60
Asn Gly Gly Ala Pro Asp Thr Thr Ala Leu Asp Glu Leu Gly Leu Ser
65 70 75 80
Lys Tyr Leu Glu Ser Asn Gly Ile Lys Val Ser Gly Leu Leu Val Leu
85 90 95
Asp Tyr Ser Lys Asp Tyr Asn His Trp Leu Ala Thr Lys Ser Leu Gly
100 105 110
Gln Trp Leu Gln Glu Glu Lys Val Pro Ala Ile Tyr Gly Val Asp Thr
115 120 125
Arg Met Leu Thr Lys Ile Ile Arg Asp Lys Gly Thr Met Leu Gly Lys
130 135 140

Ile Glu Phe Glu Gly Gln Pro Val Asp Phe Val Asp Pro Asn Lys Gln
145 150 155 160
Asn Leu Ile Ala Glu Val Ser Thr Lys Asp Val Lys Val Tyr Gly Lys
165 170 175
Gly Asn Pro Thr Lys Val Val Ala Val Asp Cys Gly Ile Lys Asn Asn
180 185 190
Val Ile Arg Leu Leu Val Lys Arg Gly Ala Glu Val His Leu Val Pro

195 200 205
Trp Asn His Asp Phe Thr Lys Met Glu Tyr Asp Gly Ile Leu Ile Ala
210 215 220
Gly Gly Pro Gly Asn Pro Ala Leu Ala Glu Pro Leu Ile Gln Asn Val
225 230 235 240
Arg Lys Ile Leu Glu Ser Asp Arg Lys Glu Pro Leu Phe Gly Ile Ser
245 250 255
Thr Gly Asn Leu Ile Thr Gly Leu Ala Ala Gly Ala Lys Thr Tyr Lys

260 265 270
Met Ser Met Ala Asn Arg Gly Gln Asn Gln Pro Val Leu Asn Ile Thr
275 280 285
Asn Lys Gln Ala Phe Ile Thr Ala Gln Asn His Gly Tyr Ala Leu Asp
290 295 300
Asn Thr Leu Pro Ala Gly Trp Lys Pro Leu Phe Val Asn Val Asn Asp
305 310 315 320
Gln Thr Asn Glu Gly Ile Met His Glu Ser Lys Pro Phe Phe Ala Val

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

385 390 395 400

Gly Glu Phe Asp Tyr Ser Gly Ser Gln Ala Val Lys Ala Met Lys Glu

405 410 415

Glu Asn Val Lys Thr Val Leu Met Asn Pro Asn Ile Ala Ser Val Gln

420 425 430

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

435 440 445

Ile Thr Pro Gln Phe Val Thr Glu Val Ile Lys Ala Glu Gln Pro Asp

450 455 460

Gly Leu Ile Leu Gly Met Gly Gly Gln Thr Ala Leu Asn Cys Gly Val

465 470 475 480

Glu Leu Phe Lys Arg Gly Val Leu Lys Glu Tyr Gly Val Lys Val Leu

485 490 495

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

500 505 510

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

515 520 525

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

530 535 540

Pro Val Met Ile Arg Ser Ala Tyr Ala Leu Gly Gly Leu Gly Ser Gly

545 550 555 560

Ile Cys Pro Asn Arg Glu Thr Leu Met Asp Leu Ser Thr Lys Ala Phe

565 570 575

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

580 585 590

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

595 600 605

Val Cys Asn Met Glu Asn Val Asp Ala Met Gly Val His Thr Gly Asp

610 615 620

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

625 630 635 640

Met Leu Arg Arg Thr Ser Ile Asn Val Val Arg His Leu Gly Ile Val

645 650 655

Gly Glu Cys Asn Ile Gln Phe Ala Leu His Pro Thr Ser Met Glu Tyr

660 665 670

Cys Ile Ile Glu Val Asn Ala Arg Leu Ser Arg Ser Ser Ala Leu Ala

675 680 685

Ser Lys Ala Thr Gly Tyr Pro Leu Ala Phe Ile Ala Ala Lys Ile Ala

690 695 700

Leu Gly Ile Pro Leu Pro Glu Ile Lys Asn Val Val Ser Gly Lys Thr

705 710 715 720

Ser Ala Cys Phe Glu Pro Ser Leu Asp Tyr Met Val Thr Lys Ile Pro

725 730 735

Arg Trp Asp Leu Asp Arg Phe His Gly Thr Ser Ser Arg Ile Gly Ser

740 745 750

Ser Met Lys Ser Val Gly Glu Val Met Ala Ile Gly Arg Thr Phe Glu

755 760 765

Glu Ser Phe Gln Lys Ala Leu Arg Met Cys His Pro Ser Ile Glu Gly

770 775 780

Phe Thr Pro Arg Leu Pro Met Asn Lys Glu Trp Pro Ser Asn Leu Asp

785 790 795 800

Leu Arg Lys Glu Leu Ser Glu Pro Ser Ser Thr Arg Ile Tyr Ala Ile

805 810 815

Ala Lys Ala Ile Asp Asp Asn Met Ser Leu Asp Glu Ile Glu Lys Leu

820 825 830

Thr Tyr Ile Asp Lys Trp Phe Leu Tyr Lys Met Arg Asp Ile Leu Asn

835 840 845

Met Glu Lys Thr Leu Lys Gly Leu Asn Ser Glu Ser Met Thr Glu Glu

850 855 860

Thr Leu Lys Arg Ala Lys Glu Ile Gly Phe Ser Asp Lys Gln Ile Ser

865 870 875 880

Lys Cys Leu Gly Leu Thr Glu Ala Gln Thr Arg Glu Leu Arg Leu Lys

885 890 895

Lys Asn Ile His Pro Trp Val Lys Gln Ile Asp Thr Leu Ala Ala Glu

900 905 910

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

915 920 925

His Asp Val Asn Phe Asp Asp His Gly Met Met Val Leu Gly Cys Gly

930 935 940

Pro Tyr His Ile Gly Ser Ser Val Glu Phe Asp Trp Cys Ala Val Ser

945 950 955 960

Ser Ile Arg Thr Leu Arg Gln Leu Gly Lys Lys Thr Val Val Val Asn

965 970 975

Cys Asn Pro Glu Thr Val Ser Thr Asp Phe Asp Glu Cys Asp Lys Leu

980 985 990

Tyr Phe Glu Glu Leu Ser Leu Glu Arg Ile Leu Asp Ile Tyr His Gln

995 1000 1005

Glu Ala Cys Gly Gly Cys Ile Ile Ser Val Gly Gly Gln Ile Pro

1010 1015 1020

Asn Asn Leu Ala Val Pro Leu Tyr Lys Asn Gly Val Lys Ile Met

1025 1030 1035

Gly Thr Ser Pro Leu Gln Ile Asp Arg Ala Glu Asp Arg Ser Ile

1040 1045 1050

Phe Ser Ala Val Leu Asp Glu Leu Lys Val Ala Gln Ala Pro Trp

1055 1060 1065

Lys Ala Val Asn Thr Leu Asn Glu Ala Leu Glu Phe Ala Lys Ser

1070 1075 1080

Val Asp Tyr Pro Cys Leu Leu Arg Pro Ser Tyr Val Leu Ser Gly

1085 1090 1095

Ser Ala Met Asn Val Val Phe Ser Glu Asp Glu Met Lys Lys Phe

1100 1105 1110

Leu Glu Glu Ala Thr Arg Val Ser Gln Glu His Pro Val Val Leu

1115 1120 1125

Thr Lys Phe Val Glu Gly Ala Arg Glu Val Glu Met Asp Ala Val

1130 1135 1140

Gly Lys Asp Gly Arg Val Ile Ser His Ala Ile Ser Glu His Val

1145 1150 1155

Glu Asp Ala Gly Val His Ser Gly Asp Ala Thr Leu Met Leu Pro

1160 1165 1170

Thr Gln Thr Ile Ser Gln Gly Ala Ile Glu Lys Val Lys Asp Ala

1175 1180 1185

Thr Arg Lys Ile Ala Lys Ala Phe Ala Ile Ser Gly Pro Phe Asn

1190 1195 1200

Val Gln Phe Leu Val Lys Gly Asn Asp Val Leu Val Ile Glu Cys

1205 1210 1215

Asn Leu Arg Ala Ser Arg Ser Phe Pro Phe Val Ser Lys Thr Leu

1220 1225 1230

Gly Val Asp Phe Ile Asp Val Ala Thr Lys Val Met Ile Gly Glu

1235 1240 1245

Asn Val Asp Glu Lys His Leu Pro Thr Leu Asp His Pro Ile Ile

1250 1255 1260

Pro Ala Asp Tyr Val Ala Ile Lys Ala Pro Met Phe Ser Trp Pro

1265 1270 1275

Arg Leu Arg Asp Ala Asp Pro Ile Leu Arg Cys Glu Met Ala Ser

1280 1285 1290

Thr Gly Glu Val Ala Cys Phe Gly Glu Gly Ile His Thr Ala Phe

1295 1300 1305

Leu Lys Ala Met Leu Ser Thr Gly Phe Lys Ile Pro Gln Lys Gly

1310 1315 1320

Ile Leu Ile Gly Ile Gln Gln Ser Phe Arg Pro Arg Phe Leu Gly

1325 1330 1335

Val Ala Glu Gln Leu His Asn Glu Gly Phe Lys Leu Phe Ala Thr

1340 1345 1350

Glu Ala Thr Ser Asp Trp Leu Asn Ala Asn Asn Val Pro Ala Thr

1355 1360 1365

Pro Val Ala Trp Pro Ser Gln Glu Gly Gln Asn Pro Ser Leu Ser

1370 1375 1380

Ser Ile Arg Lys Leu Ile Arg Asp Gly Ser Ile Asp Leu Val Ile

1385 1390 1395

Asn Leu Pro Asn Asn Asn Thr Lys Phe Val His Asp Asn Tyr Val

1400 1405 1410

Ile Arg Arg Thr Ala Val Asp Ser Gly Ile Pro Leu Leu Thr Asn

1415 1420 1425

Phe Gln Val Thr Lys Leu Phe Ala Glu Ala Val Gln Lys Ser Arg

1430 1435 1440

Lys Val Asp Ser Lys Ser Leu Phe His Tyr Arg Gln Tyr Ser Ala

1445 1450 1455

Gly Lys Ala Ala

1460

<210> 703

<211> 1053

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 703

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

1 5 10 15

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly

20 25 30

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg

35 40 45

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile

50 55 60

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His

65 70 75 80

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu

85 90 95

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu

100 105 110

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

115 120 125

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala

130 135 140

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys

145 150 155 160

Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr

165 170 175

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln

180 185 190

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

195 200 205

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys

210 215 220

Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe

225 230 235 240

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr

245 250 255

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn

260 265 270

Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe

275 280 285

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu

290 295 300

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

305 310 315 320

Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
325 330 335
Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
340 345 350

Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
355 360 365
Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
370 375 380
Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
385 390 395 400
Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
405 410 415

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
420 425 430
Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
435 440 445
Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
450 455 460
Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
465 470 475 480

Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
485 490 495
Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
500 505 510
Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
515 520 525
Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
530 535 540

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
545 550 555 560
Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys

565 570 575

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

580 585 590

Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile

595 600 605

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

610 615 620

Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp

625 630 635 640

Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu

645 650 655

Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys

660 665 670

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp

675 680 685

Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp

690 695 700

Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys

705 710 715 720

Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys

725 730 735

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu

740 745 750

Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp

755 760 765

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile

770 775 780

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu

785 790 795 800

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu

805 810 815

Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
820 825 830

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
835 840 845

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
850 855 860

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
865 870 875 880

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
885 890 895

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
900 905 910

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
915 920 925

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
930 935 940

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
945 950 955 960

Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
965 970 975

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
980 985 990

Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met
995 1000 1005

Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys
1010 1015 1020

Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
1025 1030 1035

Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
1040 1045 1050

<210> 704

<211> 1353

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 704

Ala Ala Gly Gly Ile Leu His Leu Glu Leu Leu Val Ala Val Gly Pro
1 5 10 15

Asp Val Phe Gln Ala His Gln Glu Asp Thr Glu Arg Tyr Val Leu Thr
20 25 30

Asn Leu Asn Ile Gly Ala Glu Leu Leu Arg Asp Pro Ser Leu Gly Ala
35 40 45

Gln Phe Arg Val His Leu Val Lys Met Val Ile Leu Thr Glu Pro Glu
50 55 60

Gly Ala Pro Asn Ile Thr Ala Asn Leu Thr Ser Ser Leu Leu Ser Val
65 70 75 80

Cys Gly Trp Ser Gln Thr Ile Asn Pro Glu Asp Asp Thr Asp Pro Gly
85 90 95

His Ala Asp Leu Val Leu Tyr Ile Thr Arg Phe Asp Leu Glu Leu Pro
100 105 110

Asp Gly Asn Arg Gln Val Arg Gly Val Thr Gln Leu Gly Gly Ala Cys
115 120 125

Ser Pro Thr Trp Ser Cys Leu Ile Thr Glu Asp Thr Gly Phe Asp Leu
130 135 140

Gly Val Thr Ile Ala His Glu Ile Gly His Ser Phe Gly Leu Glu His
145 150 155 160

Asp Gly Ala Pro Gly Ser Gly Cys Gly Pro Ser Gly His Val Met Ala
165 170 175

Ser Asp Gly Ala Ala Pro Arg Ala Gly Leu Ala Trp Ser Pro Cys Ser
180 185 190

Arg Arg Gln Leu Leu Ser Leu Leu Ser Ala Gly Arg Ala Arg Cys Val

195 200 205

Trp Asp Pro Pro Arg Pro Gln Pro Gly Ser Ala Gly His Pro Pro Asp

210 215 220

Ala Gln Pro Gly Leu Tyr Tyr Ser Ala Asn Glu Gln Cys Arg Val Ala

225 230 235 240

Phe Gly Pro Lys Ala Val Ala Cys Thr Phe Ala Arg Glu His Leu Asp

245 250 255

Met Cys Gln Ala Leu Ser Cys His Thr Asp Pro Leu Asp Gln Ser Ser

260 265 270

Cys Ser Arg Leu Leu Val Pro Leu Leu Asp Gly Thr Glu Cys Gly Val

275 280 285

Glu Lys Trp Cys Ser Lys Gly Arg Cys Arg Ser Leu Val Glu Leu Thr

290 295 300

Pro Ile Ala Ala Val His Gly Arg Trp Ser Ser Trp Gly Pro Arg Ser

305 310 315 320

Pro Cys Ser Arg Ser Cys Gly Gly Gly Val Val Thr Arg Arg Arg Gln

325 330 335

Cys Asn Asn Pro Arg Pro Ala Phe Gly Gly Arg Ala Cys Val Gly Ala

340 345 350

Asp Leu Gln Ala Glu Met Cys Asn Thr Gln Ala Cys Glu Lys Thr Gln

355 360 365

Leu Glu Phe Met Ser Gln Gln Cys Ala Arg Thr Asp Gly Gln Pro Leu

370 375 380

Arg Ser Ser Pro Gly Gly Ala Ser Phe Tyr His Trp Gly Ala Ala Val

385 390 395 400

Pro His Ser Gln Gly Asp Ala Leu Cys Arg His Met Cys Arg Ala Ile

405 410 415

Gly Glu Ser Phe Ile Met Lys Arg Gly Asp Ser Phe Leu Asp Gly Thr

420 425 430

Arg Cys Met Pro Ser Gly Pro Arg Glu Asp Gly Thr Leu Ser Leu Cys

435 440 445

Val Ser Gly Ser Cys Arg Thr Phe Gly Cys Asp Gly Arg Met Asp Ser
450 455 460
Gln Gln Val Trp Asp Arg Cys Gln Val Cys Gly Gly Asp Asn Ser Thr
465 470 475 480
Cys Ser Pro Arg Lys Gly Ser Phe Thr Ala Gly Arg Ala Arg Glu Tyr
485 490 495

Val Thr Phe Leu Thr Val Thr Pro Asn Leu Thr Ser Val Tyr Ile Ala
500 505 510
Asn His Arg Pro Leu Phe Thr His Leu Ala Val Arg Ile Gly Gly Arg
515 520 525
Tyr Val Val Ala Gly Lys Met Ser Ile Ser Pro Asn Thr Thr Tyr Pro
530 535 540
Ser Leu Leu Glu Asp Gly Arg Val Glu Tyr Arg Val Ala Leu Thr Glu
545 550 555 560

Asp Arg Leu Pro Arg Leu Glu Glu Ile Arg Ile Trp Gly Pro Leu Gln
565 570 575
Glu Asp Ala Asp Ile Gln Val Tyr Arg Arg Tyr Gly Glu Glu Tyr Gly
580 585 590
Asn Leu Thr Arg Pro Asp Ile Thr Phe Thr Tyr Phe Gln Pro Lys Pro
595 600 605
Arg Gln Ala Trp Val Trp Ala Ala Val Arg Gly Pro Cys Ser Val Ser
610 615 620

Cys Gly Ala Gly Leu Arg Trp Val Asn Tyr Ser Cys Leu Asp Gln Ala
625 630 635 640
Arg Lys Glu Leu Val Glu Thr Val Gln Cys Gln Gly Ser Gln Gln Pro
645 650 655
Pro Ala Trp Pro Glu Ala Cys Val Leu Glu Pro Cys Pro Pro Tyr Trp
660 665 670
Ala Val Gly Asp Phe Gly Pro Cys Ser Ala Ser Cys Gly Gly Gly Leu
675 680 685

Arg Glu Arg Pro Val Arg Cys Val Glu Ala Gln Gly Ser Leu Leu Lys

690 695 700

Thr Leu Pro Pro Ala Arg Cys Arg Ala Gly Ala Gln Gln Pro Ala Val
705 710 715 720

Ala Leu Glu Thr Cys Asn Pro Gln Pro Cys Pro Ala Arg Trp Glu Val
725 730 735

Ser Glu Pro Ser Ser Cys Thr Ser Ala Gly Gly Ala Gly Leu Ala Leu
740 745 750

Glu Asn Glu Thr Cys Val Pro Gly Ala Asp Gly Leu Glu Ala Pro Val
755 760 765

Thr Glu Gly Pro Gly Ser Val Asp Glu Lys Leu Pro Ala Pro Glu Pro
770 775 780

Cys Val Gly Met Ser Cys Pro Pro Gly Trp Gly His Leu Asp Ala Thr
785 790 795 800

Ser Ala Gly Glu Lys Ala Pro Ser Pro Trp Gly Ser Ile Arg Thr Gly
805 810 815

Ala Gln Ala Ala His Val Trp Thr Pro Ala Ala Gly Ser Cys Ser Val
820 825 830

Ser Cys Gly Arg Gly Leu Met Glu Leu Arg Phe Leu Cys Met Asp Ser
835 840 845

Ala Leu Arg Val Pro Val Gln Glu Glu Leu Cys Gly Leu Ala Ser Lys
850 855 860

Pro Gly Ser Arg Arg Glu Val Cys Gln Ala Val Pro Cys Pro Ala Arg
865 870 875 880

Trp Gln Tyr Lys Leu Ala Ala Cys Ser Val Ser Cys Gly Arg Gly Val
885 890 895

Val Arg Arg Ile Leu Tyr Cys Ala Arg Ala His Gly Glu Asp Asp Gly
900 905 910

Glu Glu Ile Leu Leu Asp Thr Gln Cys Gln Gly Leu Pro Arg Pro Glu
915 920 925

Pro Gln Glu Ala Cys Ser Leu Glu Pro Cys Pro Pro Arg Trp Lys Val
930 935 940

Met Ser Leu Gly Pro Cys Ser Ala Ser Cys Gly Leu Gly Thr Ala Arg
945 950 955 960
Arg Ser Val Ala Cys Val Gln Leu Asp Gln Gly Gln Asp Val Glu Val
965 970 975
Asp Glu Ala Ala Cys Ala Ala Leu Val Arg Pro Glu Ala Ser Val Pro
980 985 990
Cys Leu Ile Ala Asp Cys Thr Tyr Arg Trp His Val Gly Thr Trp Met
995 1000 1005

Glu Cys Ser Val Ser Cys Gly Asp Gly Ile Gln Arg Arg Arg Asp
1010 1015 1020
Thr Cys Leu Gly Pro Gln Ala Gln Ala Pro Val Pro Ala Asp Phe
1025 1030 1035
Cys Gln His Leu Pro Lys Pro Val Thr Val Arg Gly Cys Trp Ala
1040 1045 1050
Gly Pro Cys Val Gly Gln Gly Thr Pro Ser Leu Val Pro His Glu
1055 1060 1065

Glu Ala Ala Ala Pro Gly Arg Thr Thr Ala Thr Pro Ala Gly Ala
1070 1075 1080
Ser Leu Glu Trp Ser Gln Ala Arg Gly Leu Leu Phe Ser Pro Ala
1085 1090 1095
Pro Gln Pro Arg Arg Leu Leu Pro Gly Pro Gln Glu Asn Ser Val
1100 1105 1110
Gln Ser Ser Ala Cys Gly Arg Gln His Leu Glu Pro Thr Gly Thr
1115 1120 1125

Ile Asp Met Arg Gly Pro Gly Gln Ala Asp Cys Ala Val Ala Ile
1130 1135 1140
Gly Arg Pro Leu Gly Glu Val Val Thr Leu Arg Val Leu Glu Ser
1145 1150 1155
Ser Leu Asn Cys Ser Ala Gly Asp Met Leu Leu Leu Trp Gly Arg
1160 1165 1170
Leu Thr Trp Arg Lys Met Cys Arg Lys Leu Leu Asp Met Thr Phe

1175 1180 1185

Ser Ser Lys Thr Asn Thr Leu Val Val Arg Gln Arg Cys Gly Arg
1190 1195 1200
Pro Gly Gly Gly Val Leu Leu Arg Tyr Gly Ser Gln Leu Ala Pro
1205 1210 1215
Glu Thr Phe Tyr Arg Glu Cys Asp Met Gln Leu Phe Gly Pro Trp
1220 1225 1230
Gly Glu Ile Val Ser Pro Ser Leu Ser Pro Ala Thr Ser Asn Ala
1235 1240 1245

Gly Gly Cys Arg Leu Phe Ile Asn Val Ala Pro His Ala Arg Ile
1250 1255 1260
Ala Ile His Ala Leu Ala Thr Asn Met Gly Ala Gly Thr Glu Gly
1265 1270 1275
Ala Asn Ala Ser Tyr Ile Leu Ile Arg Asp Thr His Ser Leu Arg
1280 1285 1290
Thr Thr Ala Phe His Gly Gln Gln Val Leu Tyr Trp Glu Ser Glu
1295 1300 1305

Ser Ser Gln Ala Glu Met Glu Phe Ser Glu Gly Phe Leu Lys Ala
1310 1315 1320
Gln Ala Ser Leu Arg Gly Gln Tyr Trp Thr Leu Gln Ser Trp Val
1325 1330 1335
Pro Glu Met Gln Asp Pro Gln Ser Trp Lys Gly Lys Glu Gly Thr
1340 1345 1350
<210> 705
<211> 431
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"
<400> 705

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

Pro Pro Ser Gln Leu Gln Leu Pro Thr Leu Pro Leu Val Met Val Ala
65                  70                  75                  80
Pro Ser Gly Ala Arg Leu Gly Pro Leu Pro His Leu Gln Ala Leu Leu
                85                  90                  95
Gln Asp Arg Pro His Phe Met His Gln Leu Ser Thr Val Asp Ala His
            100                 105                 110
Ala Arg Thr Pro Val Leu Gln Val His Pro Leu Glu Ser Pro Ala Met
        115                 120                 125

Ile Ser Leu Thr Pro Pro Thr Thr Ala Thr Gly Val Phe Ser Leu Lys
    130                 135                 140
Ala Arg Pro Gly Leu Pro Pro Gly Ile Asn Val Ala Ser Leu Glu Trp
145                 150                 155                 160
Val Ser Arg Glu Pro Ala Leu Leu Cys Thr Phe Pro Asn Pro Ser Ala
                165                 170                 175
Pro Arg Lys Asp Ser Thr Leu Ser Ala Val Pro Gln Ser Ser Tyr Pro
            180                 185                 190

Leu Leu Ala Asn Gly Val Cys Lys Trp Pro Gly Cys Glu Lys Val Phe
        195                 200                 205
Glu Glu Pro Glu Asp Phe Leu Lys His Cys Gln Ala Asp His Leu Leu
    210                 215                 220
Asp Glu Lys Gly Arg Ala Gln Cys Leu Leu Gln Arg Glu Met Val Gln
225                 230                 235                 240
Ser Leu Glu Gln Gln Leu Val Leu Glu Lys Glu Lys Leu Ser Ala Met
```

245 250 255

Gln Ala His Leu Ala Gly Lys Met Ala Leu Thr Lys Ala Ser Ser Val
260 265 270
Ala Ser Ser Asp Lys Gly Ser Cys Cys Ile Val Ala Ala Gly Ser Gln
275 280 285
Gly Pro Val Val Pro Ala Trp Ser Gly Pro Arg Glu Ala Pro Asp Ser
290 295 300
Leu Phe Ala Val Arg Arg His Leu Trp Gly Ser His Gly Asn Ser Thr
305 310 315 320

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

Ile Arg His Asn Leu Ser Leu His Lys Cys Phe Val Arg Val Glu Ser
385 390 395 400
Glu Lys Gly Ala Val Trp Thr Val Asp Glu Leu Glu Phe Arg Lys Lys
405 410 415
Arg Ser Gln Arg Pro Ser Arg Cys Ser Asn Pro Thr Pro Gly Pro
420 425 430

<210> 706
<211> 160
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"
<400> 706

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

Leu Ser Glu Met Ile Gln Phe Tyr Leu Glu Glu Val Met Pro Gln Ala
65 70 75 80
Glu Asn Gln Asp Pro Asp Ile Lys Ala His Val Asn Ser Leu Gly Glu
85 90 95
Asn Leu Lys Thr Leu Arg Leu Arg Leu Arg Arg Cys His Arg Phe Leu
100 105 110
Pro Cys Glu Asn Lys Ser Lys Ala Val Glu Gln Val Lys Asn Ala Phe
115 120 125

Asn Lys Leu Gln Glu Lys Gly Ile Tyr Lys Ala Met Ser Glu Phe Asp
130 135 140
Ile Phe Ile Asn Tyr Ile Glu Ala Tyr Met Thr Met Lys Ile Arg Asn
145 150 155 160
<210> 707
<211> 133
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
polypeptide"
<400> 707
Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu Gln Leu Glu His
1 5 10 15

Leu Leu Leu Asp Leu Gln Met Ile Leu Asn Gly Ile Asn Asn Tyr Lys
20 25 30

Asn Pro Lys Leu Thr Arg Met Leu Thr Phe Lys Phe Tyr Met Pro Lys
35 40 45
Lys Ala Thr Glu Leu Lys His Leu Gln Cys Leu Glu Glu Glu Leu Lys
50 55 60
Pro Leu Glu Glu Val Leu Asn Leu Ala Gln Ser Lys Asn Phe His Leu
65 70 75 80

Arg Pro Arg Asp Leu Ile Ser Asn Ile Asn Val Ile Val Leu Glu Leu
85 90 95
Lys Gly Ser Glu Thr Thr Phe Met Cys Glu Tyr Ala Asp Glu Thr Ala
100 105 110
Thr Ile Val Glu Phe Leu Asn Arg Trp Ile Thr Phe Cys Gln Ser Ile
115 120 125
Ile Ser Thr Leu Thr
130

<210> 708
<400> 708
000
<210> 709
<400> 709
000
<210> 710

<400> 710
000
<210> 711
<400> 711
000
<210> 712
<400> 712
000
<210> 713
<400> 713
000
<210> 714

<400> 714

000

<210> 715

<400> 715

000

<210> 716

<400> 716

000

<210> 717

<400> 717

000

<210> 718

<211> 96

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL1 sequence"

<400> 718

Met Gln Ile Ile Thr Thr Ala Leu Val Cys Leu Leu Leu Ala Gly Met

1 5 10 15

Trp Pro Glu Asp Val Asp Ser Lys Ser Met Gln Val Pro Phe Ser Arg

20 25 30

Cys Cys Phe Ser Phe Ala Glu Gln Glu Ile Pro Leu Arg Ala Ile Leu

35 40 45

Cys Tyr Arg Asn Thr Ser Ser Ile Cys Ser Asn Glu Gly Leu Ile Phe

50 55 60

Lys Leu Lys Arg Gly Lys Glu Ala Cys Ala Leu Asp Thr Val Gly Trp

65 70 75 80

Val Gln Arg His Arg Lys Met Leu Arg His Cys Pro Ser Lys Arg Lys

85 90 95

<210> 719

<211> 50

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL2 sequence"

<400> 719

Met Lys Val Ser Ala Ala Leu Leu Cys Leu Leu Leu Ile Ala Ala Thr

1 5 10 15

Phe Ile Pro Gln Gly Leu Ala Gln Pro Asp Ala Ile Asn Ala Pro Val

20 25 30

Thr Cys Cys Tyr Asn Phe Thr Asn Arg Lys Ile Ser Val Gln Arg Leu

35 40 45

Ala Ser

50

<210> 720

<211> 92

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL3 sequence"

<400> 720

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala

1 5 10 15

Leu Cys Asn Gln Phe Ser Ala Ser Leu Ala Ala Asp Thr Pro Thr Ala

20 25 30

Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile Ala

35 40 45

Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Gly Val Ile Phe

50 55 60

Leu Thr Lys Arg Ser Arg Gln Val Cys Ala Asp Pro Ser Glu Glu Trp

65 70 75 80

Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala

85 90

<210> 721

<211> 92

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL4 sequence"

<400> 721

Met Lys Leu Cys Val Thr Val Leu Ser Leu Leu Met Leu Val Ala Ala
1 5 10 15

Phe Cys Ser Pro Ala Leu Ser Ala Pro Met Gly Ser Asp Pro Pro Thr
20 25 30

Ala Cys Cys Phe Ser Tyr Thr Ala Arg Lys Leu Pro Arg Asn Phe Val
35 40 45

Val Asp Tyr Tyr Glu Thr Ser Ser Leu Cys Ser Gln Pro Ala Val Val
50 55 60

Phe Gln Thr Lys Arg Ser Lys Gln Val Cys Ala Asp Pro Ser Glu Ser
65 70 75 80

Trp Val Gln Glu Tyr Val Tyr Asp Leu Glu Leu Asn
85 90

<210> 722

<211> 91

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL5 sequence"

<400> 722

Met Lys Val Ser Ala Ala Ala Leu Ala Val Ile Leu Ile Ala Thr Ala
1 5 10 15

Leu Cys Ala Pro Ala Ser Ala Ser Pro Tyr Ser Ser Asp Thr Thr Pro

20 25 30

Cys Cys Phe Ala Tyr Ile Ala Arg Pro Leu Pro Arg Ala His Ile Lys

35 40 45

Glu Tyr Phe Tyr Thr Ser Gly Lys Cys Ser Asn Pro Ala Val Val Phe

50 55 60

Val Thr Arg Lys Asn Arg Gln Val Cys Ala Asn Pro Glu Lys Lys Trp

65 70 75 80

Val Arg Glu Tyr Ile Asn Ser Leu Glu Met Ser

85 90

<210> 723

<211> 116

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL6 sequence"

<400> 723

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

1 5 10 15

Leu Gly Ser Gln Ala Gly Leu Ile Gln Glu Met Glu Lys Glu Asp Arg

20 25 30

Arg Tyr Asn Pro Pro Ile Ile His Gln Gly Phe Gln Asp Thr Ser Ser

35 40 45

Asp Cys Cys Phe Ser Tyr Ala Thr Gln Ile Pro Cys Lys Arg Phe Ile

50 55 60

Tyr Tyr Phe Pro Thr Ser Gly Gly Cys Ile Lys Pro Gly Ile Ile Phe

65 70 75 80

Ile Ser Arg Arg Gly Thr Gln Val Cys Ala Asp Pro Ser Asp Arg Arg

85 90 95

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

100 105 110

Lys Val Ile Ala

115

<210> 724

<211> 99

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CCL7 sequence"

<400> 724

Met Lys Ala Ser Ala Ala Leu Leu Cys Leu Leu Leu Thr Ala Ala Ala
1 5 10 15

Phe Ser Pro Gln Gly Leu Ala Gln Pro Val Gly Ile Asn Thr Ser Thr
20 25 30

Thr Cys Cys Tyr Arg Phe Ile Asn Lys Lys Ile Pro Lys Gln Arg Leu
35 40 45

Glu Ser Tyr Arg Arg Thr Thr Ser Ser His Cys Pro Arg Glu Ala Val
50 55 60

Ile Phe Lys Thr Lys Leu Asp Lys Glu Ile Cys Ala Asp Pro Thr Gln
65 70 75 80

Lys Trp Val Gln Asp Phe Met Lys His Leu Asp Lys Lys Thr Gln Thr
85 90 95

Pro Lys Leu

<210> 725

<211> 99

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CCL8 sequence"

<400> 725

Met Lys Val Ser Ala Ala Leu Leu Cys Leu Leu Leu Met Ala Ala Thr
1 5 10 15

Phe Ser Pro Gln Gly Leu Ala Gln Pro Asp Ser Val Ser Ile Pro Ile
20 25 30
Thr Cys Cys Phe Asn Val Ile Asn Arg Lys Ile Pro Ile Gln Arg Leu
35 40 45

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

<210> 726
<211> 122
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:

CCL9/CCL10 sequence"
<400> 726
Met Lys Pro Phe His Thr Ala Leu Ser Phe Leu Ile Leu Thr Thr Ala
1 5 10 15
Leu Gly Ile Trp Ala Gln Ile Thr His Ala Thr Glu Thr Lys Glu Val
20 25 30
Gln Ser Ser Leu Lys Ala Gln Gln Gly Leu Glu Ile Glu Met Phe His
35 40 45
Met Gly Phe Gln Asp Ser Ser Asp Cys Cys Leu Ser Tyr Asn Ser Arg
50 55 60

Ile Gln Cys Ser Arg Phe Ile Gly Tyr Phe Pro Thr Ser Gly Gly Cys
65 70 75 80
Thr Arg Pro Gly Ile Ile Phe Ile Ser Lys Arg Gly Phe Gln Val Cys
85 90 95

Ala Asn Pro Ser Asp Arg Arg Val Gln Arg Cys Ile Glu Arg Leu Glu

100 105 110

Gln Asn Ser Gln Pro Arg Thr Tyr Lys Gln

115 120

<210> 727

<211> 97

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL11 sequence"

<400> 727

Met Lys Val Ser Ala Ala Leu Leu Trp Leu Leu Leu Ile Ala Ala Ala

1 5 10 15

Phe Ser Pro Gln Gly Leu Ala Gly Pro Ala Ser Val Pro Thr Thr Cys

20 25 30

Cys Phe Asn Leu Ala Asn Arg Lys Ile Pro Leu Gln Arg Leu Glu Ser

35 40 45

Tyr Arg Arg Ile Thr Ser Gly Lys Cys Pro Gln Lys Ala Val Ile Phe

50 55 60

Lys Thr Lys Leu Ala Lys Asp Ile Cys Ala Asp Pro Lys Lys Lys Trp

65 70 75 80

Val Gln Asp Ser Met Lys Tyr Leu Asp Gln Lys Ser Pro Thr Pro Lys

85 90 95

Pro

<210> 728

<211> 104

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL12 sequence"

<400> 728

Met Lys Ile Ser Thr Leu Leu Cys Leu Leu Leu Ile Ala Thr Thr Ile

1 5 10 15

Ser Pro Gln Val Leu Ala Gly Pro Asp Ala Val Ser Thr Pro Val Thr

20 25 30

Cys Cys Tyr Asn Val Val Lys Gln Lys Ile His Val Arg Lys Leu Lys

35 40 45

Ser Tyr Arg Arg Ile Thr Ser Ser Gln Cys Pro Arg Glu Ala Val Ile

50 55 60

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

65 70 75 80

Trp Val Lys Asn Ser Ile Asn His Leu Asp Lys Thr Ser Gln Thr Phe

85 90 95

Ile Leu Glu Pro Ser Cys Leu Gly

100

<210> 729

<211> 98

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL13 sequence"

<400> 729

Met Lys Val Ser Ala Val Leu Leu Cys Leu Leu Leu Met Thr Ala Ala

1 5 10 15

Phe Asn Pro Gln Gly Leu Ala Gln Pro Asp Ala Leu Asn Val Pro Ser

20 25 30

Thr Cys Cys Phe Thr Phe Ser Ser Lys Lys Ile Ser Leu Gln Arg Leu

35 40 45

Lys Ser Tyr Val Ile Thr Thr Ser Arg Cys Pro Gln Lys Ala Val Ile

50 55 60

Phe Arg Thr Lys Leu Gly Lys Glu Ile Cys Ala Asp Pro Lys Glu Lys

65 70 75 80

Trp Val Gln Asn Tyr Met Lys His Leu Gly Arg Lys Ala His Thr Leu
85 90 95
Lys Thr

<210> 730
<211> 93
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CCL14 sequence"
<400> 730
Met Lys Ile Ser Val Ala Ala Ile Pro Phe Phe Leu Leu Ile Thr Ile
1 5 10 15
Ala Leu Gly Thr Lys Thr Glu Ser Ser Ser Arg Gly Pro Tyr His Pro
20 25 30

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

<210>
731
<211> 113
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CCL15 sequence"

<400> 731

Met Lys Val Ser Val Ala Ala Leu Ser Cys Leu Met Leu Val Ala Val
1 5 10 15
Leu Gly Ser Gln Ala Gln Phe Ile Asn Asp Ala Glu Thr Glu Leu Met
20 25 30
Met Ser Lys Leu Pro Leu Glu Asn Pro Val Val Leu Asn Ser Phe His
35 40 45

Phe Ala Ala Asp Cys Cys Thr Ser Tyr Ile Ser Gln Ser Ile Pro Cys
50 55 60
Ser Leu Met Lys Ser Tyr Phe Glu Thr Ser Ser Glu Cys Ser Lys Pro
65 70 75 80
Gly Val Ile Phe Leu Thr Lys Lys Gly Arg Gln Val Cys Ala Lys Pro
85 90 95
Ser Gly Pro Gly Val Gln Asp Cys Met Lys Lys Leu Lys Pro Tyr Ser
100 105 110

Ile

<210> 732

<211> 120

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CCL16 sequence"

<400> 732

Met Lys Val Ser Glu Ala Ala Leu Ser Leu Leu Val Leu Ile Leu Ile
1 5 10 15
Ile Thr Ser Ala Ser Arg Ser Gln Pro Lys Val Pro Glu Trp Val Asn
20 25 30
Thr Pro Ser Thr Cys Cys Leu Lys Tyr Tyr Glu Lys Val Leu Pro Arg
35 40 45

Arg Leu Val Val Gly Tyr Arg Lys Ala Leu Asn Cys His Leu Pro Ala

50 55 60

Ile Ile Phe Val Thr Lys Arg Asn Arg Glu Val Cys Thr Asn Pro Asn

65 70 75 80

Asp Asp Trp Val Gln Glu Tyr Ile Lys Asp Pro Asn Leu Pro Leu Leu

85 90 95

Pro Thr Arg Asn Leu Ser Thr Val Lys Ile Ile Thr Ala Lys Asn Gly

100 105 110

Gln Pro Gln Leu Leu Asn Ser Gln

115 120

<210> 733

<211> 94

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL17 sequence"

<400> 733

Met Ala Pro Leu Lys Met Leu Ala Leu Val Thr Leu Leu Leu Gly Ala

1 5 10 15

Ser Leu Gln His Ile His Ala Ala Arg Gly Thr Asn Val Gly Arg Glu

20 25 30

Cys Cys Leu Glu Tyr Phe Lys Gly Ala Ile Pro Leu Arg Lys Leu Lys

35 40 45

Thr Trp Tyr Gln Thr Ser Glu Asp Cys Ser Arg Asp Ala Ile Val Phe

50 55 60

Val Thr Val Gln Gly Arg Ala Ile Cys Ser Asp Pro Asn Asn Lys Arg

65 70 75 80

Val Lys Asn Ala Val Lys Tyr Leu Gln Ser Leu Glu Arg Ser

85 90

<210> 734

<211> 89

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CCL18 sequence"

<400> 734

Met Lys Gly Leu Ala Ala Ala Leu Leu Val Leu Val Cys Thr Met Ala
1 5 10 15

Leu Cys Ser Cys Ala Gln Val Gly Thr Asn Lys Glu Leu Cys Cys Leu
20 25 30

Val Tyr Thr Ser Trp Gln Ile Pro Gln Lys Phe Ile Val Asp Tyr Ser
35 40 45

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

Arg Gly Arg Gln Ile Cys Ala Asp Pro Asn Lys Lys Trp Val Gln Lys
65 70 75 80

Tyr Ile Ser Asp Leu Lys Leu Asn Ala
85

<210> 735

<211> 98

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CCL19 sequence"

<400> 735

Met Ala Leu Leu Leu Ala Leu Ser Leu Leu Val Leu Trp Thr Ser Pro
1 5 10 15

Ala Pro Thr Leu Ser Gly Thr Asn Asp Ala Glu Asp Cys Cys Leu Ser
20 25 30

Val Thr Gln Lys Pro Ile Pro Gly Tyr Ile Val Arg Asn Phe His Tyr
35 40 45

Leu Leu Ile Lys Asp Gly Cys Arg Val Pro Ala Val Val Phe Thr Thr
50 55 60

Leu Arg Gly Arg Gln Leu Cys Ala Pro Pro Asp Gln Pro Trp Val Glu
65 70 75 80

Arg Ile Ile Gln Arg Leu Gln Arg Thr Ser Ala Lys Met Lys Arg Arg
85 90 95
Ser Ser

<210> 736
<211> 96
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CCL20 sequence"
<400> 736
Met Cys Cys Thr Lys Ser Leu Leu Leu Ala Ala Leu Met Ser Val Leu
1 5 10 15
Leu Leu His Leu Cys Gly Glu Ser Glu Ala Ala Ser Asn Phe Asp Cys
20 25 30

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

<210> 737
<211> 134
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:

CCL21 sequence"

<400> 737

Met Ala Gln Ser Leu Ala Leu Ser Leu Leu Ile Leu Val Leu Ala Phe
1 5 10 15

Gly Ile Pro Arg Thr Gln Gly Ser Asp Gly Gly Ala Gln Asp Cys Cys
20 25 30

Leu Lys Tyr Ser Gln Arg Lys Ile Pro Ala Lys Val Val Arg Ser Tyr
35 40 45

Arg Lys Gln Glu Pro Ser Leu Gly Cys Ser Ile Pro Ala Ile Leu Phe
50 55 60

Leu Pro Arg Lys Arg Ser Gln Ala Glu Leu Cys Ala Asp Pro Lys Glu
65 70 75 80

Leu Trp Val Gln Gln Leu Met Gln His Leu Asp Lys Thr Pro Ser Pro
85 90 95

Gln Lys Pro Ala Gln Gly Cys Arg Lys Asp Arg Gly Ala Ser Lys Thr
100 105 110

Gly Lys Lys Gly Lys Gly Ser Lys Gly Cys Lys Arg Thr Glu Arg Ser
115 120 125

Gln Thr Pro Lys Gly Pro
130

<210> 738

<211> 93

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL22 sequence"

<400> 738

Met Asp Arg Leu Gln Thr Ala Leu Leu Val Val Leu Val Leu Leu Ala
1 5 10 15

Val Ala Leu Gln Ala Thr Glu Ala Gly Pro Tyr Gly Ala Asn Met Glu

20 25 30

Asp Ser Val Cys Cys Arg Asp Tyr Val Arg Tyr Arg Leu Pro Leu Arg

35 40 45

Val Val Lys His Phe Tyr Trp Thr Ser Asp Ser Cys Pro Arg Pro Gly

50 55 60

Val Val Leu Leu Thr Phe Arg Asp Lys Glu Ile Cys Ala Asp Pro Arg

65 70 75 80

Val Pro Trp Val Lys Met Ile Leu Asn Lys Leu Ser Gln

85 90

<210> 739

<211> 120

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL23 sequence"

<400> 739

Met Lys Val Ser Val Ala Ala Leu Ser Cys Leu Met Leu Val Thr Ala

1 5 10 15

Leu Gly Ser Gln Ala Arg Val Thr Lys Asp Ala Glu Thr Glu Phe Met

20 25 30

Met Ser Lys Leu Pro Leu Glu Asn Pro Val Leu Leu Asp Arg Phe His

35 40 45

Ala Thr Ser Ala Asp Cys Cys Ile Ser Tyr Thr Pro Arg Ser Ile Pro

50 55 60

Cys Ser Leu Leu Glu Ser Tyr Phe Glu Thr Asn Ser Glu Cys Ser Lys

65 70 75 80

Pro Gly Val Ile Phe Leu Thr Lys Lys Gly Arg Arg Phe Cys Ala Asn

85 90 95

Pro Ser Asp Lys Gln Val Gln Val Cys Val Arg Met Leu Lys Leu Asp

100 105 110

Thr Arg Ile Lys Thr Arg Lys Asn

115 120

<210> 740

<211> 119

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL24 sequence"

<400> 740

Met Ala Gly Leu Met Thr Ile Val Thr Ser Leu Leu Phe Leu Gly Val

1 5 10 15

Cys Ala His His Ile Ile Pro Thr Gly Ser Val Val Ile Pro Ser Pro

20 25 30

Cys Cys Met Phe Phe Val Ser Lys Arg Ile Pro Glu Asn Arg Val Val

35 40 45

Ser Tyr Gln Leu Ser Ser Arg Ser Thr Cys Leu Lys Ala Gly Val Ile

50 55 60

Phe Thr Thr Lys Lys Gly Gln Gln Phe Cys Gly Asp Pro Lys Gln Glu

65 70 75 80

Trp Val Gln Arg Tyr Met Lys Asn Leu Asp Ala Lys Gln Lys Lys Ala

85 90 95

Ser Pro Arg Ala Arg Ala Val Ala Val Lys Gly Pro Val Gln Arg Tyr

100 105 110

Pro Gly Asn Gln Thr Thr Cys

115

<210> 741

<211> 150

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL25 sequence"

<400> 741

Met Asn Leu Trp Leu Leu Ala Cys Leu Val Ala Gly Phe Leu Gly Ala

1 5 10 15

Trp Ala Pro Ala Val His Thr Gln Gly Val Phe Glu Asp Cys Cys Leu

20 25 30

Ala Tyr His Tyr Pro Ile Gly Trp Ala Val Leu Arg Arg Ala Trp Thr

35 40 45

Tyr Arg Ile Gln Glu Val Ser Gly Ser Cys Asn Leu Pro Ala Ala Ile

50 55 60

Phe Tyr Leu Pro Lys Arg His Arg Lys Val Cys Gly Asn Pro Lys Ser

65 70 75 80

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

85 90 95

Phe Ala Lys Leu His His Asn Thr Gln Thr Phe Gln Ala Gly Pro His

100 105 110

Ala Val Lys Lys Leu Ser Ser Gly Asn Ser Lys Leu Ser Ser Ser Lys

115 120 125

Phe Ser Asn Pro Ile Ser Ser Ser Lys Arg Asn Val Ser Leu Leu Ile

130 135 140

Ser Ala Asn Ser Gly Leu

145 150

<210> 742

<211> 94

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CCL26 sequence"

<400> 742

Met Met Gly Leu Ser Leu Ala Ser Ala Val Leu Leu Ala Ser Leu Leu

1 5 10 15

Ser Leu His Leu Gly Thr Ala Thr Arg Gly Ser Asp Ile Ser Lys Thr

20 25 30

Cys Cys Phe Gln Tyr Ser His Lys Pro Leu Pro Trp Thr Trp Val Arg

35 40 45

Ser Tyr Glu Phe Thr Ser Asn Ser Cys Ser Gln Arg Ala Val Ile Phe
50 55 60
Thr Thr Lys Arg Gly Lys Lys Val Cys Thr His Pro Arg Lys Lys Trp
65 70 75 80
Val Gln Lys Tyr Ile Ser Leu Leu Lys Thr Pro Lys Gln Leu
85 90

<210> 743
<211> 112
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CCL27 sequence"

<400> 743
Met Lys Gly Pro Pro Thr Phe Cys Ser Leu Leu Leu Leu Ser Leu Leu
1 5 10 15
Leu Ser Pro Asp Pro Thr Ala Ala Phe Leu Leu Pro Pro Ser Thr Ala
20 25 30
Cys Cys Thr Gln Leu Tyr Arg Lys Pro Leu Ser Asp Lys Leu Leu Arg
35 40 45
Lys Val Ile Gln Val Glu Leu Gln Glu Ala Asp Gly Asp Cys His Leu
50 55 60

Gln Ala Phe Val Leu His Leu Ala Gln Arg Ser Ile Cys Ile His Pro
65 70 75 80
Gln Asn Pro Ser Leu Ser Gln Trp Phe Glu His Gln Glu Arg Lys Leu
85 90 95
His Gly Thr Leu Pro Lys Leu Asn Phe Gly Met Leu Arg Lys Met Gly
100 105 110

<210> 744
<211> 127
<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CCL28 sequence"

<400> 744

Met Gln Gln Arg Gly Leu Ala Ile Val Ala Leu Ala Val Cys Ala Ala
1 5 10 15

Leu His Ala Ser Glu Ala Ile Leu Pro Ile Ala Ser Ser Cys Cys Thr
20 25 30

Glu Val Ser His His Ile Ser Arg Arg Leu Leu Glu Arg Val Asn Met
35 40 45

Cys Arg Ile Gln Arg Ala Asp Gly Asp Cys Asp Leu Ala Ala Val Ile
50 55 60

Leu His Val Lys Arg Arg Arg Ile Cys Val Ser Pro His Asn His Thr
65 70 75 80

Val Lys Gln Trp Met Lys Val Gln Ala Ala Lys Lys Asn Gly Lys Gly
85 90 95

Asn Val Cys His Arg Lys Lys His His Gly Lys Arg Asn Ser Asn Arg
100 105 110

Ala His Gln Gly Lys His Glu Thr Tyr Gly His Lys Thr Pro Tyr
115 120 125

<210> 745

<211> 107

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CXCL1 sequence"

<400> 745

Met Ala Arg Ala Ala Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu
1 5 10 15

Arg Val Ala Leu Leu Leu Leu Leu Leu Val Ala Ala Gly Arg Arg Ala

20 25 30

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

35 40 45

Leu Gln Gly Ile His Pro Lys Asn Ile Gln Ser Val Asn Val Lys Ser

50 55 60

Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn

65 70 75 80

Gly Arg Lys Ala Cys Leu Asn Pro Ala Ser Pro Ile Val Lys Lys Ile

85 90 95

Ile Glu Lys Met Leu Asn Ser Asp Lys Ser Asn

100 105

<210> 746

<211> 107

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CXCL2 sequence"

<400> 746

Met Ala Arg Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu

1 5 10 15

Arg Val Ala Leu Leu Leu Leu Leu Leu Val Ala Ala Ser Arg Arg Ala

20 25 30

Ala Gly Ala Pro Leu Ala Thr Glu Leu Arg Cys Gln Cys Leu Gln Thr

35 40 45

Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Lys Val Lys Ser

50 55 60

Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn

65 70 75 80

Gly Gln Lys Ala Cys Leu Asn Pro Ala Ser Pro Met Val Lys Lys Ile

85 90 95

Ile Glu Lys Met Leu Lys Asn Gly Lys Ser Asn

100 105

<210> 747

<211> 107

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CXCL3 sequence"

<400> 747

Met Ala His Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu
1 5 10 15

Arg Val Ala Leu Leu Leu Leu Leu Leu Val Ala Ala Ser Arg Arg Ala
20 25 30

Ala Gly Ala Ser Val Val Thr Glu Leu Arg Cys Gln Cys Leu Gln Thr
35 40 45

Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Asn Val Arg Ser
50 55 60

Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn
65 70 75 80

Gly Lys Lys Ala Cys Leu Asn Pro Ala Ser Pro Met Val Gln Lys Ile
85 90 95

Ile Glu Lys Ile Leu Asn Lys Gly Ser Thr Asn
100 105

<210> 748

<211> 365

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CXCL4 sequence"

<400> 748

Met Ala Tyr Tyr Glu Glu Ser Ile Ile Phe Asp Thr Asp Asn Ser Ser

1 5 10 15

Glu Glu Ser Gly Asp Phe Asp Val Asp Phe Gln Gly Pro Cys Glu Arg

20 25 30

Ala Pro Asn Tyr Asp Phe His Lys Ile Phe Leu Pro Thr Val Phe Gly

35 40 45

Ile Ile Phe Ile Leu Gly Ile Phe Gly Asn Gly Leu Val Val Ile Val

50 55 60

Met Gly Phe Gln Asn Lys Cys Lys Thr Ser Met Thr Asp Lys Tyr Arg

65 70 75 80

Leu His Leu Ser Val Ala Asp Leu Met Phe Val Leu Thr Leu Pro Phe

85 90 95

Trp Ala Val Asp Ala Ala Ser Ser Trp Tyr Phe Gly Gly Phe Leu Cys

100 105 110

Lys Ala Val His Met Ile Tyr Thr Val Asn Leu Tyr Ser Ser Val Leu

115 120 125

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

130 135 140

Thr Ala Thr Ser Ser Gln Ala Thr Arg Lys Leu Leu Ala Gly Lys Val

145 150 155 160

Ile Tyr Val Cys Val Trp Leu Pro Ala Ala Ile Leu Thr Val Pro Asp

165 170 175

Leu Val Phe Ala Ser Ser Phe Glu Val Glu Leu Asp Met Gly Gly Ser

180 185 190

Arg Met Ile Cys Gln Arg Ile Tyr Pro Leu Glu Ser Asn His Ile Trp

195 200 205

Val Ala Ala Phe Arg Phe Gln His Ile Leu Val Gly Phe Val Leu Pro

210 215 220

Gly Leu Val Ile Leu Thr Cys Tyr Cys Ile Ile Ile Ser Lys Leu Ser

225 230 235 240

Gln Gly Ser Lys Gly Gln Ala Leu Lys Arg Lys Ala Leu Lys Thr Thr

245 250 255

Val Ile Leu Ile Leu Cys Phe Phe Cys Cys Trp Leu Pro Tyr Cys Ala
260 265 270
Gly Ile Phe Val Asp Thr Leu Met Asn Leu Glu Val Ile Pro His Ser
275 280 285
Cys Ala Leu Glu Gln Gly Val Met Lys Trp Ile Ser Ile Thr Glu Val
290 295 300

Leu Ala Tyr Phe His Cys Cys Leu Asn Pro Ile Leu Tyr Ala Phe Leu
305 310 315 320
Gly Val Lys Phe Lys Lys Ser Ala Arg Asn Ala Leu Thr Phe Ser Ser
325 330 335
Arg Ser Ser His Lys Met Leu Thr Lys Lys Arg Gly Ala Ile Ser Ser
340 345 350
Ala Ser Thr Glu Ser Glu Ser Ser Ser Val Leu Tyr Ser
355 360 365

<210> 749

<211> 114
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CXCL5 sequence"
<400> 749
Met Ser Leu Leu Ser Ser Arg Ala Ala Arg Val Pro Gly Pro Ser Ser
1 5 10 15
Ser Leu Cys Ala Leu Leu Val Leu Leu Leu Leu Leu Thr Gln Pro Gly
20 25 30
Pro Ile Ala Ser Ala Gly Pro Ala Ala Ala Val Leu Arg Glu Leu Arg
35 40 45

Cys Val Cys Leu Gln Thr Thr Gln Gly Val His Pro Lys Met Ile Ser
50 55 60
Asn Leu Gln Val Phe Ala Ile Gly Pro Gln Cys Ser Lys Val Glu Val
65 70 75 80

Val Ala Ser Leu Lys Asn Gly Lys Glu Ile Cys Leu Asp Pro Glu Ala
85 90 95
Pro Phe Leu Lys Lys Val Ile Gln Lys Ile Leu Asp Gly Gly Asn Lys
100 105 110

Glu Asn

<210> 750
<211> 114
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CXCL6 sequence"
<400> 750
Met Ser Leu Pro Ser Ser Arg Ala Ala Arg Val Pro Gly Pro Ser Gly
1 5 10 15
Ser Leu Cys Ala Leu Leu Ala Leu Leu Leu Leu Leu Thr Pro Pro Gly
20 25 30
Pro Leu Ala Ser Ala Gly Pro Val Ser Ala Val Leu Thr Glu Leu Arg
35 40 45

Cys Thr Cys Leu Arg Val Thr Leu Arg Val Asn Pro Lys Thr Ile Gly
50 55 60
Lys Leu Gln Val Phe Pro Ala Gly Pro Gln Cys Ser Lys Val Glu Val
65 70 75 80
Val Ala Ser Leu Lys Asn Gly Lys Gln Val Cys Leu Asp Pro Glu Ala
85 90 95
Pro Phe Leu Lys Lys Val Ile Gln Lys Ile Leu Asp Ser Gly Asn Lys
100 105 110

Lys Asn

<210> 751
<211> 128

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CXCL7 sequence"

<400> 751

Met Ser Leu Arg Leu Asp Thr Thr Pro Ser Cys Asn Ser Ala Arg Pro

1 5 10 15

Leu His Ala Leu Gln Val Leu Leu Leu Leu Ser Leu Leu Leu Thr Ala

20 25 30

Leu Ala Ser Ser Thr Lys Gly Gln Thr Lys Arg Asn Leu Ala Lys Gly

35 40 45

Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala Glu Leu Arg Cys Met

50 55 60

Cys Ile Lys Thr Thr Ser Gly Ile His Pro Lys Asn Ile Gln Ser Leu

65 70 75 80

Glu Val Ile Gly Lys Gly Thr His Cys Asn Gln Val Glu Val Ile Ala

85 90 95

Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp Pro Asp Ala Pro Arg

100 105 110

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

115 120 125

<210> 752

<211> 99

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

CXCL8 sequence"

<400> 752

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

1 5 10 15

Ala Ala Leu Cys Glu Gly Ala Val Leu Pro Arg Ser Ala Lys Glu Leu

20 25 30

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

Glu Asn Ser

<210> 753
<211> 125
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CXCL9 sequence"
<400> 753
Met Lys Lys Ser Gly Val Leu Phe Leu Leu Gly Ile Ile Leu Leu Val
1 5 10 15
Leu Ile Gly Val Gln Gly Thr Pro Val Val Arg Lys Gly Arg Cys Ser
20 25 30
Cys Ile Ser Thr Asn Gln Gly Thr Ile His Leu Gln Ser Leu Lys Asp
35 40 45

Leu Lys Gln Phe Ala Pro Ser Pro Ser Cys Glu Lys Ile Glu Ile Ile
50 55 60
Ala Thr Leu Lys Asn Gly Val Gln Thr Cys Leu Asn Pro Asp Ser Ala
65 70 75 80
Asp Val Lys Glu Leu Ile Lys Lys Trp Glu Lys Gln Val Ser Gln Lys
85 90 95
Lys Lys Gln Lys Asn Gly Lys Lys His Gln Lys Lys Lys Val Leu Lys

100 105 110

Val Arg Lys Ser Gln Arg Ser Arg Gln Lys Lys Thr Thr
115 120 125

<210> 754

<211> 98

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CXCL10 sequence"

<400> 754

Met Asn Gln Thr Ala Ile Leu Ile Cys Cys Leu Ile Phe Leu Thr Leu
1 5 10 15
Ser Gly Ile Gln Gly Val Pro Leu Ser Arg Thr Val Arg Cys Thr Cys
20 25 30

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

Ser Pro

<210> 755

<211> 94

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CXCL11 sequence"

<400> 755

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

Ser Ile Met Tyr Pro Ser Asn Asn Cys Asp Lys Ile Glu Val Ile Ile
50 55 60
Thr Leu Lys Glu Asn Lys Gly Gln Arg Cys Leu Asn Pro Lys Ser Lys
65 70 75 80
Gln Ala Arg Leu Ile Ile Lys Lys Val Glu Arg Lys Asn Phe
85 90

<210> 756

<211> 93

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CXCL12 sequence"

<400> 756

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

Asn Asn Asn Arg Gln Val Cys Ile Asp Pro Lys Leu Lys Trp Ile Gln
65 70 75 80
Glu Tyr Leu Glu Lys Ala Leu Asn Lys Arg Phe Lys Met

85 90

<210> 757

<211> 109

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CXCL13 sequence"

<400> 757

Met Lys Phe Ile Ser Thr Ser Leu Leu Leu Met Leu Leu Val Ser Ser
1 5 10 15

Leu Ser Pro Val Gln Gly Val Leu Glu Val Tyr Tyr Thr Ser Leu Arg
20 25 30

Cys Arg Cys Val Gln Glu Ser Ser Val Phe Ile Pro Arg Arg Phe Ile
35 40 45

Asp Arg Ile Gln Ile Leu Pro Arg Gly Asn Gly Cys Pro Arg Lys Glu
50 55 60

Ile Ile Val Trp Lys Lys Asn Lys Ser Ile Val Cys Val Asp Pro Gln
65 70 75 80

Ala Glu Trp Ile Gln Arg Met Met Glu Val Leu Arg Lys Arg Ser Ser
85 90 95

Ser Thr Leu Pro Val Pro Val Phe Lys Arg Lys Ile Pro
100 105

<210> 758

<211> 111

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
CXCL14 sequence"

<400> 758

Met Ser Leu Leu Pro Arg Arg Ala Pro Pro Val Ser Met Arg Leu Leu

1 5 10 15

Ala Ala Ala Leu Leu Leu Leu Leu Leu Ala Leu Tyr Thr Ala Arg Val
20 25 30
Asp Gly Ser Lys Cys Lys Cys Ser Arg Lys Gly Pro Lys Ile Arg Tyr
35 40 45
Ser Asp Val Lys Lys Leu Glu Met Lys Pro Lys Tyr Pro His Cys Glu
50 55 60
Glu Lys Met Val Ile Ile Thr Thr Lys Ser Val Ser Arg Tyr Arg Gly
65 70 75 80

Gln Glu His Cys Leu His Pro Lys Leu Gln Ser Thr Lys Arg Phe Ile
85 90 95
Lys Trp Tyr Asn Ala Trp Asn Glu Lys Arg Arg Val Tyr Glu Glu
100 105 110

<210> 759
<211> 167
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CXCL15 sequence"
<400> 759
Met Ala Ala Gln Gly Trp Ser Met Leu Leu Leu Ala Val Leu Asn Leu
1 5 10 15

Gly Ile Phe Val Arg Pro Cys Asp Thr Gln Glu Leu Arg Cys Leu Cys
20 25 30
Ile Gln Glu His Ser Glu Phe Ile Pro Leu Lys Leu Ile Lys Asn Ile
35 40 45
Met Val Ile Phe Glu Thr Ile Tyr Cys Asn Arg Lys Glu Val Ile Ala
50 55 60
Val Pro Lys Asn Gly Ser Met Ile Cys Leu Asp Pro Asp Ala Pro Trp
65 70 75 80

Val Lys Ala Thr Val Gly Pro Ile Thr Asn Arg Phe Leu Pro Glu Asp
85 90 95
Leu Lys Gln Lys Glu Phe Pro Pro Ala Met Lys Leu Leu Tyr Ser Val
100 105 110
Glu His Glu Lys Pro Leu Tyr Leu Ser Phe Gly Arg Pro Glu Asn Lys
115 120 125
Arg Ile Phe Pro Phe Pro Ile Arg Glu Thr Ser Arg His Phe Ala Asp
130 135 140

Leu Ala His Asn Ser Asp Arg Asn Phe Leu Arg Asp Ser Ser Glu Val
145 150 155 160
Ser Leu Thr Gly Ser Asp Ala
165

<210> 760
<211> 250
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CXCL16 sequence"
<400> 760
Met Gly Arg Asp Leu Arg Pro Gly Ser Arg Val Leu Leu Leu Leu Leu
1 5 10 15
Leu Leu Leu Leu Val Tyr Leu Thr Gln Pro Gly Asn Gly Asn Glu Gly

20 25 30
Ser Val Thr Gly Ser Cys Tyr Cys Gly Lys Arg Ile Ser Ser Asp Ser
35 40 45
Pro Pro Ser Val Gln Phe Met Asn Arg Leu Arg Lys His Leu Arg Ala
50 55 60
Tyr His Arg Cys Leu Tyr Tyr Thr Arg Phe Gln Leu Leu Ser Trp Ser
65 70 75 80
Val Cys Gly Gly Asn Lys Asp Pro Trp Val Gln Glu Leu Met Ser Cys

85 90 95

Leu Asp Leu Lys Glu Cys Gly His Ala Tyr Ser Gly Ile Val Ala His
100 105 110
Gln Lys His Leu Leu Pro Thr Ser Pro Pro Ile Ser Gln Ala Ser Glu
115 120 125
Gly Ala Ser Ser Asp Ile His Thr Pro Ala Gln Met Leu Leu Ser Thr
130 135 140
Leu Gln Ser Thr Gln Arg Pro Thr Leu Pro Val Gly Ser Leu Ser Ser

145 150 155 160
Asp Lys Glu Leu Thr Arg Pro Asn Glu Thr Thr Ile His Thr Ala Gly
165 170 175
His Ser Leu Ala Ala Gly Pro Glu Ala Gly Glu Asn Gln Lys Gln Pro
180 185 190
Glu Lys Asn Ala Gly Pro Thr Ala Arg Thr Ser Ala Thr Val Pro Val
195 200 205
Leu Cys Leu Leu Ala Ile Ile Phe Ile Leu Thr Ala Ala Leu Ser Tyr

210 215 220
Val Leu Cys Lys Arg Arg Arg Gly Gln Ser Pro Gln Ser Ser Pro Asp
225 230 235 240
Leu Pro Val His Tyr Ile Pro Val Ala Pro
245 250

<210> 761
<211> 119
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
CXCL17 sequence"
<400> 761
Met Lys Val Leu Ile Ser Ser Leu Leu Leu Leu Leu Pro Leu Met Leu
1 5 10 15

Met Ser Met Val Ser Ser Ser Leu Asn Pro Gly Val Ala Arg Gly His
20 25 30

Arg Asp Arg Gly Gln Ala Ser Arg Arg Trp Leu Gln Glu Gly Gly Gln

35 40 45

Glu Cys Glu Cys Lys Asp Trp Phe Leu Arg Ala Pro Arg Arg Lys Phe

50 55 60

Met Thr Val Ser Gly Leu Pro Lys Lys Gln Cys Pro Cys Asp His Phe

65 70 75 80

Lys Gly Asn Val Lys Lys Thr Arg His Gln Arg His His Arg Lys Pro

85 90 95

Asn Lys His Ser Arg Ala Cys Gln Gln Phe Leu Lys Gln Cys Gln Leu

100 105 110

Arg Ser Phe Ala Leu Pro Leu

115

<210> 762

<211> 114

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

XCL1 sequence"

<400> 762

Met Arg Leu Leu Ile Leu Ala Leu Leu Gly Ile Cys Ser Leu Thr Ala

1 5 10 15

Tyr Ile Val Glu Gly Val Gly Ser Glu Val Ser Asp Lys Arg Thr Cys

20 25 30

Val Ser Leu Thr Thr Gln Arg Leu Pro Val Ser Arg Ile Lys Thr Tyr

35 40 45

Thr Ile Thr Glu Gly Ser Leu Arg Ala Val Ile Phe Ile Thr Lys Arg

50 55 60

Gly Leu Lys Val Cys Ala Asp Pro Gln Ala Thr Trp Val Arg Asp Val

65 70 75 80

Val Arg Ser Met Asp Arg Lys Ser Asn Thr Arg Asn Asn Met Ile Gln

85 90 95

Thr Lys Pro Thr Gly Thr Gln Gln Ser Thr Asn Thr Ala Val Thr Leu
100 105 110
Thr Gly

<210> 763
<211> 114
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
XCL2 sequence"
<400> 763
Met Arg Leu Leu Ile Leu Ala Leu Leu Gly Ile Cys Ser Leu Thr Ala

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

65 70 75 80
Val Arg Ser Met Asp Arg Lys Ser Asn Thr Arg Asn Asn Met Ile Gln
85 90 95
Thr Lys Pro Thr Gly Thr Gln Gln Ser Thr Asn Thr Ala Val Thr Leu
100 105 110
Thr Gly

<210> 764
<211> 397
<212> PRT
<213> Unknown
<220><221> source

<223> /note="Description of Unknown:

CX3CL1 sequence"

<400> 764

Met Ala Pro Ile Ser Leu Ser Trp Leu Leu Arg Leu Ala Thr Phe Cys

1 5 10 15

His Leu Thr Val Leu Leu Ala Gly Gln His His Gly Val Thr Lys Cys

20 25 30

Asn Ile Thr Cys Ser Lys Met Thr Ser Lys Ile Pro Val Ala Leu Leu

35 40 45

Ile His Tyr Gln Gln Asn Gln Ala Ser Cys Gly Lys Arg Ala Ile Ile

50 55 60

Leu Glu Thr Arg Gln His Arg Leu Phe Cys Ala Asp Pro Lys Glu Gln

65 70 75 80

Trp Val Lys Asp Ala Met Gln His Leu Asp Arg Gln Ala Ala Ala Leu

85 90 95

Thr Arg Asn Gly Gly Thr Phe Glu Lys Gln Ile Gly Glu Val Lys Pro

100 105 110

Arg Thr Thr Pro Ala Ala Gly Gly Met Asp Glu Ser Val Val Leu Glu

115 120 125

Pro Glu Ala Thr Gly Glu Ser Ser Ser Leu Glu Pro Thr Pro Ser Ser

130 135 140

Gln Glu Ala Gln Arg Ala Leu Gly Thr Ser Pro Glu Leu Pro Thr Gly

145 150 155 160

Val Thr Gly Ser Ser Gly Thr Arg Leu Pro Pro Thr Pro Lys Ala Gln

165 170 175

Asp Gly Gly Pro Val Gly Thr Glu Leu Phe Arg Val Pro Pro Val Ser

180 185 190

Thr Ala Ala Thr Trp Gln Ser Ser Ala Pro His Gln Pro Gly Pro Ser

195 200 205

Leu Trp Ala Glu Ala Lys Thr Ser Glu Ala Pro Ser Thr Gln Asp Pro

210 215 220

Ser Thr Gln Ala Ser Thr Ala Ser Ser Pro Ala Pro Glu Glu Asn Ala
225 230 235 240
Pro Ser Glu Gly Gln Arg Val Trp Gly Gln Gly Gln Ser Pro Arg Pro
245 250 255
Glu Asn Ser Leu Glu Arg Glu Glu Met Gly Pro Val Pro Ala His Thr

260 265 270
Asp Ala Phe Gln Asp Trp Gly Pro Gly Ser Met Ala His Val Ser Val
275 280 285
Val Pro Val Ser Ser Glu Gly Thr Pro Ser Arg Glu Pro Val Ala Ser
290 295 300
Gly Ser Trp Thr Pro Lys Ala Glu Glu Pro Ile His Ala Thr Met Asp
305 310 315 320
Pro Gln Arg Leu Gly Val Leu Ile Thr Pro Val Pro Asp Ala Gln Ala

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

385 390 395
<210> 765
<211> 223
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
Tbet sequence"
<400> 765
His Leu Leu Trp Ser Lys Phe Asn Gln His Gln Thr Glu Met Ile Ile
1 5 10 15

Thr Lys Gln Gly Arg Arg Met Phe Pro Phe Leu Ser Phe Thr Val Ala
20 25 30
Gly Leu Glu Pro Thr Ser His Tyr Arg Met Phe Val Asp Val Val Leu

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

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

165 170 175
Phe Ala Lys Gly Phe Arg Glu Asn Phe Glu Ser Met Tyr Ala Ser Val
180 185 190
Asp Thr Ser Val Pro Ser Pro Pro Gly Pro Asn Cys Gln Leu Leu Gly
195 200 205
Gly Asp Pro Tyr Ser Pro Leu Leu Ser Asn Gln Tyr Pro Val Pro
210 215 220
<210> 766
<211> 443
<212> PRT
<213> Unknown
<220
><221> source

<223> /note="Description of Unknown:

GATA3 sequence"

<400> 766

Met Glu Val Thr Ala Asp Gln Pro Arg Trp Val Ser His His His Pro

1 5 10 15

Ala Val Leu Asn Gly Gln His Pro Asp Thr His His Pro Gly Leu Ser

20 25 30

His Ser Tyr Met Asp Ala Ala Gln Tyr Pro Leu Pro Glu Glu Val Asp

35 40 45

Val Leu Phe Asn Ile Asp Gly Gln Gly Asn His Val Pro Pro Tyr Tyr

50 55 60

Gly Asn Ser Val Arg Ala Thr Val Gln Arg Tyr Pro Pro Thr His His

65 70 75 80

Gly Ser Gln Val Cys Arg Pro Pro Leu Leu His Gly Ser Leu Pro Trp

85 90 95

Leu Asp Gly Gly Lys Ala Leu Gly Ser His His Thr Ala Ser Pro Trp

100 105 110

Asn Leu Ser Pro Phe Ser Lys Thr Ser Ile His His Gly Ser Pro Gly

115 120 125

Pro Leu Ser Val Tyr Pro Pro Ala Ser Ser Ser Ser Leu Ser Gly Gly

130 135 140

His Ala Ser Pro His Leu Phe Thr Phe Pro Pro Thr Pro Pro Lys Asp

145 150 155 160

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

165 170 175

Arg Gln Asp Glu Lys Glu Cys Leu Lys Tyr Gln Val Pro Leu Pro Asp

180 185 190

Ser Met Lys Leu Glu Ser Ser His Ser Arg Gly Ser Met Thr Ala Leu

195 200 205

Gly Gly Ala Ser Ser Ser Thr His His Pro Ile Thr Thr Tyr Pro Pro

210 215 220

Tyr Val Pro Glu Tyr Ser Ser Gly Leu Phe Pro Pro Ser Ser Leu Leu

225 230 235 240

Gly Gly Ser Pro Thr Gly Phe Gly Cys Lys Ser Arg Pro Lys Ala Arg

245 250 255

Ser Ser Thr Gly Arg Glu Cys Val Asn Cys Gly Ala Thr Ser Thr Pro

260 265 270

Leu Trp Arg Arg Asp Gly Thr Gly His Tyr Leu Cys Asn Ala Cys Gly

275 280 285

Leu Tyr His Lys Met Asn Gly Gln Asn Arg Pro Leu Ile Lys Pro Lys

290 295 300

Arg Arg Leu Ser Ala Ala Arg Arg Ala Gly Thr Ser Cys Ala Asn Cys

305 310 315 320

Gln Thr Thr Thr Thr Thr Leu Trp Arg Arg Asn Ala Asn Gly Asp Pro

325 330 335

Val Cys Asn Ala Cys Gly Leu Tyr Tyr Lys Leu His Asn Ile Asn Arg

340 345 350

Pro Leu Thr Met Lys Lys Glu Gly Ile Gln Thr Arg Asn Arg Lys Met

355 360 365

Ser Ser Lys Ser Lys Lys Cys Lys Lys Val His Asp Ser Leu Glu Asp

370 375 380

Phe Pro Lys Asn Ser Ser Phe Asn Pro Ala Ala Leu Ser Arg His Met

385 390 395 400

Ser Ser Leu Ser His Ile Ser Pro Phe Ser His Ser Ser His Met Leu

405 410 415

Thr Thr Pro Thr Pro Met His Pro Pro Ser Ser Leu Ser Phe Gly Pro

420 425 430

His His Pro Ser Ser Met Val Thr Ala Met Gly

435 440

<210> 767

<211> 518

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
RORgt sequence"

<400> 767

Met Asp Arg Ala Pro Gln Arg Gln His Arg Ala Ser Arg Glu Leu Leu
1 5 10 15

Ala Ala Lys Lys Thr His Thr Ser Gln Ile Glu Val Ile Pro Cys Lys
20 25 30

Ile Cys Gly Asp Lys Ser Ser Gly Ile His Tyr Gly Val Ile Thr Cys
35 40 45

Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Gln Arg Cys Asn Ala Ala
50 55 60

Tyr Ser Cys Thr Arg Gln Gln Asn Cys Pro Ile Asp Arg Thr Ser Arg
65 70 75 80

Asn Arg Cys Gln His Cys Arg Leu Gln Lys Cys Leu Ala Leu Gly Met
85 90 95

Ser Arg Asp Ala Val Lys Phe Gly Arg Met Ser Lys Lys Gln Arg Asp
100 105 110

Ser Leu His Ala Glu Val Gln Lys Gln Leu Gln Gln Arg Gln Gln Gln
115 120 125

Gln Gln Glu Pro Val Val Lys Thr Pro Pro Ala Gly Ala Gln Gly Ala
130 135 140

Asp Thr Leu Thr Tyr Thr Leu Gly Leu Pro Asp Gly Gln Leu Pro Leu
145 150 155 160

Gly Ser Ser Pro Asp Leu Pro Glu Ala Ser Ala Cys Pro Pro Gly Leu
165 170 175

Leu Lys Ala Ser Gly Ser Gly Pro Ser Tyr Ser Asn Asn Leu Ala Lys
180 185 190

Ala Gly Leu Asn Gly Ala Ser Cys His Leu Glu Tyr Ser Pro Glu Arg
195 200 205

Gly Lys Ala Glu Gly Arg Glu Ser Phe Tyr Ser Thr Gly Ser Gln Leu
210 215 220

Thr Pro Asp Arg Cys Gly Leu Arg Phe Glu Glu His Arg His Pro Gly
225 230 235 240

Leu Gly Glu Leu Gly Gln Gly Pro Asp Ser Tyr Gly Ser Pro Ser Phe
245 250 255

Arg Ser Thr Pro Glu Ala Pro Tyr Ala Ser Leu Thr Glu Ile Glu His
260 265 270

Leu Val Gln Ser Val Cys Lys Ser Tyr Arg Glu Thr Cys Gln Leu Arg
275 280 285

Leu Glu Asp Leu Leu Arg Gln Arg Ser Asn Ile Phe Ser Arg Glu Glu
290 295 300

Val Thr Gly Tyr Gln Arg Lys Ser Met Trp Glu Met Trp Glu Arg Cys
305 310 315 320

Ala His His Leu Thr Glu Ala Ile Gln Tyr Val Val Glu Phe Ala Lys
325 330 335

Arg Leu Ser Gly Phe Met Glu Leu Cys Gln Asn Asp Gln Ile Val Leu
340 345 350

Leu Lys Ala Gly Ala Met Glu Val Val Leu Val Arg Met Cys Arg Ala
355 360 365

Tyr Asn Ala Asp Asn Arg Thr Val Phe Phe Glu Gly Lys Tyr Gly Gly
370 375 380

Met Glu Leu Phe Arg Ala Leu Gly Cys Ser Glu Leu Ile Ser Ser Ile
385 390 395 400

Phe Asp Phe Ser His Ser Leu Ser Ala Leu His Phe Ser Glu Asp Glu
405 410 415

Ile Ala Leu Tyr Thr Ala Leu Val Leu Ile Asn Ala His Arg Pro Gly
420 425 430

Leu Gln Glu Lys Arg Lys Val Glu Gln Leu Gln Tyr Asn Leu Glu Leu
435 440 445

Ala Phe His His His Leu Cys Lys Thr His Arg Gln Ser Ile Leu Ala
450 455 460

Lys Leu Pro Pro Lys Gly Lys Leu Arg Ser Leu Cys Ser Gln His Val

465 470 475 480

Glu Arg Leu Gln Ile Phe Gln His Leu His Pro Ile Val Val Gln Ala

485 490 495

Ala Phe Pro Pro Leu Tyr Lys Glu Leu Phe Ser Thr Glu Thr Glu Ser

500 505 510

Pro Val Gly Leu Ser Lys

515

<210> 768

<211> 268

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Cd25 sequence"

<400> 768

Met Glu Pro Cys Leu Leu Met Trp Gly Ile Leu Thr Phe Ile Thr Val

1 5 10 15

Ser Gly Tyr Thr Thr Asp Leu Cys Asp Asp Asp Pro Pro Asn Leu Lys

20 25 30

His Ala Thr Phe Lys Ala Leu Thr Tyr Lys Thr Gly Thr Val Leu Asn

35 40 45

Cys Asp Cys Glu Arg Gly Phe Arg Arg Ile Ser Ser Tyr Met His Cys

50 55 60

Thr Gly Asn Ser Ser His Ala Ser Trp Glu Asn Lys Cys Gln Cys Lys

65 70 75 80

Ser Ile Ser Pro Glu Asn Arg Lys Gly Lys Val Thr Thr Lys Pro Glu

85 90 95

Glu Gln Lys Gly Glu Asn Pro Thr Glu Met Gln Ser Gln Thr Pro Pro

100 105 110

Met Asp Glu Val Asp Leu Val Gly His Cys Arg Glu Pro Pro Pro Trp

115 120 125

Glu His Glu Asn Ser Lys Arg Ile Tyr His Phe Val Val Gly Gln Thr

130 135 140

Leu His Tyr Gln Cys Met Gln Gly Phe Thr Ala Leu His Arg Gly Pro

145 150 155 160

Ala Lys Ser Ile Cys Lys Thr Ile Phe Gly Lys Thr Arg Trp Thr Gln

165 170 175

Pro Pro Leu Lys Cys Ile Ser Glu Ser Gln Phe Pro Asp Asp Glu Glu

180 185 190

Leu Gln Ala Ser Thr Asp Ala Pro Ala Gly Arg Asp Thr Ser Ser Pro

195 200 205

Phe Ile Thr Thr Ser Thr Pro Asp Phe His Lys His Thr Glu Val Ala

210 215 220

Thr Thr Met Glu Ser Phe Ile Phe Thr Thr Glu Tyr Gln Ile Ala Val

225 230 235 240

Ala Ser Cys Val Leu Leu Leu Ile Ser Ile Val Leu Val Ser Gly Leu

245 250 255

Thr Trp Gln Arg Arg Arg Arg Lys Ser Arg Thr Ile

260 265

<210> 769

<211> 163

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Skp1 sequence"

<400> 769

Met Pro Ser Ile Lys Leu Gln Ser Ser Asp Gly Glu Ile Phe Glu Val

1 5 10 15

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

20 25 30

Asp Leu Gly Met Asp Asp Glu Gly Asp Asp Asp Pro Val Pro Leu Pro

35 40 45

Asn Val Asn Ala Ala Ile Leu Lys Lys Val Ile Gln Trp Cys Thr His

50 55 60

His Lys Asp Asp Pro Pro Pro Pro Glu Asp Asp Glu Asn Lys Glu Lys
65 70 75 80
Arg Thr Asp Asp Ile Pro Val Trp Asp Gln Glu Phe Leu Lys Val Asp
85 90 95
Gln Gly Thr Leu Phe Glu Leu Ile Leu Ala Ala Asn Tyr Leu Asp Ile
100 105 110
Lys Gly Leu Leu Asp Val Thr Cys Lys Thr Val Ala Asn Met Ile Lys
115 120 125

Gly Lys Thr Pro Glu Glu Ile Arg Lys Thr Phe Asn Ile Lys Asn Asp
130 135 140
Phe Thr Glu Glu Glu Glu Ala Gln Val Arg Lys Glu Asn Gln Trp Cys
145 150 155 160
Glu Glu Lys

<210> 770
<211> 161
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
Spy sequence"
<400> 770
Met Arg Lys Leu Thr Ala Leu Phe Val Ala Ser Thr Leu Ala Leu Gly
1 5 10 15

Ala Ala Asn Leu Ala His Ala Ala Asp Thr Thr Thr Ala Ala Pro Ala
20 25 30
Asp Ala Lys Pro Met Met His His Lys Gly Lys Phe Gly Pro His Gln
35 40 45
Asp Met Met Phe Lys Asp Leu Asn Leu Thr Asp Ala Gln Lys Gln Gln
50 55 60
Ile Arg Glu Ile Met Lys Gly Gln Arg Asp Gln Met Lys Arg Pro Pro

65 70 75 80

Leu Glu Glu Arg Arg Ala Met His Asp Ile Ile Ala Ser Asp Thr Phe
85 90 95
Asp Lys Ala Lys Ala Glu Ala Gln Ile Ala Lys Met Glu Glu Gln Arg
100 105 110
Lys Ala Asn Met Leu Ala His Met Glu Thr Gln Asn Lys Ile Tyr Asn
115 120 125
Ile Leu Thr Pro Glu Gln Lys Lys Gln Phe Asn Ala Asn Phe Glu Lys
130 135 140

Arg Leu Thr Glu Arg Pro Ala Ala Lys Gly Lys Met Pro Ala Thr Ala
145 150 155 160
Glu

<210> 771
<211> 270
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
Fkpa sequence"
<400> 771
Met Lys Ser Leu Phe Lys Val Thr Leu Leu Ala Thr Thr Met Ala Val
1 5 10 15
Ala Leu His Ala Pro Ile Thr Phe Ala Ala Glu Ala Ala Lys Pro Ala
20 25 30

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

85 90 95

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

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

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

<210> 772
<211> 428
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
Sura sequence"

<400> 772

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

Thr Leu Arg His Gln Ile Met Glu Arg Leu Ile Met Asp Gln Ile Ile
65                  70                  75                  80
Leu Gln Met Gly Gln Lys Met Gly Val Lys Ile Ser Asp Glu Gln Leu
                85                  90                  95
Asp Gln Ala Ile Ala Asn Ile Ala Lys Gln Asn Asn Met Thr Leu Asp
            100                 105                 110
Gln Met Arg Ser Arg Leu Ala Tyr Asp Gly Leu Asn Tyr Asn Thr Tyr
        115                 120                 125

Arg Asn Gln Ile Arg Lys Glu Met Ile Ile Ser Glu Val Arg Asn Asn
    130                 135                 140
Glu Val Arg Arg Arg Ile Thr Ile Leu Pro Gln Glu Val Glu Ser Leu
145                 150                 155                 160
Ala Gln Gln Val Gly Asn Gln Asn Asp Ala Ser Thr Glu Leu Asn Leu
                165                 170                 175
Ser His Ile Leu Ile Pro Leu Pro Glu Asn Pro Thr Ser Asp Gln Val
            180                 185                 190

Asn Glu Ala Glu Ser Gln Ala Arg Ala Ile Val Asp Gln Ala Arg Asn
        195                 200                 205
Gly Ala Asp Phe Gly Lys Leu Ala Ile Ala His Ser Ala Asp Gln Gln
    210                 215                 220
Ala Leu Asn Gly Gly Gln Met Gly Trp Gly Arg Ile Gln Glu Leu Pro
225                 230                 235                 240
Gly Ile Phe Ala Gln Ala Leu Ser Thr Ala Lys Lys Gly Asp Ile Val
```

245 250 255

Gly Pro Ile Arg Ser Gly Val Gly Phe His Ile Leu Lys Val Asn Asp
260 265 270
Leu Arg Gly Glu Ser Lys Asn Ile Ser Val Thr Glu Val His Ala Arg
275 280 285
His Ile Leu Leu Lys Pro Ser Pro Ile Met Thr Asp Glu Gln Ala Arg
290 295 300
Val Lys Leu Glu Gln Ile Ala Ala Asp Ile Lys Ser Gly Lys Thr Thr
305 310 315 320

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

Asn Val Asp Lys Thr Asp Ala Ala Gln Lys Asp Arg Ala Tyr Arg Met
385 390 395 400
Leu Met Asn Arg Lys Phe Ser Glu Glu Ala Ala Ser Trp Met Gln Glu
405 410 415
Gln Arg Ala Ser Ala Tyr Val Lys Ile Leu Ser Asn
420 425

<210> 773
<211> 566
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
Hsp60 sequence"
<400> 773

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

Ile Thr Lys Asp Gly Val Thr Val Ala Lys Ala Ile Thr Leu Glu Asp
65                  70                  75                  80
Lys Phe Glu Asp Leu Gly Ala Lys Leu Leu Gln Glu Val Ala Ser Lys
                85                  90                  95
Thr Asn Glu Ser Ala Gly Asp Gly Thr Thr Ser Ala Thr Val Leu Gly
            100                 105                 110
Arg Ser Ile Phe Thr Glu Ser Val Lys Asn Val Ala Ala Gly Cys Asn
        115                 120                 125

Pro Met Asp Leu Arg Arg Gly Ser Gln Ala Ala Val Glu Ala Val Ile
    130                 135                 140
Glu Phe Leu Gln Lys Asn Lys Lys Glu Ile Thr Thr Ser Glu Glu Ile
145                 150                 155                 160
Ala Gln Val Ala Thr Ile Ser Ala Asn Gly Asp Lys His Ile Gly Asp
                165                 170                 175
Leu Leu Ala Asn Ala Met Glu Lys Val Gly Lys Glu Gly Val Ile Thr
            180                 185                 190

Val Lys Glu Gly Lys Thr Leu Glu Asp Glu Leu Glu Val Thr Glu Gly
        195                 200                 205
Met Lys Phe Asp Arg Gly Phe Ile Ser Pro Tyr Phe Ile Thr Asn Thr
    210                 215                 220
Lys Thr Gly Lys Val Glu Phe Glu Asn Pro Leu Ile Leu Leu Ser Glu
225                 230                 235                 240
Lys Lys Ile Ser Ser Ile Gln Asp Ile Leu Pro Ser Leu Glu Leu Ser
```

245 250 255

Asn Gln Thr Arg Arg Pro Leu Leu Ile Ile Ala Glu Asp Val Asp Gly
260 265 270
Glu Ala Leu Ala Ala Cys Ile Leu Asn Lys Leu Arg Gly Gln Val Gln
275 280 285
Val Cys Ala Val Lys Ala Pro Gly Phe Gly Asp Asn Arg Lys Asn Thr
290 295 300
Leu Gly Asp Ile Ala Ile Leu Ser Gly Gly Thr Val Phe Thr Glu Glu
305 310 315 320

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

Arg Leu Ala Lys Leu Ser Gly Gly Val Ala Val Ile Lys Val Gly Gly
385 390 395 400
Ala Ser Glu Val Glu Val Gly Glu Lys Lys Asp Arg Tyr Glu Asp Ala
405 410 415
Leu Asn Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Leu Pro Gly Gly
420 425 430
Gly Thr Ala Leu Ile Lys Ala Thr Lys Ile Leu Asp Glu Val Lys Glu
435 440 445

Lys Ala Val Asn Phe Asp Gln Lys Leu Gly Val Asp Thr Ile Arg Ala
450 455 460
Ala Ile Thr Lys Pro Ala Lys Arg Ile Ile Glu Asn Ala Gly Glu Glu
465 470 475 480
Gly Ala Val Ile Val Gly Lys Ile Tyr Asp Glu Pro Glu Phe Asn Lys
485 490 495

Gly Tyr Asp Ser Gln Lys Gly Glu Phe Thr Asp Met Ile Ala Ala Gly
500 505 510

Ile Ile Asp Pro Phe Lys Val Val Lys Asn Gly Leu Val Asp Ala Ser
515 520 525
Gly Val Ala Ser Leu Leu Ala Thr Thr Glu Cys Ala Ile Val Asp Ala
530 535 540
Pro Gln Pro Lys Gly Ser Pro Ala Ala Pro Pro Ala Pro Gly Met Gly
545 550 555 560
Gly Met Pro Gly Met Phe
565

<210> 774
<211> 705
<212> PRT
<213> Unknown
<220><221>
source
<223> /note="Description of Unknown:
Hsp70 sequence"
<400> 774
Met Ala Pro Ala Val Gly Ile Asp Leu Gly Thr Thr Tyr Ser Cys Val
1 5 10 15
Gly Ile Phe Arg Asp Asp Arg Ile Glu Ile Ile Ala Asn Asp Gln Gly
20 25 30
Asn Arg Thr Thr Pro Ser Phe Val Ala Phe Thr Asp Thr Glu Arg Leu
35 40 45
Ile Gly Asp Ala Ala Lys Asn Gln Val Ala Met Asn Pro Ala Asn Thr

50 55 60
Val Phe Asp Ala Lys Arg Leu Ile Gly Arg Lys Phe Ala Asp Pro Glu
65 70 75 80
Val Gln Ala Asp Met Lys His Phe Pro Phe Lys Ile Thr Asp Lys Gly
85 90 95
Gly Lys Pro Val Ile Gln Val Glu Phe Lys Gly Glu Thr Lys Glu Phe
100 105 110

Thr Pro Glu Glu Ile Ser Ser Met Val Leu Thr Lys Met Arg Glu Thr

115 120 125

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

130 135 140

Pro Ala Tyr Phe Asn Asp Ser Gln Arg Gln Ala Thr Lys Asp Ala Gly

145 150 155 160

Leu Ile Ala Gly Leu Asn Val Leu Arg Ile Ile Asn Glu Pro Thr Ala

165 170 175

Ala Ala Ile Ala Tyr Gly Leu Asp Lys Lys Ala Asp Gly Glu Arg Asn

180 185 190

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

195 200 205

Thr Ile Glu Glu Gly Ile Phe Glu Val Lys Ser Thr Ala Gly Asp Thr

210 215 220

His Leu Gly Gly Glu Asp Phe Asp Asn Arg Leu Val Asn His Phe Val

225 230 235 240

Ser Glu Phe Lys Arg Lys Phe Lys Lys Ile Ser Pro Ala Glu Arg Ala

245 250 255

Arg Ala Leu Arg Arg Ser Pro Thr Ala Cys Glu Arg Ala Lys Arg Thr

260 265 270

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

275 280 285

Gly Ile Asp Phe Tyr Thr Ser Ile Thr Arg Ala Arg Phe Glu Glu Leu

290 295 300

Cys Gln Asp Leu Phe Arg Ser Thr Met Glu Pro Val Glu Arg Val Leu

305 310 315 320

Arg Asp Ala Lys Ile Asp Lys Ser Ser Val His Glu Ile Val Leu Val

325 330 335

Gly Gly Ser Thr Arg Ile Pro Arg Ile Gln Lys Leu Val Ser Asp Phe

340 345 350

Phe Asn Gly Lys Glu Pro Asn Lys Ser Ile Asn Pro Asp Glu Ala Val

355 360 365

Ala Tyr Gly Ala Ala Val Gln Ala Ala Ile Leu Ser Gly Asp Thr Ser

370 375 380

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

385 390 395 400

Leu Gly Ile Glu Thr Ala Gly Gly Val Met Thr Pro Leu Ile Lys Arg

405 410 415

Asn Thr Thr Ile Pro Thr Lys Lys Ser Glu Thr Phe Ser Thr Phe Ser

420 425 430

Asp Asn Gln Pro Gly Val Leu Ile Gln Val Phe Glu Gly Glu Arg Ala

435 440 445

Arg Thr Lys Asp Asn Asn Leu Leu Gly Lys Phe Glu Leu Thr Gly Ile

450 455 460

Pro Arg Ala Arg Gly Val Pro Gln Ile Glu Val Thr Phe Asp Val Asp

465 470 475 480

Ala Asn Gly Ile Met Asn Val Ser Ala Leu Glu Lys Gly Thr Arg Lys

485 490 495

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

500 505 510

Glu Ile Glu Arg Met Leu Ala Glu Ala Glu Lys Tyr Lys Ala Glu Asp

515 520 525

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

530 535 540

Ala Tyr Ser Leu Arg Asn Ser Leu Arg His Ser Lys Val Asp Glu Lys

545 550 555 560

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

565 570 575

Val Gln Trp Leu Asp Glu Asn Gln Thr Ala Thr Lys Glu Glu Tyr Glu

580 585 590

Ser Gln Gln Lys Glu Leu Glu Ala Val Ala Asn Pro Ile Met Met Lys

595 600 605

Phe Tyr Ala Gly Gly Glu Gly Ala Pro Gly Gly Phe Pro Gly Ala Gly

610 615 620

Gly Pro Gly Gly Phe Pro Gly Gly Pro Gly Ala Gly His Ala Ser Gly

625 630 635 640

Gly Gly Asp Asp Gly Pro Thr Val Glu Glu Val Asp Leu Lys Phe Pro

645 650 655

Met Leu Pro Leu Pro Trp Gln Leu Ser Val Arg Lys Met His Arg Pro

660 665 670

Phe Phe Leu Phe Leu Leu Phe Leu Ile Phe Leu Ile Phe Leu Ile Leu

675 680 685

Phe Leu Phe Tyr Phe Phe Leu Pro Val Arg Phe Asn Glu Ser Cys Phe

690 695 700

Ser

705

<210> 775

<211> 548

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GroEL sequence"

<400> 775

Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala Arg Val Lys Met

1 5 10 15

Leu Arg Gly Val Asn Val Leu Ala Asp Ala Val Lys Val Thr Leu Gly

20 25 30

Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe Gly Ala Pro Thr

35 40 45

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

50 55 60

Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Ser Lys

65 70 75 80

```
Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala
                85                  90                  95
Gln Ala Ile Ile Thr Glu Gly Leu Lys Ala Val Ala Ala Gly Met Asn

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

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

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

    290                 295                 300
Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp Leu Gly Gln Ala
305                 310                 315                 320
Lys Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile Ile Asp Gly Val
```

325 330 335

Gly Glu Glu Ala Ala Ile Gln Gly Arg Val Ala Gln Ile Arg Gln Gln

340 345 350

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

355 360 365

Val Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala

370 375 380

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

385 390 395 400

His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val Ala Gly Gly Gly

405 410 415

Val Ala Leu Ile Arg Val Ala Ser Lys Leu Ala Asp Leu Arg Gly Gln

420 425 430

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

435 440 445

Ala Pro Leu Arg Gln Ile Val Leu Asn Cys Gly Glu Glu Pro Ser Val

450 455 460

Val Ala Asn Thr Val Lys Gly Gly Asp Gly Asn Tyr Gly Tyr Asn Ala

465 470 475 480

Ala Thr Glu Glu Tyr Gly Asn Met Ile Asp Met Gly Ile Leu Asp Pro

485 490 495

Thr Lys Val Thr Arg Ser Ala Leu Gln Tyr Ala Ala Ser Val Ala Gly

500 505 510

Leu Met Ile Thr Thr Glu Cys Met Val Thr Asp Leu Pro Lys Asn Asp

515 520 525

Ala Ala Asp Leu Gly Ala Ala Gly Gly Met Gly Gly Met Gly Gly Met

530 535 540

Gly Gly Met Met

545

<210> 776

<211> 97

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

GroES sequence"

<400> 776

Met Asn Ile Arg Pro Leu His Asp Arg Val Ile Val Lys Arg Lys Glu
1 5 10 15

Val Glu Thr Lys Ser Ala Gly Gly Ile Val Leu Thr Gly Ser Ala Ala
20 25 30

Ala Lys Ser Thr Arg Gly Glu Val Leu Ala Val Gly Asn Gly Arg Ile
35 40 45

Leu Glu Asn Gly Glu Val Lys Pro Leu Asp Val Lys Val Gly Asp Ile
50 55 60

Val Ile Phe Asn Asp Gly Tyr Gly Val Lys Ser Glu Lys Ile Asp Asn
65 70 75 80

Glu Glu Val Leu Ile Met Ser Glu Ser Asp Ile Leu Ala Ile Val Glu
85 90 95

Ala

<210> 777

<211> 732

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hsp90-alpha sequence"

<400> 777

Met Pro Glu Glu Thr Gln Thr Gln Asp Gln Pro Met Glu Glu Glu Glu
1 5 10 15

Val Glu Thr Phe Ala Phe Gln Ala Glu Ile Ala Gln Leu Met Ser Leu
20 25 30

Ile Ile Asn Thr Phe Tyr Ser Asn Lys Glu Ile Phe Leu Arg Glu Leu

35 40 45

Ile Ser Asn Ser Ser Asp Ala Leu Asp Lys Ile Arg Tyr Glu Ser Leu

50 55 60

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

65 70 75 80

Ile Pro Asn Lys Gln Asp Arg Thr Leu Thr Ile Val Asp Thr Gly Ile

85 90 95

Gly Met Thr Lys Ala Asp Leu Ile Asn Asn Leu Gly Thr Ile Ala Lys

100 105 110

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

115 120 125

Ser Met Ile Gly Gln Phe Gly Val Gly Phe Tyr Ser Ala Tyr Leu Val

130 135 140

Ala Glu Lys Val Thr Val Ile Thr Lys His Asn Asp Asp Glu Gln Tyr

145 150 155 160

Ala Trp Glu Ser Ser Ala Gly Gly Ser Phe Thr Val Arg Thr Asp Thr

165 170 175

Gly Glu Pro Met Gly Arg Gly Thr Lys Val Ile Leu His Leu Lys Glu

180 185 190

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

195 200 205

Lys His Ser Gln Phe Ile Gly Tyr Pro Ile Thr Leu Phe Val Glu Lys

210 215 220

Glu Arg Asp Lys Glu Val Ser Asp Asp Glu Ala Glu Glu Lys Glu Asp

225 230 235 240

Lys Glu Glu Glu Lys Glu Lys Glu Glu Lys Glu Ser Glu Asp Lys Pro

245 250 255

Glu Ile Glu Asp Val Gly Ser Asp Glu Glu Glu Glu Lys Lys Asp Gly

260 265 270

Asp Lys Lys Lys Lys Lys Lys Ile Lys Glu Lys Tyr Ile Asp Gln Glu

275 280 285

Glu Leu Asn Lys Thr Lys Pro Ile Trp Thr Arg Asn Pro Asp Asp Ile
290 295 300
Thr Asn Glu Glu Tyr Gly Glu Phe Tyr Lys Ser Leu Thr Asn Asp Trp
305 310 315 320

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

Ile Arg Gly Val Val Asp Ser Glu Asp Leu Pro Leu Asn Ile Ser Arg
385 390 395 400
Glu Met Leu Gln Gln Ser Lys Ile Leu Lys Val Ile Arg Lys Asn Leu
405 410 415
Val Lys Lys Cys Leu Glu Leu Phe Thr Glu Leu Ala Glu Asp Lys Glu
420 425 430
Asn Tyr Lys Lys Phe Tyr Glu Gln Phe Ser Lys Asn Ile Lys Leu Gly
435 440 445

Ile His Glu Asp Ser Gln Asn Arg Lys Lys Leu Ser Glu Leu Leu Arg
450 455 460
Tyr Tyr Thr Ser Ala Ser Gly Asp Glu Met Val Ser Leu Lys Asp Tyr
465 470 475 480
Cys Thr Arg Met Lys Glu Asn Gln Lys His Ile Tyr Tyr Ile Thr Gly
485 490 495
Glu Thr Lys Asp Gln Val Ala Asn Ser Ala Phe Val Glu Arg Leu Arg
500 505 510

Lys His Gly Leu Glu Val Ile Tyr Met Ile Glu Pro Ile Asp Glu Tyr
515 520 525
Cys Val Gln Gln Leu Lys Glu Phe Glu Gly Lys Thr Leu Val Ser Val

530 535 540

Thr Lys Glu Gly Leu Glu Leu Pro Glu Asp Glu Glu Glu Lys Lys Lys
545 550 555 560

Gln Glu Glu Lys Lys Thr Lys Phe Glu Asn Leu Cys Lys Ile Met Lys
565 570 575

Asp Ile Leu Glu Lys Lys Val Glu Lys Val Val Val Ser Asn Arg Leu
580 585 590

Val Thr Ser Pro Cys Cys Ile Val Thr Ser Thr Tyr Gly Trp Thr Ala
595 600 605

Asn Met Glu Arg Ile Met Lys Ala Gln Ala Leu Arg Asp Asn Ser Thr
610 615 620

Met Gly Tyr Met Ala Ala Lys Lys His Leu Glu Ile Asn Pro Asp His
625 630 635 640

Ser Ile Ile Glu Thr Leu Arg Gln Lys Ala Glu Ala Asp Lys Asn Asp
645 650 655

Lys Ser Val Lys Asp Leu Val Ile Leu Leu Tyr Glu Thr Ala Leu Leu
660 665 670

Ser Ser Gly Phe Ser Leu Glu Asp Pro Gln Thr His Ala Asn Arg Ile
675 680 685

Tyr Arg Met Ile Lys Leu Gly Leu Gly Ile Asp Glu Asp Asp Pro Thr
690 695 700

Ala Asp Asp Thr Ser Ala Ala Val Thr Glu Glu Met Pro Pro Leu Glu
705 710 715 720

Gly Asp Asp Asp Thr Ser Arg Met Glu Glu Val Asp
725 730

<210> 778

<211> 624

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:
HtpG sequence"

<400> 778

Met Lys Gly Gln Glu Thr Arg Gly Phe Gln Ser Glu Val Lys Gln Leu
1 5 10 15

Leu His Leu Met Ile His Ser Leu Tyr Ser Asn Lys Glu Ile Phe Leu
20 25 30

Arg Glu Leu Ile Ser Asn Ala Ser Asp Ala Ala Asp Lys Leu Arg Phe
35 40 45

Arg Ala Leu Ser Asn Pro Asp Leu Tyr Glu Gly Asp Gly Glu Leu Arg
50 55 60

Val Arg Val Ser Phe Asp Lys Asp Lys Arg Thr Leu Thr Ile Ser Asp
65 70 75 80

Asn Gly Val Gly Met Thr Arg Asp Glu Val Ile Asp His Leu Gly Thr
85 90 95

Ile Ala Lys Ser Gly Thr Lys Ser Phe Leu Glu Ser Leu Gly Ser Asp
100 105 110

Gln Ala Lys Asp Ser Gln Leu Ile Gly Gln Phe Gly Val Gly Phe Tyr
115 120 125

Ser Ala Phe Ile Val Ala Asp Lys Val Thr Val Arg Thr Arg Ala Ala
130 135 140

Gly Glu Lys Pro Glu Asn Gly Val Phe Trp Glu Ser Ala Gly Glu Gly
145 150 155 160

Glu Tyr Thr Val Ala Asp Ile Thr Lys Glu Asp Arg Gly Thr Glu Ile
165 170 175

Thr Leu His Leu Arg Glu Gly Glu Asp Glu Phe Leu Asp Asp Trp Arg
180 185 190

Val Arg Ser Ile Ile Ser Lys Tyr Ser Asp His Ile Ala Leu Pro Val
195 200 205

Glu Ile Glu Lys Arg Glu Glu Lys Asp Gly Glu Thr Val Ile Ser Trp
210 215 220

Glu Lys Ile Asn Lys Ala Gln Ala Leu Trp Thr Arg Asn Lys Ser Glu
225 230 235 240

Ile Thr Asp Glu Glu Tyr Lys Glu Phe Tyr Lys His Ile Ala His Asp
245 250 255
Phe Asn Asp Pro Leu Thr Trp Ser His Asn Arg Val Glu Gly Lys Gln
260 265 270
Glu Tyr Thr Ser Leu Leu Tyr Ile Pro Ser Gln Ala Pro Trp Asp Met

275 280 285
Trp Asn Arg Asp His Lys His Gly Leu Lys Leu Tyr Val Gln Arg Val
290 295 300
Phe Ile Met Asp Asp Ala Glu Gln Phe Met Pro Asn Tyr Leu Arg Phe
305 310 315 320
Val Arg Gly Leu Ile Asp Ser Ser Asp Leu Pro Leu Asn Val Ser Arg
325 330 335
Glu Ile Leu Gln Asp Ser Thr Val Thr Arg Asn Leu Arg Asn Ala Leu

340 345 350
Thr Lys Arg Val Leu Gln Met Leu Glu Lys Leu Ala Lys Asp Asp Ala
355 360 365
Glu Lys Tyr Gln Thr Phe Trp Gln Gln Phe Gly Leu Val Leu Lys Glu
370 375 380
Gly Pro Ala Glu Asp Phe Ala Asn Gln Glu Ala Ile Ala Lys Leu Leu
385 390 395 400
Arg Phe Ala Ser Thr His Thr Asp Ser Ser Ala Gln Thr Val Ser Leu

405 410 415
Glu Asp Tyr Val Ser Arg Met Lys Glu Gly Gln Glu Lys Ile Tyr Tyr
420 425 430
Ile Thr Ala Asp Ser Tyr Ala Ala Ala Lys Ser Ser Pro His Leu Glu
435 440 445
Leu Leu Arg Lys Lys Gly Ile Glu Val Leu Leu Leu Ser Asp Arg Ile
450 455 460
Asp Glu Trp Met Met Asn Tyr Leu Thr Glu Phe Asp Gly Lys Pro Phe

465 470 475 480
Gln Ser Val Ser Lys Val Asp Glu Ser Leu Glu Lys Leu Ala Asp Glu

485 490 495

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

500 505 510

Asp Arg Val Lys Ala Leu Leu Gly Glu Arg Val Lys Asp Val Arg Leu

515 520 525

Thr His Arg Leu Thr Asp Thr Pro Ala Ile Val Ser Thr Asp Ala Asp

530 535 540

Glu Met Ser Thr Gln Met Ala Lys Leu Phe Ala Ala Ala Gly Gln Lys

545 550 555 560

Val Pro Glu Val Lys Tyr Ile Phe Glu Leu Asn Pro Asp His Val Leu

565 570 575

Val Lys Arg Ala Ala Asp Thr Glu Asp Glu Ala Lys Phe Ser Glu Trp

580 585 590

Val Glu Leu Leu Leu Asp Gln Ala Leu Leu Ala Glu Arg Gly Thr Leu

595 600 605

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

610 615 620

<210> 779

<211> 340

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

Hsp40 sequence"

<400> 779

Met Gly Lys Asp Tyr Tyr Gln Thr Leu Gly Leu Ala Arg Gly Ala Ser

1 5 10 15

Asp Glu Glu Ile Lys Arg Ala Tyr Arg Arg Gln Ala Leu Arg Tyr His

20 25 30

Pro Asp Lys Asn Lys Glu Pro Gly Ala Glu Glu Lys Phe Lys Glu Ile

35 40 45

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

50 55 60

Asp Arg Tyr Gly Glu Glu Gly Leu Lys Gly Ser Gly Pro Ser Gly Gly

65 70 75 80

Ser Gly Gly Gly Ala Asn Gly Thr Ser Phe Ser Tyr Thr Phe His Gly

85 90 95

Asp Pro His Ala Met Phe Ala Glu Phe Phe Gly Gly Arg Asn Pro Phe

100 105 110

Asp Thr Phe Phe Gly Gln Arg Asn Gly Glu Glu Gly Met Asp Ile Asp

115 120 125

Asp Pro Phe Ser Gly Phe Pro Met Gly Met Gly Gly Phe Thr Asn Val

130 135 140

Asn Phe Gly Arg Ser Arg Ser Ala Gln Glu Pro Ala Arg Lys Lys Gln

145 150 155 160

Asp Pro Pro Val Thr His Asp Leu Arg Val Ser Leu Glu Glu Ile Tyr

165 170 175

Ser Gly Cys Thr Lys Lys Met Lys Ile Ser His Lys Arg Leu Asn Pro

180 185 190

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

195 200 205

Lys Lys Gly Trp Lys Glu Gly Thr Lys Ile Thr Phe Pro Lys Glu Gly

210 215 220

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

225 230 235 240

Asp Lys Pro His Asn Ile Phe Lys Arg Asp Gly Ser Asp Val Ile Tyr

245 250 255

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

260 265 270

Val Pro Thr Leu Asp Gly Arg Thr Ile Pro Val Val Phe Lys Asp Val

275 280 285

Ile Arg Pro Gly Met Arg Arg Lys Val Pro Gly Glu Gly Leu Pro Leu

290 295 300

Pro Lys Thr Pro Glu Lys Arg Gly Asp Leu Ile Ile Glu Phe Glu Val
305 310 315 320
Ile Phe Pro Glu Arg Ile Pro Gln Thr Ser Arg Thr Val Leu Glu Gln
325 330 335
Val Leu Pro Ile
340

<210> 780

<211> 277
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
ClpP sequence"
<400> 780
Met Trp Pro Gly Ile Leu Val Gly Gly Ala Arg Val Ala Ser Cys Arg
1 5 10 15
Tyr Pro Ala Leu Gly Pro Arg Leu Ala Ala His Phe Pro Ala Gln Arg
20 25 30
Pro Pro Gln Arg Thr Leu Gln Asn Gly Leu Ala Leu Gln Arg Cys Leu
35 40 45

His Ala Thr Ala Thr Arg Ala Leu Pro Leu Ile Pro Ile Val Val Glu
50 55 60
Gln Thr Gly Arg Gly Glu Arg Ala Tyr Asp Ile Tyr Ser Arg Leu Leu
65 70 75 80
Arg Glu Arg Ile Val Cys Val Met Gly Pro Ile Asp Asp Ser Val Ala
85 90 95
Ser Leu Val Ile Ala Gln Leu Leu Phe Leu Gln Ser Glu Ser Asn Lys
100 105 110

Lys Pro Ile His Met Tyr Ile Asn Ser Pro Gly Gly Val Val Thr Ala
115 120 125
Gly Leu Ala Ile Tyr Asp Thr Met Gln Tyr Ile Leu Asn Pro Ile Cys
130 135 140

Thr Trp Cys Val Gly Gln Ala Ala Ser Met Gly Ser Leu Leu Leu Ala
145 150 155 160
Ala Gly Thr Pro Gly Met Arg His Ser Leu Pro Asn Ser Arg Ile Met
165 170 175

Ile His Gln Pro Ser Gly Gly Ala Arg Gly Gln Ala Thr Asp Ile Ala
180 185 190
Ile Gln Ala Glu Glu Ile Met Lys Leu Lys Lys Gln Leu Tyr Asn Ile
195 200 205
Tyr Ala Lys His Thr Lys Gln Ser Leu Gln Val Ile Glu Ser Ala Met
210 215 220
Glu Arg Asp Arg Tyr Met Ser Pro Met Glu Ala Gln Glu Phe Gly Ile
225 230 235 240

Leu Asp Lys Val Leu Val His Pro Pro Gln Asp Gly Glu Asp Glu Pro
245 250 255
Thr Leu Val Gln Lys Glu Pro Val Glu Ala Ala Pro Ala Ala Glu Pro
260 265 270
Val Pro Ala Ser Thr
275

<210> 781
<211> 633
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:
ClpX sequence"
<400> 781
Met Pro Ser Cys Gly Ala Cys Thr Cys Gly Ala Ala Ala Val Arg Leu

1 5 10 15
Ile Thr Ser Ser Leu Ala Ser Ala Gln Arg Gly Ile Ser Gly Gly Arg
20 25 30
Ile His Met Ser Val Leu Gly Arg Leu Gly Thr Phe Glu Thr Gln Ile
35 40 45

```
Leu Gln Arg Ala Pro Leu Arg Ser Phe Thr Glu Thr Pro Ala Tyr Phe
    50                  55                  60
Ala Ser Lys Asp Gly Ile Ser Lys Asp Gly Ser Gly Asp Gly Asn Lys

65                  70                  75                  80
Lys Ser Ala Ser Glu Gly Ser Ser Lys Lys Ser Gly Ser Gly Asn Ser
                85                  90                  95
Gly Lys Gly Gly Asn Gln Leu Arg Cys Pro Lys Cys Gly Asp Leu Cys
            100                 105                 110
Thr His Val Glu Thr Phe Val Ser Ser Thr Arg Phe Val Lys Cys Glu
        115                 120                 125
Lys Cys His His Phe Phe Val Val Leu Ser Glu Ala Asp Ser Lys Lys

    130                 135                 140
Ser Ile Ile Lys Glu Pro Glu Ser Ala Ala Glu Ala Val Lys Leu Ala
145                 150                 155                 160
Phe Gln Gln Lys Pro Pro Pro Pro Pro Lys Lys Ile Tyr Asn Tyr Leu
                165                 170                 175
Asp Lys Tyr Val Val Gly Gln Ser Phe Ala Lys Lys Val Leu Ser Val
            180                 185                 190
Ala Val Tyr Asn His Tyr Lys Arg Ile Tyr Asn Asn Ile Pro Ala Asn

        195                 200                 205
Leu Arg Gln Gln Ala Glu Val Glu Lys Gln Thr Ser Leu Thr Pro Arg
    210                 215                 220
Glu Leu Glu Ile Arg Arg Arg Glu Asp Glu Tyr Arg Phe Thr Lys Leu
225                 230                 235                 240
Leu Gln Ile Ala Gly Ile Ser Pro His Gly Asn Ala Leu Gly Ala Ser
                245                 250                 255
Met Gln Gln Gln Val Asn Gln Gln Ile Pro Gln Glu Lys Arg Gly Gly

            260                 265                 270
Glu Val Leu Asp Ser Ser His Asp Asp Ile Lys Leu Glu Lys Ser Asn
        275                 280                 285
Ile Leu Leu Leu Gly Pro Thr Gly Ser Gly Lys Thr Leu Leu Ala Gln
```

290 295 300

Thr Leu Ala Lys Cys Leu Asp Val Pro Phe Ala Ile Cys Asp Cys Thr

305 310 315 320

Thr Leu Thr Gln Ala Gly Tyr Val Gly Glu Asp Ile Glu Ser Val Ile

325 330 335

Ala Lys Leu Leu Gln Asp Ala Asn Tyr Asn Val Glu Lys Ala Gln Gln

340 345 350

Gly Ile Val Phe Leu Asp Glu Val Asp Lys Ile Gly Ser Val Pro Gly

355 360 365

Ile His Gln Leu Arg Asp Val Gly Gly Glu Gly Val Gln Gln Gly Leu

370 375 380

Leu Lys Leu Leu Glu Gly Thr Ile Val Asn Val Pro Glu Lys Asn Ser

385 390 395 400

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

405 410 415

Phe Val Ala Ser Gly Ala Phe Asn Gly Leu Asp Arg Ile Ile Ser Arg

420 425 430

Arg Lys Asn Glu Lys Tyr Leu Gly Phe Gly Thr Pro Ser Asn Leu Gly

435 440 445

Lys Gly Arg Arg Ala Ala Ala Ala Ala Asp Leu Ala Asn Arg Ser Gly

450 455 460

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

465 470 475 480

His Val Glu Ala Arg Asp Leu Ile Glu Phe Gly Met Ile Pro Glu Phe

485 490 495

Val Gly Arg Leu Pro Val Val Val Pro Leu His Ser Leu Asp Glu Lys

500 505 510

Thr Leu Val Gln Ile Leu Thr Glu Pro Arg Asn Ala Val Ile Pro Gln

515 520 525

Tyr Gln Ala Leu Phe Ser Met Asp Lys Cys Glu Leu Asn Val Thr Glu

530 535 540

Asp Ala Leu Lys Ala Ile Ala Arg Leu Ala Leu Glu Arg Lys Thr Gly
545 550 555 560
Ala Arg Gly Leu Arg Ser Ile Met Glu Lys Leu Leu Leu Glu Pro Met
565 570 575
Phe Glu Val Pro Asn Ser Asp Ile Val Cys Val Glu Val Asp Lys Glu

580 585 590
Val Val Glu Gly Lys Lys Glu Pro Gly Tyr Ile Arg Ala Pro Thr Lys
595 600 605
Glu Ser Ser Glu Glu Glu Tyr Asp Ser Gly Val Glu Glu Glu Gly Trp
610 615 620
Pro Arg Gln Ala Asp Ala Ala Asn Ser
625 630
<210> 782
<211> 316
<212> PRT
<213> Unknown
<220><221> source
<223> /note="Description of Unknown:

CDKN1C sequence"
<400> 782
Met Ser Asp Ala Ser Leu Arg Ser Thr Ser Thr Met Glu Arg Leu Val
1 5 10 15
Ala Arg Gly Thr Phe Pro Val Leu Val Arg Thr Ser Ala Cys Arg Ser
20 25 30
Leu Phe Gly Pro Val Asp His Glu Glu Leu Ser Arg Glu Leu Gln Ala
35 40 45
Arg Leu Ala Glu Leu Asn Ala Glu Asp Gln Asn Arg Trp Asp Tyr Asp
50 55 60

Phe Gln Gln Asp Met Pro Leu Arg Gly Pro Gly Arg Leu Gln Trp Thr
65 70 75 80
Glu Val Asp Ser Asp Ser Val Pro Ala Phe Tyr Arg Glu Thr Val Gln
85 90 95

Val Gly Arg Cys Arg Leu Leu Leu Ala Pro Arg Pro Val Ala Val Ala
100 105 110
Val Ala Val Ser Pro Pro Leu Glu Pro Ala Ala Glu Ser Leu Asp Gly
115 120 125

Leu Glu Glu Ala Pro Glu Gln Leu Pro Ser Val Pro Val Pro Ala Pro
130 135 140
Ala Ser Thr Pro Pro Pro Val Pro Val Leu Ala Pro Ala Pro Ala Pro
145 150 155 160
Ala Pro Ala Pro Val Ala Ala Pro Val Ala Ala Pro Val Ala Val Ala
165 170 175
Val Leu Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro
180 185 190

Ala Pro Val Ala Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro
195 200 205
Ala Pro Ala Pro Ala Pro Ala Pro Asp Ala Ala Pro Gln Glu Ser Ala
210 215 220
Glu Gln Gly Ala Asn Gln Gly Gln Arg Gly Gln Glu Pro Leu Ala Asp
225 230 235 240
Gln Leu His Ser Gly Ile Ser Gly Arg Pro Ala Ala Gly Thr Ala Ala
245 250 255

Ala Ser Ala Asn Gly Ala Ala Ile Lys Lys Leu Ser Gly Pro Leu Ile
260 265 270
Ser Asp Phe Phe Ala Lys Arg Lys Arg Ser Ala Pro Glu Lys Ser Ser
275 280 285
Gly Asp Val Pro Ala Pro Cys Pro Ser Pro Ser Ala Ala Pro Gly Val
290 295 300
Gly Ser Val Glu Gln Thr Pro Arg Lys Arg Leu Arg
305 310 315
<210> 783
<
211> 200
<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

BAX inhibitor sequence"

<400> 783

Met Asn Ile Phe Asp Arg Lys Ile Asn Phe Asp Ala Leu Leu Lys Phe
1 5 10 15

Ser His Ile Thr Pro Ser Thr Gln Gln His Leu Lys Lys Val Tyr Ala
20 25 30

Ser Phe Ala Leu Cys Met Phe Val Ala Ala Ala Gly Ala Tyr Val His
35 40 45

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

Leu Ile Leu Met Ile Trp Leu Met Ala Thr Pro His Ser His Glu Thr
65 70 75 80

Glu Gln Lys Arg Leu Gly Leu Leu Ala Gly Phe Ala Phe Leu Thr Gly
85 90 95

Val Gly Leu Gly Pro Ala Leu Glu Phe Cys Ile Ala Val Asn Pro Ser
100 105 110

Ile Leu Pro Thr Ala Phe Met Gly Thr Ala Met Ile Phe Thr Cys Phe
115 120 125

Thr Leu Ser Ala Leu Tyr Ala Arg Arg Arg Ser Tyr Leu Phe Leu Gly
130 135 140

Gly Ile Leu Met Ser Ala Leu Ser Leu Leu Leu Leu Ser Ser Leu Gly
145 150 155 160

Asn Val Phe Phe Gly Ser Ile Trp Leu Phe Gln Ala Asn Leu Tyr Val
165 170 175

Gly Leu Val Val Met Cys Gly Phe Val Leu Phe Asp Thr Gln Leu Ile
180 185 190

Ile Glu Lys Ala Glu His Gly Asp
195 200

<210> 784

<211> 464

<212> PRT

<213> Unknown

<220><221> source

<223> /note="Description of Unknown:

codA sequence"

<400> 784

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

1 5 10 15

Glu Glu Gly Leu Trp Gln Ile His Leu Gln Asp Gly Lys Ile Ser Ala

20 25 30

Ile Asp Ala Gln Ser Gly Val Met Pro Ile Thr Glu Asn Ser Leu Asp

35 40 45

Ala Glu Gln Gly Leu Val Ile Pro Pro Phe Val Glu Pro His Ile His

50 55 60

Leu Asp Thr Thr Gln Thr Ala Gly Gln Pro Asn Trp Asn Gln Ser Gly

65 70 75 80

Thr Leu Phe Glu Gly Ile Glu Arg Trp Ala Glu Arg Lys Ala Leu Leu

85 90 95

Thr His Asp Asp Val Lys Gln Arg Ala Trp Gln Thr Leu Lys Trp Gln

100 105 110

Ile Ala Asn Gly Ile Gln His Val Arg Thr His Val Asp Val Ser Asp

115 120 125

Ala Thr Leu Thr Ala Leu Lys Ala Met Leu Glu Val Lys Gln Glu Val

130 135 140

Ala Pro Trp Ile Asp Leu Gln Ile Val Ala Phe Pro Gln Glu Gly Ile

145 150 155 160

Gln Asp Tyr Ile Trp His Cys Ile Asp Leu Phe Leu Asp Phe Ile Thr

165 170 175

Val Phe Arg Lys Leu Met Met Ile Leu Ala Met Asn Glu Lys Asp Lys

180 185 190

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

195 200 205

Glu Ala Leu Arg Leu Gly Ala Asp Val Val Gly Ala Ile Pro His Phe

210 215 220

Glu Phe Thr Arg Glu Tyr Gly Val Glu Ser Leu His Lys Thr Phe Ala

225 230 235 240

Leu Ala Gln Lys Tyr Asp Arg Leu Ile Asp Val His Cys Asp Glu Ile

245 250 255

Asp Asp Glu Gln Ser Arg Phe Val Glu Thr Val Ala Ala Leu Ala His

260 265 270

His Glu Gly Met Gly Ala Arg Val Thr Ala Ser His Thr Thr Ala Met

275 280 285

His Ser Tyr Asn Gly Ala Tyr Thr Ser Arg Leu Phe Arg Leu Leu Lys

290 295 300

Met Ser Gly Ile Asn Phe Val Ala Asn Pro Leu Val Asn Ile His Leu

305 310 315 320

Gln Gly Arg Phe Asp Thr Tyr Pro Lys Arg Arg Gly Ile Thr Arg Val

325 330 335

Lys Glu Met Leu Glu Ser Gly Ile Asn Val Cys Phe Gly His Asp Asp

340 345 350

Val Phe Asp Pro Trp Tyr Pro Leu Gly Thr Ala Asn Met Leu Gln Val

355 360 365

Leu His Met Gly Leu His Val Cys Gln Leu Met Gly Tyr Gly Gln Ile

370 375 380

Asn Asp Gly Leu Asn Leu Ile Thr His His Ser Ala Arg Thr Leu Asn

385 390 395 400

Leu Gln Asp Tyr Gly Ile Ala Ala Gly Asn Ser Ala Asn Leu Ile Ile

405 410 415

Leu Pro Ala Glu Asn Gly Phe Asp Ala Leu Arg Arg Gln Val Pro Val

420 425 430

Arg Tyr Ser Val Arg Gly Gly Lys Val Ile Ala Ser Thr Gln Pro Ala

435 440 445

Gln Thr Thr Val Tyr Leu Glu Gln Pro Glu Ala Ile Asp Tyr Lys Arg
450 455 460

<210> 785

<211> 42

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 785

Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met
1 5 10 15

Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe
20 25 30

Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
35 40

<210> 786

<211> 112

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 786

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
1 5 10 15

Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
20 25 30

Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
35 40 45

Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
50 55 60

Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
65 70 75 80
Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
85 90 95
Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
100 105 110

<210> 787
<211> 241
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 787
Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Glu Lys Pro Gly Ala
1 5 10 15
Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asp Tyr
20 25 30
Tyr Met His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe
50 55 60
Gln Gly Arg Val Thr Met Thr Arg Asp Thr Ser Ile Ser Thr Ala Tyr
65 70 75 80
Met Glu Leu Ser Arg Leu Arg Ser Asp Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Ala Ser Gly Trp Asp Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr
100 105 110

Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly
115 120 125
Gly Ser Gly Gly Gly Gly Ser Asp Ile Val Met Thr Gln Ser Pro Ser
130 135 140

Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala
145 150 155 160
Ser Gln Ser Ile Arg Tyr Tyr Leu Ser Trp Tyr Gln Gln Lys Pro Gly
165 170 175

Lys Ala Pro Lys Leu Leu Ile Tyr Thr Ala Ser Ile Leu Gln Asn Gly
180 185 190
Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu
195 200 205
Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Leu
210 215 220
Gln Thr Tyr Thr Thr Pro Asp Phe Gly Pro Gly Thr Lys Val Glu Ile
225 230 235 240

Lys

<210> 788
<211> 120
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic
polypeptide"
<400> 788
Glu Val Lys Leu Gln Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln
1 5 10 15
Ser Leu Ser Val Thr Cys Thr Val Ser Gly Val Ser Leu Pro Asp Tyr
20 25 30
Gly Val Ser Trp Ile Arg Gln Pro Pro Arg Lys Gly Leu Glu Trp Leu

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

Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Ile Tyr Tyr Cys Ala
85 90 95
Lys His Tyr Tyr Tyr Gly Gly Ser Tyr Ala Met Asp Tyr Trp Gly Gln

100 105 110
Gly Thr Ser Val Thr Val Ser Ser
115 120

<210> 789
<211> 107
<212> PRT
<213> Artificial Sequence
<220><221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"
<400> 789
Asp Ile Gln Met Thr Gln Thr Thr Ser Ser Leu Ser Ala Ser Leu Gly
1 5 10 15
Asp Arg Val Thr Ile Ser Cys Arg Ala Ser Gln Asp Ile Ser Lys Tyr
20 25 30

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

Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Thr
100 105

<210> 790
<211> 218
<212> PRT
<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 790

Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Lys Gly Val Ser Thr Ser
20 25 30

Gly Tyr Ser Tyr Leu His Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro
35 40 45

Arg Leu Leu Ile Tyr Leu Ala Ser Tyr Leu Glu Ser Gly Val Pro Ala
50 55 60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser
65 70 75 80

Ser Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln His Ser Arg
85 90 95

Asp Leu Pro Leu Thr Phe Gly Gly Gly Thr Lys Val Glu Ile Lys Arg
100 105 110

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
115 120 125

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
130 135 140

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
145 150 155 160

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
165 170 175

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
180 185 190

His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
195 200 205

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
210 215

<210> 791

<211> 80

<212> RNA

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<220>

<221> source

<223> /note="See specification as filed for detailed description of substitutions and preferred embodiments"

<400> 791

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60

aaaaaaaaaa aaaaaaaaaa 80

<210> 792

<211> 5

<212> PRT

<213> Artificial Sequence

<220><221

> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 792

Gly Gly Gly Gly Ser

1 5

<210> 793

<211> 15

<212> PRT

<213> Artificial Sequence

<220><221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<400> 793

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15